

Minnesota Studies in the  
**PHILOSOPHY OF SCIENCE**

---

VOLUME II

*Concepts, Theories, and the Mind-Body Problem*

EDITED BY

HERBERT FEIGL, MICHAEL SCRIVEN  
AND GROVER MAXWELL

FOR THE MINNESOTA CENTER FOR PHILOSOPHY OF SCIENCE

1958

UNIVERSITY OF MINNESOTA PRESS, MINNEAPOLIS

## *The "Mental" and the "Physical"*

### I. A Preliminary Survey of Some Perplexities and Their Repression

Tough-minded scientists tend to relegate the mind-body problem to the limbo of speculative metaphysics. Perhaps after trying a bit, but with questionable success to square themselves with the puzzle, they usually take one or the other of two attitudes. Either the puzzle is left to the philosophers to worry about, or else it is bluntly declared a pseudoproblem not worth pondering by anybody. Yet, the perplexities crop up again and again, often quite unexpectedly, if not in central issues of substantive scientific research, then certainly, and at least in connection with the attempts to formulate adequately and consistently the problems, the results, and the programs of scientific inquiry. The disputes regarding the very subject matter and definition of psychology furnish a poignant illustration. Is it *mental experience* or is it *behavior*?

The behaviorist revolution in psychology, as well as its opposite philosophical counterpart, the phenomenalist point of view in epistemology, each in its way, tried to obviate the problem. But all sorts of perplexities keep bedeviling both parties. The problem may be repressed, but repression produces symptoms, logical symptoms such as paradoxes or inconsistencies in this case. The behaviorist psychologist assimilates his method to that of the "objective" natural sciences. [Scientific psychology, as the well known saying goes, having first lost its soul, later its consciousness, seems finally to lose its mind altogether.] Behaviorism, now after more than forty years of development, shows of course many signs of mitigation of its originally rather harsh and radical position. It has availed itself of various clothings from the storehouse of philosophical garments. But despite the considerably greater scientific and logical sophistication in recent treatments of the issue, it is somewhat

depressing to note that the main philosophical positions still are these: materialism, mentalism, mind-body interactionism, evolutionary emergence theories, psychoneurophysiological parallelism (epiphenomenalism, isomorphism, double aspect theories), and neutral monism. Characteristically, the phenomenalist and the behaviorist positions, refined descendants or variants respectively of the mentalistic and the materialistic philosophies, have been most forcefully advocated by the positivists of the last and of the present century. Positivism, more distinctly than any other point of view, with its notorious phobia of metaphysical problems and its marked tendency toward reductionism, was always ready to diagnose the mind-body puzzle as a *Scheinproblem*. Small wonder then that phenomenism (or neutral monism) on the one hand, and physicalism on the other, have been the favored positions in various phases of the history of the positivistic outlook.

In the philosophy of the enlightenment of the eighteenth century we find the outspoken and clear-headed phenomenism of Hume, but also the equally explicit, though more "simplistic" French materialism, especially of Baron d'Holbach. The German positivists of the nineteenth century, led by Mach and Avenarius, were essentially Humeans. And so was Bertrand Russell in one of the earlier phases of his epistemological odyssey. It was the combined influence of Russell's phenomenism (or neutral monism) and of the logic of *Principia Mathematica* which led Carnap in his early work *Der Logische Aufbau der Welt* (1928) to elaborate in considerable detail and with remarkable precision a logical reconstruction of the relation between psychological and physical concepts. He chose as a basis for this reconstruction a set of neutral experiential data and showed how the concepts of various scientific disciplines can be constituted as logical constructions erected on a basis of concepts which refer to elements and relations of that (subjectless) raw material of immediate experience. Carnap's attempt was thus a culminating point in the series of positivistic-phenomenalist epistemologies. But certain grave objections and difficulties soon made Carnap abandon this scheme and replace it by another, different in basis and structure. His new reconstruction is physicalistic in that the basic elements and relations are the designata of an intersubjective observation language (viz., the physicalistic thing-language). The difference in logical structure is due mainly to the recognition that the Russellian hierarchy of types does not adequately explicate the category mistakes which

undoubtedly give rise to some (though by no means all) mind-body puzzles.

The physicalist views of Lashley (192), Carnap (62, 64, 66, 67), Hempel (146), Black (37, 38), Quine (268), Ryle (294), Skinner (321), and Wilfrid Sellars (315), though differing in many more or less important respects among each other, are primarily motivated by a basic doubt about the possibility of a purely phenomenal language. The observation language of everyday life, we are told, is rooted in the intersubjective terms whose usage we acquired in the learning situations of a common, public context of labeling things, properties, relations, states, events, processes, and dispositions. Subjective or "mentalist" terms, this group of thinkers claims, are introduced and their usage learned on an intersubjective basis. Remove this intersubjective basis and you have deprived psychological concepts not only of their scientific significance, but you are left with nothing more than ineffable raw feels or with exclamations devoid of cognitive significance.

But the problems will not completely yield to this reductive approach. Introspection, though admittedly often unreliable, does enable us to describe elements, aspects, and configurations in the phenomenal fields of direct experience. When the doctor asks me whether I have a pain in my chest, whether my mood is gloomy, or whether I can read the fine print, he can afford to be a behaviorist and test for these various experiences in a perfectly objective manner. But I have (or do not have) the pain, the depressed mood, or the visual sensations; and I can report them on the basis of direct experience and introspection. Thus the question arises inevitably: how are the raw feels related to behavioral (or neurophysiological) states? Or, if we prefer the formal mode of speech to the material mode, what are the logical relations of raw-feel-talk (phenomenal terms, if not phenomenal language) to the terms and statements in the language of behavior (or of neurophysiology)?

