The Physical Foundations of Biology and the Problems of Psychophysics*

Alfred Gierer

Max Planck Institute for Developmental Biology
Spemannstr. 35, D- 72076 Tübingen

Abstract:

Full applicability of physics to human biology does not necessarily imply that one can uncover a comprehensive, algorithmic correlation between physical brain states and corresponding mental states. The argument takes into account that information processing is finite in principle in a finite world. Presumably the brain-mind-relation cannot be resolved in all essential aspects, particularly when high degrees of abstraction or self-analytical processes are involved. Our conjecture plausibly unifies the universal validity of physics and a logical limitation of human thought, and it does not regard consciousness - the most basic human experience - as a marginal phenomenon.

Key words: brain / mind / consciousness / human / decodability / decision theory / Gödel / Heisenberg / Frayn

---

(*) Reformated version of the article published in RATIO XII, 1970, No. 1, p 47-64 (keywords and abstract are added to the original text of 1970). RATIO appeared up to 1987 in both a German and an English version.
The German title of this article: Alfred Gierer, Der physikalische Grundlegungsversuch in der Biologie und das psychophysische Problem. RATIO XII, Heft 1, 1970, S. 40-54.

INTRODUCTION

The results of modern biology, in particular of molecular genetics, developmental physiology and neurophysiology, strongly indicate that the laws of physics apply to the entire biological domain, including man and the higher animals.

In the energy range which is essential for biological and chemical processes, physics can be regarded as concluded by quantum mechanics. Within the limits of indeterminacy formulated by Heisenberg's uncertainty principle the fundamental equations of quantum mechanics yield a full description of the observable events in space and time. More particularly, they enable us to calculate in advance the probabilities of all positions and momenta at any future time for any closed system for which the positions and momenta have been measured at a given time within the limits of possible accuracy. In the case of macroscopic objects which consist of many atoms these predictions mostly prove to be very accurate. If physics is applicable to the entire domain of biology including the nervous system of man, then one is led to the conjecture that with its help it should be possible to give a complete explanation of all objectively observable properties of man.

However, this argument is opposed by another which denies the possibility of a complete explanation of man by the natural sciences. It seems plausible that human thinking cannot completely comprehend itself although it is a property of the human nervous system which in turn is subject to physical laws. In the following section I shall discuss some of the logical, sociological and other arguments which can be adduced to support this view. The point of the thesis is that human thinking cannot simultaneously be the means and the object of an exhaustive analysis. The problem of the application of thinking to itself is connected with the well-known contradictions which arise when concepts are applied self-referentially, e.g. in the liar paradox.

We thus have plausible physical arguments for and plausible logical arguments against the view that the properties of man can be completely analysed in scientific terms. In this paper I shall attempt a solution of this question by advancing a hypothesis about the psychophysical problem based on a discussion of physical as well as logical arguments. There are numerous theories about the 'mind-body problem' and I do not wish to imply that certain points made by me in outlining the hypothesis are new. Some of them are involved in every serious discussion of the problem. I shall, therefore, abstain from quoting references, except for a few instances of particularly stimulating work. I shall base my considerations on a finitist approach which takes account of the fundamental finiteness of the number of possible operations. I shall try to show that it may be impossible to establish a complete theory of the psychophysical relation, even though it is assumed that the physical laws apply to the entire biological domain and that there is a one-one correspondence between physical and psychological states.

2. THE LIMITS OF SCIENTIFIC EXPLANATION

If physics applies to all nature including the biological domain, then the limits of objective scientific explanation are essentially determined by three factors: (a) The limits of physics which are given by the indeterminacy principle. (b) The limits of formalizable thinking which are particularly characterized by the decidability problems of mathematical logic. (c) The numerical limitation of the possible physical and mental operations by the finiteness of human possibilities, ultimately by the limited size and duration of the universe.

(a) The limits of physics

The limits of objective physical analysis are given by the indeterminacy principle which is connected with the unavoidable impingement of the measuring process upon the measured object. Quantum theory allows us to make probability statements about the results of future measurements on the basis of results of measurements made at a given time. Atomic and molecular processes are necessarily subject to the indeterminacy principle: chemical and physical processes involving individual molecules and atoms cannot, as a matter of principle, be determined accurately in advance by measurement and calculations; we can only calculate their probability. In the case of macroscopic objects the quantum theoretical indeterminacy does not in general affect matters, since the averaging of values for the numerous atoms and molecules of the object permits very accurate measurement and calculations. However, both the inorganic and the organic domain include macroscopic phenomena which are determined by individual atomic events and thus subject to the indeterminacy principle. This does not only apply to artificial measuring arrangements with amplifier effect, such as e.g. Geiger counters, but also to phenomena connected with formation of nuclei in general, such as occur e.g. in crystallization processes and in turbulence phenomena. Thus it is e.g. to be expected, that the weather-to which such processes are highly relevant-is subject to quantum theoretical indeterminacy at least in the long run.