No matter how sophisticated we may be in logical analysis or epistemology, the old perplexities center precisely around this point and they will not down. Many philosophical positions at least since the eighteenth century were primarily motivated, I strongly suspect, by the wish to avoid the mind-body problem. Moreover, the central significance of the problem for any *Weltanschauung* burdens its clarification with powerful emotions, be they engendered by materialistic, idealistic

or theological prepossessions. Schopenhauer rightly viewed the mind-body problem as the "Weltknoten" (world knot). It is truly a cluster of intricate puzzles—some scientific, some epistemological, some syntactical, some semantical, and some pragmatic. Closely related to these are the equally sensitive and controversial issues regarding teleology, purpose, intentionality, and free will.

I am convinced, along with many contemporary philosophical analysts and logicians of science, that all of these problems have been unnecessarily complicated by conceptual confusions, and to that extent are gratuitous puzzles and pseudoproblems. But I feel that we have not yet done full justice to any of them. Repression by reductionist philosophies (positivism, phenomenalism, logical behaviorism, operationism) is fortunately going out of fashion and is being replaced by much more detailed and painstaking analyses, of both the (Wittgensteinian) "ordinary language" and the (Carnapian) "reconstructionist" types.

Collingwood once said "people are apt to be ticklish in their absolute presuppositions; [they] blow up right in your face, because you have put your finger on one of their absolute presuppositions." One might add that philosophers are hypersensitive also in their repressed perplexities. A puzzle which does not resolve itself within a given favored philosophical frame is repressed very much in the manner in which unresolved intrapersonal conflicts are repressed. I surmise that psychologically the first kind may be subsumed under the second. Scholars erect certain ideas so strongly and their outlook becomes so ego involved that they erect elaborate barricades of defenses, merely to protect their pet ideas from the blows (or the slower corrosive effects) of criticism. No one can be sure that he is not doing this sort of thing in a particular case, and I claim no exception for myself. The best one can do is to proceed with candor and to subject oneself to ruthless criticism as often as feasible and fruitful. Techniques of self-scrutiny are nothing new in philosophy, but implemented by modern depth-psychological tools they could surely be made much more effective. In this candid spirit, I shall begin by putting my cards quite openly on the table; in the next two sections I shall indicate what I consider the sort of requirements for an adequate solution of the mind-body problems. I have no doubt whatever that some philosophers or psychologists will differ from me even in these first stages. All I can do then is to try, first to make these requirements as plausible as I can, and second, to analyze and evaluate

the assets and the liabilities of some of the various proposed solutions as fully as space permits.

## II. The Scientific and the Philosophical Strands in the Mind-Body Tangle

A first indispensable step toward a clarification of the issues is to separate the scientific from the epistemological questions pertaining to the relations of the mental to the physical. Epistemology is here understood in the modern sense of a logical analysis of concepts and statements and of the closely related logical reconstruction of the validation of knowledge claims. Some of the pertinent statements themselves are, however, essentially of a scientific nature in that they fall under the jurisdiction of empirical evidence. It is right here where we find a fundamental parting of the ways. Biologists, psychologists (and with them, many philosophers) hold deep convictions, one way or another, on the autonomy or non-autonomy of the mental. The strongest contrast is to be found between those who hold interactionistic views regarding the mental and the physical, and those who reject interactionism and hence espouse either parallelism (e.g., in its currently favored form, isomorphism) or some emphatically monistic view. Interactionism as well as parallelism are of course forms of dualism. The main difference and dispute between these two points of view is at present not fully decided by the evidence. But I think this is an issue to which empirical evidence is ultimately and in principle relevant.

Vitalists or interactionists like Driesch, McDougall, J. B. Pratt, Ducasse, Kapp, *et al.* hold that biological concepts and laws are not reducible to the laws of physics, and hence—a *fortiori*—that psychological concepts and laws are likewise irreducible. Usually this doctrine is combined with a theory of the emergent novelty of life and mind. But there are others who restrict emergence to the mental, i.e. they hold a reducibility view in regard to the biological facts. "Reducibility" is here understood to mean the same as "explainability"; and has no necessary connection with the introducibility (empirical anchorage) of biological or psychological concepts on the basis of physicalistic *observation* terms. As Carnap (67) has pointed out clearly, the thesis (*his* thesis) of the unity of the *language* of science does not in any way prejudice the issue of the unitary explainability of biological and psychological facts (or laws) on the basis of physical *theory*. Philosophers should certainly not

assume that such a basic scientific issue can be settled *merely* by logical analysis. It is *logically* conceivable that biological, psychological, and social phenomena (as well as their regularities) may not be explainable in terms of those physical or physicochemical laws (and theoretical assumptions) which are sufficient for the explanation and prediction of inorganic phenomena (and their regularities).

Logical parallels to such irreducibilities are clearly evident even *within* physics. The "mechanistic" (Newtonian) premises of explanation are now viewed as entirely insufficient for the explanation of electromagnetic radiation, of the dynamics of intra-molecular and intra-atomic processes, and of the interaction of electromagnetic radiation and the particles of matter. Nineteenth century physics added the fundamentally new concepts and laws of electromagnetics; and these in turn were drastically modified and supplemented by the relativity and quantum theories of our century. It is conceivable that homologous emendations may be required for the explanation of the phenomena of life and mind. Contemporary dualists, be they vitalists, emergentists, interactionists, or parallelists, maintain that such an enrichment of the conceptual system of science will be indispensable. Their arguments are based primarily on the traditionally captivating evidence of teleological processes, purposive behavior, psychosomatics, and the mnemonic and intentional features of perception, cognition, thought, desire, and volition. And some apparently very persuasive arguments point simply to the existence (occurrence) of immediate experience, i.e., the raw feels or hard data of the directly given. They maintain that these data, though *related* to behavior and neurophysiological processes, are not *reducible* to, or *definable* in terms of, purely physical concepts; and that their occurrence is not predictable or explainable on the basis of physical laws and physical descriptions only.

At this point the distinction between the scientific and the philosophical aspects of the mind-body problems becomes imperative. "Irreducibility" may mean non-derivability from a specified set of premises; but in other contexts it may mean non-translatability (non-synonymy, non-equivalence in the logical sense). To illustrate: many physical phenomena of sound or heat are derivable from the kinetic theory of molecular motion. In this sense certain parts of acoustics and of thermodynamics are reducible to mechanics, with a high degree of approximation at least within a certain limited range of the relevant variables.

But the phenomena of heat radiation (and similarly those of optics, electricity, magnetism, and chemistry) are not reducible to mechanics. Whitehead speculatively maintained that the laws pertaining to the motion of electrons in living organisms differ fundamentally from the laws of electrons in the context of inorganic lifeless bodies. In a similar vein the physicist Elsasser (95, 96, 97), following some suggestions contained in Bergson's views on organic life and memory, regards the physical laws as special or limiting cases of biological laws. This is a drastic reversal of the "Victorian" outlook according to which macro-regularities are (usually) explainable in terms of basic micro-laws.\*

As a student of the history and the methodology of modern science, and impressed as I am with the recent advances of biophysics, biochemistry, and neurophysiology, I am inclined to believe strongly in the fruitfulness of the physicalistic research program (involving micro-explanations) for biology and psychology. But qua analytic philosopher my intellectual conscience demands that I do not prejudge the issues of reducibility (explainability) in an a priori manner. Beyond the sketchy empirically oriented arguments which I am going to submit presently, I shall address myself later on primarily to the logical and epistemological aspects of the mind-body problem.

Along empirical lines I believe there are differences, in principle capable of test, between parallelism and interactionism (and/or emergentism). Psycho-neurophysiological parallelism is here understood as postulating a one-one, or at least a one-many, simultaneity-correspondence between the mental and the physical. Parallelism as customarily conceived clearly rules out a many-one or a many-many correspondence. This latter type of correspondence, if I may speak for a moment about the motivation rather than the evidential substantiation (confirmation), is generally unpalatable to the scientific (especially the "Victorian") point of view, because it would obviously limit the predictability of mental events from neurophysiological states of the organism. But given a "dictionary," i.e., more properly speaking, a set of laws correlating in one-one or many-one fashion physical and mental states, physical determinism is not abrogated.

\* I have dealt elsewhere (106, 108, 112, 113, 115, 116) with the logic and methodology of such explanations. See also the important articles by E. Nagel (230, 232); Hempel and Oppenheim (152); Kemeny and Oppenheim (177); Oppenheim and Putnam (in the present volume).

Two important qualifying remarks are in order here: (1) By "physical determinism" I mean, of course, that degree of precise and specific in-principle-predictability that even modern quantum physics would allow as regards the macro- and some of the micro-processes in organisms. (2) By "physical" I mean \* the type of concepts and laws which suffice in principle for the explanation and prediction of inorganic processes. If emergentism is not required for the phenomena of organic life, "physical" would mean those concepts and laws sufficient for the explanation of inorganic as well as of biological phenomena. In accordance with the terminology of Meehl and Sellars (221), I shall henceforth designate this concept by "physical<sub>2</sub>" in contradistinction to "physical<sub>1</sub>", which is practically synonymous with "scientific", i.e., with being an essential part of the coherent and adequate descriptive and explanatory account of the spatio-temporal-causal world.

In view of what was said above about the empirical character of the interaction and the emergence problems, the concepts of mental states might well be physical<sub>1</sub> concepts, in that they could be introduced on the basis of the intersubjective observation language of common life (and this includes the observation language of science). Just as the concept of the magnetic field, while not denoting anything directly observable, can be introduced with the help of postulates and correspondence rules (cf. Carnap, 73), so it is conceivable that concepts of vital forces, entelechies, "diathetes" (cf. Kapp, 172, 173, 174), and mental events might be given their respective meanings by postulates and correspondence rules. Of course, the question remains whether such ("emergent") concepts are really needed and whether they will do the expected job in the explanation and prediction of the behavior of organisms, subhuman or human. My personal view, admittedly tentative and based on the progress and partial success of physicalistic micro-explanation (implemented by Gestalt and cybernetic considerations), is to the effect that physical<sub>2</sub> laws will prove sufficient. But, having abandoned the all too narrow old meaning criteria of the earlier logical positivists, I would not for a moment wish to suggest that the doctrines of emergence or of interactionism are scientifically meaningless.

Let us then return to the empirically testable difference between interactionism and emergentism on the one hand, and parallelism on

\* In this context only; other meanings of "physical" will be listed and discussed in sections IV and V.

the other. An obvious and picturesque analogy or model for the interactionist view may be suggested here to provide a more vivid background. Billiard balls are in motion on a billiard table, and their motions are, we assume, predictable on the basis of mechanical laws (Newton's, supplemented by the laws of friction and of partially elastic collision). But imagine now a mischievous boy standing by, once and again pushing this or that ball or lifting some ball from the table. The mechanical laws, combined with a statement of initial conditions for the balls and the table, at a given moment, will then no longer suffice for the prediction of the course of the balls. The system in this case is of course an open one. If we could proceed to a larger closed system including the boy, with information about his shifting desires and so forth, deterministic predictability might be restored. (Since prediction of the boy's actions is precisely the issue at stake, I shall not beg any questions here and shall leave the boy's behavior unexamined for the moment.) This model is merely to illustrate a good clear meaning of "interaction". The boy watches the balls and his actions are in part influenced by their momentary distribution and motions on the board. The events on the board are in turn influenced by the boy's actions. From the point of view of ordinary usage, it is proper to employ the word "interaction" perhaps only when we deal with causal relations directed both ways between two continuants (things, organisms, persons, etc.).

But even a theory of emergence, such as the one suggested, though not definitely endorsed, by Meehl and Sellars (221), is confirmable in principle by showing that physical<sub>2</sub> determinism does not hold. Mental states or raw feels, be they regarded as states of an interacting substantial mind (or soul) or as values of emergent scientific variables, would in any case entail a breach in physical<sub>2</sub> determinism. The system of neurophysiological events inasmuch as it is describable in physical<sub>2</sub> terms would have to be regarded as open not only in the usual way, i.e., in regard to the extraneural, let alone extradermal, events, but it would also be open in regard to the set of mental events with which they are assumed to be causally (functionally) related in a way that would make them radically different from a set of mere epiphenomena. Now, while it is admittedly difficult at present to test for the implied breach in physical<sub>2</sub> determinism, the idea is not metaphysical in the objectionable sense that empirical evidence could not conceivably confirm or disconfirm it.