The organic domain in particular includes processes in which molecular events have macroscopic effects. Thus e.g. mutations are molecular processes which are subject to the principle of indeterminacy. In this context sexual re-combination is of particular interest. The bi-sexual reproduction of higher organisms produces by means of a special kind of chromosome distribution and recombination a being with unique hereditary characteristics and with a unique combination of biological properties (except in the case of uniovular twins where the same hereditary characteristics are transmitted to two organisms). Since the chromosome distribution and re-combination are molecular processes, the properties of future organisms are subject to the principle of indeterminacy; they are thus in principle unpredictable.

We may also assume that certain diseases and disturbances of development are subject to this principle. We do not yet know whether this also applies to some of the processes whose functions affect the nervous system, but it seems a plausible assumption. Trial-and-error behaviour and some spontaneous reactions may e.g. be partly determined by random generators in the nervous system whose function rests on indeterminate molecular processes. In this case some aspects of behaviour of higher organisms would, in the short run, also be subject to the quantum indeterminacy. (It is to be stressed in this context that this indeterminacy does not correspond to those hypotheses which try to establish a direct relationship between free will and physically indeterminate processes in the nervous system.)

Over a period and probably even in the short run all the data mentioned above-weather, diseases, the constitution of other organisms and perhaps certain aspects of their behaviour-influence in many direct and indirect ways the environmental conditions relevant to a person; these conditions are thus subject, in an important way, to quantum indeterminacy. Contrary to the deterministic assumptions of classical physics, according to which every event at any time, including human destiny, was already prefigured and contained in the atomic configuration when the world began, quantum theory leads to the assumption of an open, indeterminate future of the human environment with a great many essentially different possibilities.

(b) The limits of formalizable thinking

Let us assume that the complete scientific analysis of a person must include an analysis of his thinking which in turn is to be regarded as a function of the nervous system. Quite apart from a detailed knowledge of neurophysiology, there exists a number of general logical grounds and arguments against the possibility of a complete analysis of man. These I shall now discuss.

One argument rests on the assumption that the information- content of the human thought system is limited. To be able to provide a complete analysis, the analyst would have to have at his disposal all functionally relevant parameters of the person who is to be analyzed, including the data of his information storage system. The information store of the analyst is not by itself adequate for this task, since it is of the same size and kind as that of the analysed person. This limitation is, however, not absolute, since the analyst can employ additional technical aids to store information and process data.

A more serious objection takes into account the effect of the analysis in bringing about changes in the analysed person. (The special problem involved in the mutual interaction in the physical analysis of the nervous system will be discussed in the next section.) In general, every question put to the person who is to be analysed, and indeed the very fact that he knows that he is being analysed, will affect his state (e.g. by embarrassing him). The results of the analysis may be correct, but they do not represent the general case (without analysis). Thus an exhaustive self-analysis is in a general sense just as impossible as an exhaustive analysis of other people, in as far as the latter become aware of the kind of analysis or of the fact that it is being carried out.

Another argument concerns the reaction following the announcement of the result of the analysis, particularly in the case of scientific forecasts (e.g. Mackay). One may adopt the view that private insights are not scientific insights, i.e. that only their inherent general accessibility invests them with scientific character. But forecasts which have been made known may influence their object; thus there are instances in which, because of this, two contradictory forecasts are wrong (successful pragmatic pessimism), as well as instances in which different forecasts are correct (successful pragmatic optimism). All this leads to the conclusion that the reactive effect of the analysis and of its results on the state of the analysed person sets effective limits to the possibility of a scientific analysis of man by man, if we presuppose that the analysed person is aware of the analysis or knows its results.

However, leaving aside the limitations which we have just outlined, we may ask the further question, namely whether, in the absence of every reactive effect on the analysed person, complete analysability is possible. Would a complete analysis of human thinking be possible on the assumption that we could discover everything about a person's state, including his nervous system, without any reciprocal effect?