Much depends in this issue upon just how the "interactors" or the "emergents" are conceived. Traditional vitalism, culturally and historically perhaps a descendant of more primitive forms of animism, stresses the capricious nature of *vis vitalis* and of *anima*. (In our model the boy by the billiard table is assumed to exercise "free" choice.) But interaction need not be indeterministic in the wider system. The wind and the waves of the sea genuinely interact; even if the wind's influence is quantitatively greater, the waves do have some effect upon the air currents nearby. But though precise prediction of detail is practically extremely difficult because of the enormous complexities of the situation, this type of interaction is in principle deterministically \* analyzable in terms of the functional relations of the two types of variables. Even the individual "free" or "capricious" momentary choices of our boy might be predictable in principle; but here the practical feasibility is far beyond the horizon of current psychology. At best only some statistical regularities might be formulated.

Determinism, inasmuch as it is allowed for by current physical theory, is also the presupposition of the sophisticated conception of emergence as presented in the essay by Meehl and Sellars. Here we have no interacting things or substances, but scientific variables intertwined in such a way that certain values in the range of one set of variables are functionally so related to the values of the variables in the other set, that the relations in the second set are nomologically different from what they would be if the values of the first set are zero. More concretely, once mental states have emerged, their very occurrence is supposed to alter the functional relations between the neurophysiological (physical<sub>2</sub>) variables in a manner in principle susceptible to confirmation. While my (scientific) predilections are completely incompatible with this ingenious and fanciful assumption, I do consider it scientifically meaningful. I just place my bets regarding the future of psychophysiology in the "Victorian" direction. And I admit I may be woefully wrong.†

\* Again it is only to the extent that hydrodynamics and aerodynamics for macro-processes are (approximately) deterministic.

† In his earlier formulations of the general theory of relativity Einstein endorsed the so-called Mach principle, according to which centrifugal and inertial forces are the effects of accelerations relative to the total masses of the fixed-stars-universe. But, impelled by what he considered cogent physical and mathematical arguments, he later ascribed those effects to a relatively independently existing "Führungsfeld" (guiding field). I mention this merely as a somewhat remote logically parallel case from an entirely different domain of science. Naturally, my expectation here is that something

With the foregoing remarks I hope to have indicated clearly enough that I consider these basic issues as essentially scientific rather than philosophical. But a full clarification and analysis of the precise meanings and implications of, respectively, parallelism, isomorphism, interactionism, and the various forms, naive or sophisticated, of emergentism is a *philosophical* task. I shall now develop the philosophical explication of the factual-empirical meaning of these assorted doctrines a little further and bring out their salient epistemological points. Parallelism and isomorphism, now that we have recovered from the excesses of positivism and behaviorism, are generally considered as inductively confirmable hypotheses. Reserving more penetrating epistemological analyses, especially of the "immediate experience" and "other minds" problems, until later, I assume for the present purpose and in the vein of the recent positions of Ayer (15, 18) and Pap (243, 248) that the  $\psi$ - $\Phi$  (i.e., psycho-neurophysiological) relations or correspondences can be empirically investigated; and that mental states (raw feels) may by analogy be ascribed to other human beings (and higher animals), even if in the case of those "others" they are inaccessible to direct confirmation.

Parallelism, then, in its strongest form assumes a one-to-one correspondence of the  $\psi$ 's to the  $\Phi$ 's. It is empirically extremely likely that these correspondences are not "atomistic" in the sense that there is a separate law of correspondence between each discernible  $\psi_1$  and its correlate  $\Phi_1$ . It is quite plausible that, for example, different intensities of a phenomenally given tone (e.g., middle C), at least within a given range, are correlated with corresponding values in a limited range of some variable(s) of the neural processes in the temporal lobe of the brain.

Isomorphism as understood by the Gestalt psychologists (Wertheimer, Köhler, and Koffka) and the cyberneticists (Wiener, McCulloch, Pitts, etc.) assumes an even more complete one-one correspondence between the elements, relations, and configurations of the phenomenal fields with their counterparts in the neurophysiological fields which characterize portions of cerebral, and especially cortical, processes. As mentioned before, this sort of approach would also countenance a one-many correspondence of  $\psi$ 's and  $\Phi$ 's. In that case, mental states would of Mach's principle, even if in strongly modified form, will be salvaged. Powerful inertial forces as effects of a self-existent metrical field seem extremely implausible to me.

(with the help of the  $\psi$ - $\Phi$  "dictionary") still be uniquely inferable from neurophysiological descriptions. But many-one or many-many correspondences, even if expressed in terms of statistical laws, would seriously restrict such inferences from specific  $\Phi$ 's to specific  $\psi$ 's. I know of no good empirical reasons for assuming anything but one-one correspondence; or one-many if very exact and detailed  $\Phi$ -descriptions are used, and if account is to be taken of the limited introspective discernibility of the  $\psi$ 's from one another.

Interactionism, as I understand (but reject) it, would entail a many-one or many-many correspondence. Arthur Pap (242, p. 277), however, argued that there is no empirically confirmable difference between parallelism and interactionism. This, he thought, is because lawful relations or functional dependencies are the modern scientific equivalent of the cause-effect relation. Temporal succession, he maintains, is not a criterion of causal connection. While I admit that the most general conception of the causal relation is simply that of a (synthetic) sufficient condition,\* and is thus free of any connotation regarding the temporal succession of cause and effect; and though I also agree that in the case of  $\psi$ - $\Phi$  relation it would seem rather fantastic to assume anything like a time difference, I think that the interaction hypothesis differs in its empirical meaning from parallelism or isomorphism in that it entails a breach of physical<sub>2</sub> determinism for the  $\Phi$ 's. This, if true, could in principle be confirmed by autocerebroscopic evidence. For example, the experience of volitions as directly introspected would not be correlated in one-to-one (or one-many) fashion with simultaneous cortical states as observed (really inferred) by looking upon the screen of a cerebroscope,† and regularly succeeded by certain processes in the efferent nerves of the brain, ultimately affecting my muscles or glands, and thus ensuing in some act of behavior. This is the sort of most direct evidence one could ever hope for, as regards the confirmation of  $\psi$ - $\Phi$  action. If the idea of interaction, i.e., action both ways between the  $\psi$ 's and the  $\Phi$ 's, is entertained, then there should be sensations (produced by the

\* And in the laws of classical mechanics and electrodynamics of sufficient and necessary condition.