The logical arguments against these possibilities can be based on the decidability theorems of mathematics which can be expressed roughly as follows in ordinary language: Every formal system which is rich enough to serve as a foundation of logic and number theory includes correct general propositions which can be formulated by means of the system, but cannot be proved within it. There can be no machine which, for every proposition which can be formulated within the system, could decide in a finite number of steps according to a general (algorithmic) method, whether the proposition is correct or not. Among the undecidable propositions is the proposition about the consistency of the system; this proposition cannot, therefore, be proved within the system. It is possible to think of a richer system in which the undecidable propositions of the poorer system (including the proposition about its consistency) become decidable; but the extended system in turn yields new undecidable propositions, including the proposition about its own consistency. An absolute, unconditional consistency proof is, in the case of complex logical systems, impossible.

These results have led to contradictory arguments about the nature of human mental processes; they range from the often quoted aphorism attributed to Weil: 'God exists because mathematics is consistent, and the devil exists because we cannot prove it', to the view that no conclusion at all can be drawn about the faculty of human cognition (cf. Lorenzen).

A detailed study of the problem was undertaken by Stegmüller; the contradictions of the various views can even be regarded as in some sense confirming his conclusion. It states that all formal thinking rests on non-formal presuppositions which must be justified by evidence; that is to say that intuitive thinking can never be completely formalized and human thought can never be given an absolutely sound foundation.

If we accept this conclusion and if we regard human thinking as a property or function of the nervous system, we realize that a complete formalization of human thinking cannot be achieved even by means of a full analysis of the nervous system. Thus either a full analysis of the physical processes in the nervous system is impossible, or it is possible but does not-for reasons which we shall discuss below-lead to the kind of theory of the relation between mental and neurophysiological processes which would allow a formalization of all thinking.

(c) The finiteness of the universe, limited formulae and limited proofs

According to current cosmological ideas, the duration and extent of the universe are finite. But even if we postulate a periodic universe or a 'steady state' universe or (finitely or infinitely) many universes, the exchange and storage of information is still only possible within the duration or period of one universe. This means that the total number of realizable physical and mental operations is also limited. The largest conceivable computer is smaller than the universe itself. A (very generous) estimate of the upper bound of realizable operations might be the number of stable elementary particles in the universe of the order of magnitude (10^{80}) multiplied by the duration of the universe in elementary time intervals (of the order of magnitude 10^{40}). (Operations of lesser duration than the elementary time interval would destroy the stability of elementary particles.) The resulting number is of the order of magnitude $Z = 10^{120}$ [1] This number is in fact much smaller in the case of the human domain, if for no other reason than that the brevity of human life limits not only the temporal but also the spatial range of information exchange to 100 light-years. In practice the number is determined by the size and operational speed of the groups of electronic computers and human brains which can be co-ordinated with each other. The exact value of this number is not very important. In the following discussion we shall assume that in the entire cosmic domain no experiments or calculations with more than Z operational steps, no propositions with more than Z signs, no catalogues with more than Z items, etc. can be realized; such

[1] This is the same number as that of the 'Uralternatives' in the universe in C. F. von Weiszäcker's schema of a finitist quantum theory.

calculations are thus inadmissible even as parts of thought experiments. It might be possible to speculate about how the world would look from the point of view of an imaginary supercosmic computer. But such a question appears meaningless within a theory of science and we shall not deal with it.

The very large numbers of the order of magnitude Z do not, of course, ever occur in reality, but they occur readily in quite ordinary problems as a combinatorial result; that is to say as the number of possibilities. The number of the possible compositions of a passenger list of an aircraft, of possible combinations of hereditary human properties, of possible one-page poems, etc., are all larger than Z. For the human nervous system which consists of some 10^{10} cells, the number of essentially different possible states is very great indeed. (We regard those states as essentially different whose differentiation is a conditions for a good approximation of the physical description of the reactions.) It is not yet possible to calculate the number of possible states because, on the one hand, the state of the individual nerve cells is determined by a number of as yet not fully understood parameters and, on the other hand, because the function of the nervous system does not, on the whole, seem to be dependent on individual nerve cells, but manifests a high degree of redundancy. The likely number is of an order of magnitude between 10^{10^7} and $10^{10^{10}}$ and in any case much larger than Z.