† This, for the time being, of course, must remain a piece of science fiction (conceived in analogy to the doctors' fluoroscope) with the help of which I would be able to ascertain the detailed configurations of my cortical nerve currents while introspectively noting other direct experiences, such as the auditory experiences of music, or my thoughts, emotions, or desires.

chain of processes usually assumed in the causal theory of perception, but) not strictly correlated with the terminal cortical events.

Characteristically, philosophers have been emphasizing much more the action of "mind on matter"—as in voluntary behavior, or in the roles of pleasure, pain, and attention—than that of "matter on mind." This asymmetrical attitude usually comes from preoccupation with the freewill puzzle, or related to this, from some remnants of theological ideas in the doctrines of an ideal ("noumenal") self. But the freewill puzzle—even if some details of its moral aspects still await more clarification—has in its scientific aspects been satisfactorily resolved by making the indispensable distinctions between causality and compulsion (and indeterminism and free choice). The perennial confusions underlying the freewill perplexity, truly a scandal in philosophy, have been brilliantly exposed by empiricist philosophers.\*

The main reasons why most psychophysicists (and along with them many philosophers) reject the hypothesis of  $\psi$ - $\Phi$ -many-one or many-many correspondence are these:

1. Normal inductive extrapolation from the successes of psychophysiology to date makes it plausible that an adequate theory of animal and human behavior can be provided on a neurophysiological basis. Most physiologists therefore favor  $\psi$ - $\Phi$  parallelism or epiphenomenalism. Parallelism, I repeat, is here understood as the assertion of the one-one (or, at least, one-many)  $\psi$ - $\Phi$  correspondence, and not, as by Wundt and some philosophers, as the doctrine of double causation, i.e., involving parallel series of events with temporal-causal relations corresponding (contemporaneously) to one another on both sides. Causality in the mental series is by far too spotty to constitute a "chain" of events sufficiently regular to be deterministic by itself. Epiphenomenalism in a value-neutral scientific sense may be understood as the hypothesis of a one-one correlation of  $\psi$ 's to (some, not all)  $\Phi$ 's, with determinism (or as much of it as allowed for by modern physics) holding for the  $\Phi$ -series, and of course the "dangling" nomological relations connecting the  $\Phi$ 's with the  $\psi$ 's. According to this conception voluntary action as well as psychosomatic processes, such as hysteria, neurotic symptoms,

\* Hobbes, Locke, and especially Hume, Mill, Sidgwick, Russell, Schlick (301); and Dickinson S. Miller, cf. the superb article he published under the pseudonym "R. E. Hobart" (157). See also C. L. Stevenson (329); University of California Associates (339); A. K. Stout (330); and Francis Raab (271).

and psychogenic organic diseases (e.g., gastric ulcers) may ultimately quite plausibly be explained by the causal effects of cerebral states and processes upon various other parts of the organism; only the cerebral states themselves being correlated with conscious (or unconscious\*) mental states.

2. While the cultural and historical roots of the epiphenomenalist doctrine may be the same as those of traditional materialism, we can disentangle what is methodologically sound and fruitful in the materialistic point of view from what is cognitively false, confused, or meaningless. The fundamental methodological reason for the rejection of interactionism, or the (equivalent) adoption of  $\psi$ - $\Phi$ -one-one (or one-many) correspondence as a working hypothesis or research program, however, is this: If the  $\psi$ 's are not inferable on the basis of intersubjectively accessible (observed, or usually, inferred)  $\Phi$ 's, then their role is suspiciously like that of a *deus ex machina*. The German biologist-philosopher Driesch admitted this candidly, and thereby gave his case for vitalism away. He said that the intentions of the entelechy could be inferred only *post factum*, but could not be predicted from antecedent physical conditions. This is just like the case, in our crude analogy, of the capricious boy at the billiard table. After he has removed a ball we may say that he intended (perhaps!) to avoid a collision of the red ball with the white one. According to the vitalist interactionist doctrine, the volitions of the boy are in principle unpredictable on the basis of any and all antecedent conditions in his organism and the environment. Interactionism so conceived assumes causal relations between the elements in the series of mental states, the series of physical states, but also some crossing from the set of mental states to the physical ones and vice versa. In the model of the wind and the waves, we have precisely this sort of schema exemplified. But notice the crucial difference. A closed system (or a system with known initial and boundary conditions) is here conceivable in which all relevant variables are ascertainable intersubjectively and antecedently to the prediction of later states of the system; whereas in the case of  $\psi$ - $\Phi$  interaction, intersubjective and antecedent confirmation of the  $\psi$ -states is *ex hypothesi* excluded.

The flavor of the theological arguments from design and of primitive animistic explanations of nature and human behavior permeates inter-

\* The terminological question whether to speak of the unconscious as "mental" will be discussed in sections IV and V.

actionistic explanations. They are at best *ex post facto* explanations. This sort of explanation, while not as satisfactory as explanations that also have *predictive* power, is nevertheless quite legitimate and is frequently the best we can provide in complex situations. Earthquakes are notoriously unpredictable (i.e., *practically* unpredictable), but once we observe a certain case of large scale destruction, its explanation in terms of an earthquake is perfectly legitimate even if the precise location of each piece of rubble in the shambles is far from predictable. Biologists are satisfied with evolutionary (retrospective) explanations of the emergence of a new species, even though they could never have predicted this emergence in any specific detail. Given the species in the Cambrian epoch, and given the principles of genetics and of Neo-Darwinian evolution, nobody could inductively infer the emergence of the chimpanzee or of the orchid; nevertheless, the very partial explanations of the theory of evolution are scientifically significant, acceptable, and helpful. Explanations of historical phenomena like wars, revolutions, and new forms of art furnish another illustration for the same type of *ex post facto* explanations. Finally, for an example in the psychological domain, if we find that a man has written dozens of letters of application for a certain type of job, we infer that he was impelled by a desire for such a job, even if we could not have predicted the occurrence of this desire on the basis of antecedent and intersubjectively confirmable conditions.