If we start from the assumption that the number of realizable mental and physical operations cannot exceed Z, it follows that in the case of complex data it is not only practically but even theoretically impossible to examine all possibilities individually, i.e. to start with a catalogue of all possibilities. Propositions and formulae about more than Z possibilities must, if they are to be meaningful, contain a limited number of signs or operations which is much smaller than Z. The definition of the signs must, of course, again be possible in fewer than Z individual steps; it must also be possible to carry out the operations in fewer than Z individual steps. In other words, meaningful propositions about complex data are not lists of individual propositions, but limited, general relations, in which certain properties occur while other properties are left out of consideration. We shall call such propositions limited formulae. If we say that a limited formula is valid for more than Z possibilities, then it is not enough to have a method which allows us to decide for every possibility whether the proposition is correct or whether a certain property is present. To prove it we must find a limited proof which proves the general validity of the limited formula in a number of steps which is much smaller than Z. One would expect such proofs to be available for some though not for all true limited formulae.

Furthermore, we have to bear in mind that for complex formulae (which consist of more than several hundred signs) there are also more than Z possibilities. Even if a limited proof is available for such a formula, there will not always be a way (apart from lucky coincidence) to hit upon this true formula, since it is in principle impossible to test all formulae of this degree of complexity.

In complex systems, such as e.g. the nervous system, the restriction of scientifically meaningful general proposition to limited formulae leads to the following possibilities: there may exist true limited formulae to describe the complicated relations between environment, neurophysiological processes, and behaviour for which no limited proof can be found, i.e. which cannot be proved within the limits of the universe, even if every special case of the formula can be decided individually and even though an imaginary supercosmic computer might be able to give a proof. There may also exist limited true formulae which, once they have been formulated, can be proved or shown to be probable, but which cannot be found by a formal method, but at best by chance (a chance which, under the circumstances, would be extremely unlikely). Thus true general propositions may exist which-by reason of the necessarily finite number of possible operations-can either not be found or not be proved.

3. THE POSSIBILITY OF A PHYSICAL ANALYSIS OF MAN

If physics is applicable to man in every domain, then it follows immediately from the fundamental equations of physics that the physical state of a person at a given time and the accurate sequence of the influences of the environment determine all objectively observable properties for every future moment within the limits set by the principle of indeterminacy. The capacity of the nervous system to function presupposes that the parameters relevant to its function (except e.g. possible random generators for trial-and-error procedures such as might also be used in electronic computers) are not, as a rule, subject to the indeterminacy principle; this implies that human behaviour is in the main-with the possible exception of a few statistical kinds of behaviour-determined by the laws of physics inasmuch as a certain sequence of environmental influences is given. But this presupposes that the relevant physical parameters of the initial state can be discovered without disturbing the nervous system. We are here not concerned with the practical possibilities which certainly do not exist at present; but rather with the question whether even for a consistent thought experiment certain limits are necessarily set to the discovery of the parameters. It might e.g. be relevant to long-term memory whether among the numerous similar molecules of the nervous system a certain molecule, such as a protein or a nucleic acid, defined by a specific sequence of amino acids or nucleotides, occurs at least once. Whether a complete sequential analysis of all the molecules which occur in the nervous system can be carried out without creating a disturbance and without being subject to the indeterminacy principle, is a question which requires further examination. So far, however, there are no plausible grounds for such an assumption, and it is probable that the relevant parameters of the nervous system can in general be physically analysed. On this assumption the fundamental equations of physics yield for any given sequence of environmental influences (including words and gestures) a definite sequence of reactions.

In a thought experiment it is possible to conceive an analogue computer which completely maps the physical state relevant to the functions of a person, and which would allow one to test and predetermine all reactions to a sequence of environmental impressions. The reactions include all physically determinable processes, including answers to questions and other verbal statements, as well as those reactions (such as e.g. blushing) which allow us to draw more or less accurate conclusions about the person's mental state (e.g. anger).

It is by no means necessary that the imaginary analogue computer should be inconceivably big. The number of nerve cells in the human brain is of the order of magnitude 10^{10} (or 10^{11}). If a nerve cell is represented in an element of the size of one cubic millimeter the machine would be approximately as big as a house. The machine would be characterized not so much by its size as by the great complexity of the connections between the elements. But even this complexity remains within stateable limits. The human brain (at birth) is a product of physiological development; its functionally relevant pattern is determined by the person's genes which are contained in the linear sequence of the order of magnitude 10^{10} nucleotides of DNA. In the course of life the state of the nervous system is constantly being changed by learning and experience. If we accept that the information-content of long-range memory is in the region of about 10^{10} 'bits', then this would increase the complexity of the imaginary analogue computer from birth onwards but would not greatly affect its order of magnitude. Strictly speaking we would also have to take into account that to look at the nervous system in isolation is not enough; but the inclusion of the other parts of the human body would not materially alter the size or complexity of the analogue computer. The functionally complete plan of the human nervous system can thus theoretically be recorded by a number of bits of the order of magnitude 10^{10} , i.e. in a library of some 10,000 volumes, albeit in a complicated and as yet unknown code. Thus, though the number of possible operations is limited, it is in principle possible to map the human nervous system on an analogue computer. Our thought experiment is therefore quite legitimate. The validity of physics for human biology implies that the objectively observable reactions of man to any given sequence of outside impressions is physically determined, since they are theoretically calculable by an analogue machine (with the exception of possible statistical aspects of behaviour which are subject to the indeterminacy principle and are probably in the majority of cases not very significant).

4. BEHAVIOURAL FORMULAE

The two preceding sections outline arguments in support of logical limits to the analysability of man on the one hand, and in support of physical determinism on the other. I shall now try to show that this does not involve a contradiction and propose a hypothesis which takes into account both types of argument. The hypothesis sets out from the consideration that though the reaction of a person to any given sequence of outside influences is determined physically, we must yet take into account the indeterminacy of the future of a person's external conditions. The open future includes a very large number of essentially different alternatives, a number which exceeds - at least in the medium and long range-the cosmological number Z - i.e. the greatest possible number of operations in the universe. The individual events which really occur in the medium or long range are always improbable. This does not mean, of course, that science cannot make probable (and at times even certain) predictions: but these calculations are not a function of the many individually possible sequences of events in the outside world but only of their general properties. Any possible knowledge of future behaviour is confined to those general relations which can be represented in limited formulae. Such limited formulae which express a relation between a certain state of a person (particularly of a person's nervous system), properties of sequences of outside influences and properties of future states, including the behaviour of the person, we shall call behavioural formulae.

Thus a behavioural formula while referring to a certain person in a certain state yet refers to all or very many possible sequences of outside influences. It can contain predictions of certainties or of probabilities. The simplest form would be an unconditional statement about the probability of a certain reaction at a certain future time which represents the average of the probabilities of all possible outside influences. But formulae which are functions of the outside influences are vastly more interesting. (An example of a behavioural formula would be a possibly complicated description of the conditions accompanying future events under which a certain person would vote for a certain party in next year's elections.)

If we now in a thought experiment apply the physical laws to an analogue computer of the human nervous system (Section 3), it becomes theoretically possible to test the correctness or incorrectness of every formula with respect to every single possibility of sequences of outside influences. But if there are more than Z possibilities, then we cannot test them. In the open environment there is a very great manifold of possible sequences of outside influences which is greater than Z. A given state of the nervous system can accordingly merge into a very great manifold of different states which in turn determine the complex behaviour. Even with the help of the imaginary analogue computer it is impossible to test all these possibilities for a behavioural formula. Nor is it possible to test the provability of all conceivable behavioural formulae, since their number is also much greater than Z. It therefore follows that the complete applicability of physics to man and the complete mapping of man on an analogue computer do not as such lead to a general method by means of which all true behavioural formulae could be discovered or proved. On the contrary, we may suppose that true formulae exist which cannot be discovered or cannot be proved.

The undecidability theorems of mathematics state that in complex logical systems it is not the case that a finite deduction or proof is available for every general proposition about infinite data. Nor is it to be expected that a deduction or proof in fewer than Z steps exists for all behavioural formulae about the possible states of the human nervous system which, though finite, are very varied and greatly exceed Z. By analogy to the undecidability problems in mathematics which (expressed informally and rather vaguely) occur when a high degree of abstraction and self-referential statements are involved, one may conjecture that deduction and proof of a behavioural formula from the state of the nervous system will be lacking mainly in those instances where highly abstract and self-analytical thought processes correspond to the processes in the nervous system. If this conjecture is correct, it would mean that there is a connection between the impossibility of determining the behavioral formulae on the basis of the physical state of the nervous system and incomplete formalizability of the thought- and other processes which correspond to this state. I need not stress that this analogy is purely conjectural; nor can its formulation lay claim to formal precision; it would require much more searching study to substantiate it.