It is important, however, to notice again the decisive difference between explanations for which it is at least in principle conceivable that they could be predictive (as well as retrodictive), and those which *ex hypothesi* are *only* retrodictive. Scientists are predominantly interested in enlarging the scope of predictive explanations. The opposition against vitalism then stems from a reluctance to admit defeat as regards predictability. And the opposition against  $\psi$ - $\Phi$  interactionism stems furthermore from the reluctance to admit antecedents which are only subjectively accessible into the premises (regarding initial conditions) for predictive inferences. Expressing the same idea positively, we may say that it is part of the methodology or of the over-all working hypothesis of modern science that prediction, to the extent that it is possible at all (taking account of the basic quantum indeterminacies), is always in principle possible starting from *intersubjectively* confirmable statements about initial conditions. Scientists have, on the whole, adjusted themselves to the limitations involved in *statistical* prediction

and *probabilistic* explanation. Very likely nothing better will ever be forthcoming in any area except in the few where classical determinism holds with a high degree of approximation. Of course, a logical distinction should be made between those cases in which the restriction to probabilistic predictability is a consequence of the complexity of the situation, and those in which the *theoretical* postulates of a given domain are themselves formulations of statistical laws. Although one can never be sure that this distinction is correctly drawn or that the dividing line will remain in the same place during the progress of science, the distinction can be drawn tentatively in the light of theories well confirmed at a given time.

But scientists are radically opposed to the admission of *purely subjective* factors or data (conceived as in principle inaccessible to intersubjective confirmation) as a basis for prediction or explanation. This would indeed be scientifically meaningless, if not even statistical relations of subjective states to antecedent or consequent intersubjective observables could be assumed. If they are assumed, then the subjective states are not *purely* subjective or "private" in the radical sense intended by some interactionists. The "emergent" raw feels in the interpretation by Meehl and Sellars are of course subjective only in the sense that they can be the objects of direct introspective verification, but they are also intersubjective (physical<sub>1</sub>) in the sense that they can be assumed (posited, inferred, hypothetically constructed) by scientists who do not have the same sort of raw feels in the repertory of their own direct experience. This is so, for example, in the case of a congenitally blind scientist, equipped with modern electronic instruments who could establish the (behavioristic) psychology of vision for subjects endowed with eyesight. The blind scientist could thus confirm all sorts of statements about visual sensations and qualities—which in his knowledge would be represented by "hypothetical constructs." But if *ex hypothesi* all connections of the subjective raw feels with the intersubjectively accessible facts are radically severed, then such raw feels are, I should say by definition, excluded from the scope of science. The question whether discourse about such absolutely private raw feels makes sense in any sense of "sense" will be discussed later.

The upshot of this longish discussion on the difference between the scientific and the philosophical components of the mind-body problems is this: If interactionism or any genuine emergence hypotheses

are sensibly formulated, they have empirical content and entail incisive limitations of the scope of physical<sub>2</sub> determinism. Interactionism is more difficult to formulate sensibly than is the (Meehl-Sellars) emergence hypothesis. In one form it requires substances (things, continuants or systems of such) for a normal use of the term "interaction," and in this form there seems little scientific evidence that would support it. I have read a great many arguments by metaphysicians attempting to support the idea of a totally (or partially) immaterial "self." But I have never been able to discern any good cognitive reasons beneath their emotionally and pictorially highly charged phrases. Whatever role the self (in Freudian terms perhaps the total superego, ego, and id-structure) may play in the determination of human conduct, it may yet very well be explained by a more or less stable structure of dispositions due to some constitutionally inherited, maturationally and environmentally modified, and continually modulated structure of the organism (especially the nervous and endocrine systems).

In another form interactionism (without a self) would require "spontaneously" arising mental states, i.e., an indeterminism not even limited by statistical regularities, and this again is neither supported by empirical evidence, nor advisable as a regulative idea for research. Nor is it required for the solution of the freewill problem, or for an account of the causal efficacy of mental events in the course of behavior. As regards the emergence hypothesis (à la Meehl and Sellars), this clearly makes sense, but whether it is really needed for the explanation of behavior is an open question. In the spirit of the normal procedures of scientific induction and theory construction I remain conservative in thinking that the rule of parsimony (Ockham's razor, or Newton's first *regula philosophandi*) warns us not to multiply entities (factors, variables) beyond necessity. If the necessity should become evident in the progress of research, I shall cheerfully accept this enrichment of the conceptual apparatus of science; or, ontologically speaking, this discovery of new entities in our world. In the meantime, I remain skeptical about emergence, i.e., optimistic about the prospects of physical<sub>2</sub> determinism. And, as I shall argue from the point of view of epistemology in sections IV and V, the sheer existence of raw feels is not a good reason for holding an emergence doctrine.

Another philosophical issue which needs careful separation from the scientific problems among the mind-body tangles is that of the "inten-

tionality" of the mental. (For expository reasons the discussion of this issue will be reserved for section IV F.)

### III. Requirements and Desiderata for an Adequate Solution of the Mind-Body Problem. A Concise Statement of the Major Issues

If the title of this section were not already a bit too long, I should have added, "as I view these requirements and desiderata, and as I conceive the adequacy of a solution." All I can say by way of extenuation of my personal biases in this matter is that I have concerned myself seriously and repeatedly with the problem for about thirty-six years; that I have studied most of the contributions from thinkers of many lands in modern and recent philosophy and science; and that this is my fourth published attempt to arrive at an all around satisfactory clarification. There have often been moments of despair when I tried ineffectively to do justice to the many (apparently) conflicting but impressive claims coming from ever so many quarters. It is, then, with a heavy sense of intellectual responsibility and not without some misgivings that I proceed to enumerate the following requirements, desiderata, and considerations which seem to me the conditions (or at least some of the conditions) that may serve as criteria of adequacy for a solution of the problem; a solution that is to be satisfactory from the point of view of contemporary science as well as in the light of modern philosophical analysis. I concede unblushingly that in some respects I share here the attitude of some of the (shall I say, epistemologically not too naive) metaphysicians who have wrestled with the problem and have tried to provide a solution that is *synoptic* in that it would render a just, consistent, and coherent account of all relevant aspects and facets of the issue.

Here, then, is my list of requirements and desiderata (or "conservanda" and "explicanda"):

1. The terms "mental" and "physical" are precariously ambiguous and vague. Hence a first prerequisite for the clarification and the adequate settlement of the main issues is an analytical study of the meanings of each of these two key terms, and a comparative critical appraisal of the merits and demerits of their various definitions and connotations. Due attention will also be given to the (partly) terminological question as to whether to include under "mental" beside the directly experienced

and introspectible also the unconscious states and processes of depth-psychological theories (Freudian or Neo-Freudian). All this will be undertaken in the next section of this essay.

2. In the light of what was said in the preceding section about the scientific (empirical) components of the mind-body problem, an analysis of the mind-body relation is to be sought which does justice to the arguments for the sort of mind-body unity which impresses itself increasingly upon the majority of psychologists, psychophysicists, and psychiatrists of our time. Although the question of evolutionary as well as of logical "emergence" cannot be decided by a priori philosophical considerations, vitalistic and interactionist doctrines appear on empirical and methodological grounds as suspect and undesirable. Just what the alleged facts of parapsychology (telepathy, clairvoyance, precognition, psychokinesis, etc.) may imply for the mind-body problem is still quite unclear. Here too, it seems to me, any speculations along the lines of interactionism are—to put it mildly—premature, and any theological interpretations amount to jumping to completely unwarranted conclusions. My own attitude in regard to the experiments (statistical designs) on extrasensory perception, etc. is that of the "open mind." The book by Soal and Bateman (325) and its discussion by M. Scriven (305) present evidence and arguments which can not lazily or cavalierly be shrugged off. The chances of explaining the "facts" away as due to experimental or statistical error, let alone as outright hoax or fraud, seem now rather remote. But even granting these facts, I think that efforts should be made to explain them first by revisions and emendations in the *physical* theory of behavior before we indulge in speculations about immaterial souls or selves. These remarks clearly reveal my bias in favor of a naturalistic, if not monistic, position. That and how this position differs from "crass materialism," the bugbear of idealistic and spiritualistic metaphysicians, will be explained later on.

3. Any solution of the mind-body problem worth consideration should render an adequate account of the *efficacy* of mental states, events, and processes in the behavior of human (and also some subhuman) organisms. It is not tender-mindedness or metaphysical confusions, I trust, which impel this repudiation of a materialistically oriented epiphenomenalism. Admittedly, the testimony of direct experience and of introspection is fallible. But to maintain that planning, deliberation, preference, choice, volition, pleasure, pain, displeasure, love, hatred, at-

tention, vigilance, enthusiasm, grief, indignation, expectations, remembrances, hopes, wishes, etc. are not among the causal factors which determine human behavior, is to fly in the face of the commonest of evidence, or else to deviate in a strange and unjustifiable way from the ordinary use of language. The task is neither to repudiate these obvious facts, nor to rule out this manner of describing them. The task is rather to analyze the logical status of this sort of description in its relation to behavioral and/or neurophysiological descriptions. In the pursuit of this objective it will of course be necessary to avoid both interactionism and epiphenomenalism; and it will moreover be desirable to formulate the solution in such a way that it does not presuppose emergentism (in the sense of physical<sub>2</sub> indeterminism), although the door to a scientifically formulated emergentism need not be closed.

In this same connection justice should be rendered to what is meaningful and scientifically defensible in the notion of free will or free choice. If our personality-as-it-is at the moment of choice expresses itself in the choice made; if our choices accord with our most deeply felt desires, i.e., if they are not imposed upon us by some sort of compulsion, coercion, or constraints such as by brute physical force, by other persons (or even only by components of our personality which we do not acknowledge as the "core" deemed centrally our "self"), then we are "free" in the sense that we are the doers of our deeds, the choosers of our choices, the makers of our decisions. In other words, it is in this case that our central personality structure is a link in the causal chain of our behavior, predominantly, even if not exclusively, effective in the determination of our conduct. This sort of freedom (in the superb formulation of R. E. Hobart-Dickinson Miller) "*involves determinism and [is] inconceivable without it.*" \*

4. A most important logical requirement for the analysis of the mind-body problem is the recognition of the synthetic or empirical character of the statements regarding the correlation of psychological to neurophysiological states. It has been pointed out time and again † that the early reductionistic logical behaviorism failed to produce an adequate and plausible construal of mentalistic concepts by explicit definition on the basis of purely behavioral concepts. (In the less adequate material

\* Cf. R. E. Hobart (157).

† Cf. F. Kaufmann (175), N. Jacobs (163), C. I. Lewis (196), E. Nagel (230), A. Pap (243), et al.

mode this might be put by bluntly saying that mind is not identifiable with behavior.) For a long time, however, I was tempted to identify, in the sense of *logical* identity, the mental with the neurophysiological, or rather with certain configurational aspects of the neural processes. It was in this sense that I (103) suggested a *double-language* theory of the mental and the physical. But if this theory is understood as holding a *logical translatability* (analytic transformability) of statements in the one language into statements in the other, this will certainly not do. Interlinguistic translations like “*Il ne fait pas beau temps*” into “The weather is not fine” are analytic if the respective meanings are fixed with the help of syntactical and semantical metalanguages common to both French and English. Similarly the geocentric description of the pure kinematics of the planetary system is analytically translatable into the corresponding heliocentric description, precisely because we avail ourselves here of transformation rules in a four-dimensional geometry (i.e., kinematics).