It is obvious that the behavioural formulae assert less and thus have a greater degree of indeterminacy than an extremely detailed physical analysis of the relation between given states of the nervous system, outside influences and reactions, such as might be provided by a supercosmic computer; however, such an analysis may be regarded as meaningless from the point of view of scientific theory.

5. ASPECTS OF THE MIND-BODY PROBLEM

A central problem of man's knowledge of himself is the so called mind-body problem. Even its formulation-quite apart from its possible solution-presents great difficulties. If certain philosophical assumptions are made the problem appears spurious; but this does not do away with the logical and empirical questions which lie concealed behind the conventional dualistic conception of the mind-body problem; they reappear in a different formulation and a different context. For a detailed discussion of various aspects, I would refer the reader to Feigl's full exposition and analysis of the psychophysiological problem[2]. In the following sections we shall regard the relation between what is subjectively given and what is physically measurable as of fundamental importance.

Man is capable of objective observation, ultimately by means of his senses, be it directly or by means of measuring instruments; he is capable of consciously or unconsciously analysing and processing the observed data. It is probable that the scientific regularities of the observed data are in principle reducible to the laws of physics. Beyond this man is endowed with immediate self-awareness in the form of feelings, thoughts, inclinations, etc. without the mediation of his senses. These insights contribute to his mental state. This state can, at least partially, be expressed by speech, gestures and other forms of behaviour, and communicated to others. The elements (in a direct as well as in a metaphorical sense the vocabulary) of this expression are partly innate, partly learned and partly given by analogical reasoning from one's own to other minds. Even very complex mental states are communicable by combining different elements of expression and can be communicated even when the experience of the complex state cannot be directly reproduced.

Let us therefore assume that the mental state which is immediately given to the experiencing mind can at least partially be communicated and can thus be the object of scientific inquiry. We shall not discuss the question whether and in what sense the mind can have other private experiences which are not communicable intersubjectively. This does not affect our conclusions.

We shall also leave the question open as to what constitutes a criterion for the existence of consciousness. We may suppose that one person is aware of the mind of another only if the latter can communicate itself intersubjectively, i.e. if it is capable of human expression in some form or other; the phenomenon of communicability is confined to man and, in a deficient form, to higher animals. Questions about the mind of other organic, technical or other objects would be meaningless or undecidable. However, it is conceivable, though not very plausible, that the theory of consciousness will one day become part of a more general and more complete physics with quite different criteria than those which we can conceive of at present.

A fundamental aspect of the mind-body problem is the correlation between the mental state which is immediately given to consciousness and can yet be described by an intersubjective language, and the physical state of the experiencing person, particularly of his nervous system which is objectively measurable. One thought experiment to discover such correlations is 'autocerebroscopy' (cf. Feigl); it consists in registering independently mental experiences (thoughts, feelings, etc.) and the objectively measurable processes in the ner-

[2] The approach of this paper differs in some respects from Feigl's considerations but it is not necessarily inconsistent with his identity hypothesis.

vous system, and in subsequently trying to analyse the relation between them.

Expressed in the language of the conventional dualism, the main doctrine concerning this psychophysical relation is the psychophysical parallelism which asserts the deducibility of mental from physical states and is thus in keeping with the assumption that physics is completely applicable to biology. A similar assertion results from monistic concepts. The opposite view is that of the interactionist theories which postulate that the mental state exercises a somehow determining influence upon the physical state, thus throwing doubt on the validity or (at least) the completeness of physical laws when applied to biology.

According to Feigl, the psychophysical problem is not purely philosophical or logical, but has also purely empirical aspects. Thus e.g. psychophysical parallelism or the unlimited applicability of the known physical laws to the nervous system could be proved or refuted only with the help of psychological and neurophysiological data.

6. A HYPOTHESIS CONCERNING THE PSYCHOPHYSICAL PROBLEM

The preceding discussion permits the inference that parallelism and interaction are not genuine alternatives with regard to the psychophysical problem.

Assuming physics to be fully applicable to man, human behaviour is physically determined (within the limits of the principle of indeterminacy) for any given sequence of outside influences; from the physical reactions we can (within the usual margin of error) draw retrospective conclusions about the mental state. The objectively observable forms of expression (e.g. laughter) of the mental state (e.g. joy) are on these assumptions theoretically calculable results of physical processes. In a thought experiment, the calculation can be left to a suitably programmed analogue computer (cf. Section 3), which can also be asked questions. But it is important that in complicated cases the machine's answers to questions should not allow an unambiguous retrospective conclusion about the mental state, since the occurrence and sequence of the questions can influence the answers, and since questions can also receive false or paradoxical answers or no answers at all. But the possibilities of prediction of future physical states including behaviour which are supplied by the knowledge of the present physical state are complete in the sense that they cannot be enriched by the knowledge of the immediately given mental state.

However, the case of a person exposed to his indeterminate environment and facing an open future is different. As we have shown in the last section, the possible scientific predictions are in this case confined to behavioural formulae and one would expect that not all behavioural formulae which are true of a person are deducible from the state of his nervous system.

Mental states which can i.a. be regarded as dispositions to future behaviour are closely related to behavioural formulae. We can draw conclusions from a given behavioural formula about the mental state and from a given mental state (e.g. thoughts, feelings, inclinations, decisions) we can construct behavioural formulae which allow us to make probable or certain predictions about behaviour.

It follows that the arguments which apply to the deducibility of behavioural formulae also apply to the deducibility of the mental state. There is no logical reason why we should expect every mental state to be deducible from the structure of the nervous system in fewer than Z steps, i.e. in limited formulae. We may, on the contrary, conjecture that such a deduction is impossible in complex cases, particularly when thought processes of a high degree of abstraction or self-analytical participation are involved. If this hypothesis is true, then there is no necessary incompatibility between the two arguments outlined in the earlier sections of which one denies the complete formalizability of human thinking and the other asserts the complete analysability of the human nervous system in physical terms, since the connection between physical states of the nervous system on the one hand and mental states on the other need not necessarily be completely describable by limited formulae. The proposed hypothesis in a sense denies the complete decidability of the psychophysical problem and thereby the possibility of constructing a complete scientific theory of the relations between physical and mental states.

7. THE LIMITATION OF PSYCHO-PHYSICAL PARALLELISM

The hypothesis outlined in the preceding section restricts the scope of psychophysical parallelism without, however, accepting the interactionist theory and without denying the two fundamental physical assumptions of parallelism. One of the assumptions is that physics is completely applicable to human biology. This assumption has received extensive empirical confirmation by the findings of molecular biology, neurophysiology, biological cybernetics and other branches of biology. The second assumption asserts that mental states are a function of the physical states of the nervous system. Somewhat more precisely this means that the relation between mental states and physical states is either a one-one or a one-many but never a many-one correspondence, i.e. that two different mental states can never correspond to one physical state (cf. Feigl). This assumption is indirectly supported by numerous findings about the psychological effects of chemo-physical changes in the nervous system, as well as by those branches of cybernetics which rest on the possibility of establishing a one-one correspondence between formalizable thought and physical processes. While it may be premature to speak of satisfactory empirical confirmation, we may regard this assumption as very probably correct.

It is not necessarily the case that the existence of a one-one correlation between mental and physical states is sufficient for the calculation of all general and complex psychophysical relations between physical and mental states if we take into account the fundamental limitation of the number of possible operations.

If man were exposed to a predetermined sequence of outside influences (or to a surveyable number of alternatives), then psychophysical parallelism might be correct, since all objectively observable reactions are in principle calculable on a physical analogue machine in such a way that the knowledge of the mental state does not enrich the physical description and thus need not enter into it. That which can be intersubjectively communicated about the mental state would be confined to what can be deduced from physical behaviour.

From the point of view of a computer whose size greatly exceeds that of the universe, parallelism could also be correct in the case of an open future (however in this context there arises the question about the internal consistency of such a machine); but if we limit scientifically meaningful thought operations to those within our universe, then psychophysical parallelism is probably not completely valid in the general case, i.e. in the case of an open future. The mental state can be largely, but not wholly, deduced from the physical state. Knowledge of the mental and of the physical state allows more far-reaching objective predictions than the knowledge of the physical state alone. This is not a consequence of the interaction between the mental and the physical states as the interactionist theory assumes. Is is rather a result of our inability to draw, with the finite means at our disposal, all the conclusions from the measured physical state, so that the additional knowledge of the mental state can increase our total knowledge and thereby also the degree of computability of future states.

I do not intend to discuss here the difficult question whether and in what sense the above state of affairs allows us to speak of a determining influence of the mental state. If the question is to be answered affirmatively, this would by no means imply that the mental state influences physical states beyond or outside the terms of the known physical regularities either by changing or suspending the physical laws or by determining processes which are physically subject to the principle of indeterminacy.

From the point of view of a complete psychophysical parallelism it might be objected that machines could be constructed which would have self-analytical functions and other properties more or less analogous to consciousness and for which a complete formal theory would be available. But such an objection lacks cogency because it presupposes complete formalizability without proving it. In the nature of things it is easier to give a formal description of certain partial aspects of consciousness than to refer to something that is not formally describable. For this reason we shall outline a speculative idea about properties of the nervous system which may possibly not be representable by limited formulae. According to a note by Lorenz a situation is particularly prone to enter into our conscious- ness if it contains a conflict which cannot easily be resolved, e.g. if two different kinds of behaviour correspond to one complicated situation. In terms of neurophysiology this could mean that two contradictory pieces of information arrive in quick succession in some area of the nervous system ('yes' and 'no' about a certain datum). But if the production of such contradictions in neurophysiological processes plays an essential part in determining what reaches consciousness then the formal criteria of contradiction could constitute the presupposition of a complete psychophysical theory. By analogy to the decision problems of mathematics, one doubts whether such criteria are obtainable in the form of limited formulae in cases where the enquiries are sufficiently complicated. It is to be noted that the last-mentioned consideration, which rests on analogy, is quite speculative. Its validity could only be confirmed by a much more detailed study; it is merely intended to show that it is equally inadmissible to regard formalizability as a matter of course.

8. THE PROBLEM OF FREE WILL

The concept of a behavioural formula, which we introduced earlier, is very closely connected with another aspect of the mental state, namely the will. The direction of the will, and a corresponding decision, can essentially determine a behavioural formula. Therefore, the will and the decision are again not necessarily (and probably not always) deducible from the physical state of the nervous system in fewer than Z steps, in so far as will and decision concern an environment with an open future. We may therefore assume that it is in principle impossible to objectively deduce or predict human will and decisions in every case though it may be possible in principle to determine the actual reactions in the case of a given sequence of outside influences.

The freedom of the will which we experience in our selves means, in the first instance, that our decisions are not determined by outside factors alone, but also by factors within ourselves. It does not, however, mean that the inner factors are not physical. Nearly all philosophers of science agree that the question whether the inner factors are themselves determined by physics is irrelevant to the problem of the freedom of the will. For the functions of the nervous system which are relevant to thinking determination is a prerequisite of free will, since the nature of the latter is based on deliberation and not on chance. But the question remains whether and to what extent our decisions are determined for an observer who has extensive means for the analysis and the manipulation of environmental conditions. Part of the answer to this question must certainly be affirmative. But our earlier considerations render it probable that-assuming an open future-a complete analysis of human will and human decisions by an outsider is in principle impossible, even with the help of extensive means of analysis.

9. CONCLUSIONS

In this paper I attempted to show that complete applicability of physics to human biology does not necessarily imply that there is a finitist and at the same time complete correlation between physical and mental states. Further, I proposed a hypothesis which denies in principle the possibility of a complete theory of the psycho-physical relation. In the present state of our empirical, neurophysiological and psychological knowledge the proposed hypothesis can be neither proved nor refuted. More detailed logical enquiries, e.g. about decision problems in cases where in the light of biological facts, the number of data is finite and the number of operations limited, might also prove relevant. The hypothesis also concerns only some of the various aspects of the psychophysical problem; it does not explain why and under which conditions consciousness exists, and what is it.

Our conjecture is mainly based on the fact that it plausibly unifies the universal validity of physics on the one hand and a logical limitation of human thinking (e.g. with respect to itself) on the other, and that it moreover does not regard consciousness-the most basic human experience-as a marginal phenomenon.

BIBLIOGRAPHY

H. FEIGL, *The Mental and the Physical*, Minnesota Studies II, p. 370. Minneapolis, Minn., 1958.

P. LORENZEN, *Metamathematik*, p. 132. Mannheim, 1962.

D. M. MACKAY, in *Man and his Future*, p. 153. London, 1963.

W. STEGMULLER, *Metaphysik, Wissenschaft, Skepsis*. Frankfurt and Vienna, 1954

W. STEGMULLER, *Unvollständigkeit und Unterscheidbarkeit*, Vienna, 1959.

Max-Planck-Institut für Virusforschung
Tübingen, Germany
(Translation from German)