But the question which mental states correspond to which cerebral states is in *some* sense (to be analyzed epistemologically later on) an empirical question. If this were not so, the intriguing and very unfinished science of psychophysiology could be pursued and completed by purely a priori reasoning. Ancient and primitive people had a fair amount of informal and practical psychological knowledge, but the fact that mental states are closely associated with cerebral states was unknown to them. Aristotle held that the seat of our feelings and emotions is the heart (and this has survived in the traditions of poetic discourse). But to say that Aristotle was wrong means that we have now empirical evidence which proves that the emotions are linked to brain processes. It is therefore imperative to preserve the *synthetic* character of the assertion of this knowledge claim, whatever specifically may prove to be its most clarifying formulation.

If any of my readers should be hard-boiled behaviorists or “crass” materialists, it will be difficult to convince them that there is a problem at all. I can do no more than to ask them such persuasive or *ad hominem* questions as, Don't you want anesthesia if the surgeon is to operate on you? And if so, what you want prevented is the occurrence of the (very!) raw feels of pain, is it not? If you have genuine concern and compassion for your fellow human beings (as well as perhaps for your dogs, horses, etc.), what is it that you object to among the con-

sequences of cruel treatments? Is it not the pains experienced by these “others”? It could not be merely their physical mutilation and consequent malfunctioning. Moral condemnation of wanton cruelty presupposes the meaningfulness of the ascription of direct experience to others. Subjective experience in *this* sense cannot be *logically* identical with states of the organism; i.e., phenomenal terms could not explicitly be defined on the basis of physical<sub>1</sub> or physical<sub>2</sub> terms.

It should be noted that we repudiate the *logical translatability* thesis not because of the possibility, definitely contemplated, of a one-many- $\psi$ - $\Phi$  correspondence. One could always formulate such a correspondence with the help of a general equivalence between statements containing single  $\psi$ -predicates on the one side and disjunctions of statements containing several and various  $\Phi$ -predicates on the other. It is rather the *logical necessity* of the equivalence which is here rejected. The equivalence must be construed as logically contingent.]

5. Consonant with the spirit of the preceding discussions, but now to be stated explicitly, are three very closely related *epistemological requirements*. To list them first very briefly, they are:

- (a) the need for a *criterion of scientific meaningfulness* based on *intersubjective confirmability*;
- (b) the recognition that epistemology, in order to provide an adequate reconstruction of the confirmation of knowledge claims must employ the notion of immediate experience as a confirmation basis; (the “given” cannot be entirely a myth!) “Acquaintance” and “Knowledge by Acquaintance,” however, require careful scrutiny;
- (c) the indispensability of a *realistic*, as contrasted with operationalistic or phenomenalist, interpretation of empirical knowledge in general, and of scientific theories in particular.

(ad a) It is generally agreed that scientific knowledge claims must not only be intersubjectively communicable (intelligible), but also intersubjectively testable. The following considerations will illustrate the point. If the stream of my conscious experience continued beyond the death and decay of my body, then this may be verifiable by me (in some, none too clear, sense of “me”; but I shall let this pass for the moment). If such survival were, however, not even extremely indirectly or incompletely confirmable by others; if it were in no way lawfully connected with, and thus not inferable from, any feature of life (mine or that of others) before death, then, while the statement in question may

be said to have subjective meaning, it could not become part of science in the sense in which "science" is commonly understood.\*

(ad b) Recent behavioristic and physicalistic arguments to the contrary notwithstanding, I am still convinced that purely phenomenal statements make sense and are the ultimate epistemic basis of the confirmation (or disconfirmation) of knowledge claims. By this I do not at all wish to suggest that phenomenal statements are infallible ("incorrigible"), nor that they necessarily have a higher degree of certainty than intersubjectively confirmable statements about the ordinary objects of our common life environment. I grant that, especially for the purposes of the philosophy of science, it is more useful to choose the physicalistic thing language for the confirmation basis of knowledge claims. But when I judge, e.g., that a certain pain is increasing, or that I hear a certain ringing sound (no matter whether this sound-as-experienced is causally due to a doorbell, a police car siren, to "buzzing in my ear," or to a hallucination), then that certain *it* which may later find its place in the causal structure of the world is first of all, and taken by itself, a *datum of direct experience*. Whether I get to it "post-analytically," or whether I simply have it, pre-analytically; that is to say, whether I arrive at it by a kind of analysis starting from "seeming," "appearing," "looks like" ("sounds like," etc.) sentences; or whether I can by simultaneous introspection (self-observation) or immediate retro-spection, ascertain the occurrence of a certain datum, I have no doubt that talk about phenomenal data and phenomenal fields makes sense; and that in a rational reconstruction of the confirmation of ordinary observation statements, we can (if we wish) penetrate to this deepest level of evidence.†

I have not been convinced by the arguments of Popper (258) that the search for "hard data" is doomed to failure, that the "given" is like a bottomless swamp. Nor am I convinced that a purely private language ‡ is inconceivable. Of course, if by "language" one means an in-

\* For a fuller discussion of the scientific meaning criterion cf. my articles (103, 105, 109, 110, 114, 116) and Carnap (64, 67, 73). For stimulating discussions of the meaning of "disembodied minds" see Aldrich (6) and Lewy (199).

† For persuasive arguments along these lines, cf. B. Russell (284, 287); H. H. Price (264); C. I. Lewis (195, 197, 198); Ayer (12, 13, 18); N. Goodman (135, 136, 137). For an incisive critique of the "incorrigibility" arguments, cf. K. R. Popper (258); R. Carnap (62, 64); H. Reichenbach (273, 276); M. Black (38); J. Epstein (98).

‡ Cf. the symposium by Ayer and Rhees (16, 278).

strument of interpersonal communication, then the idea of an absolutely private language is self-contradictory. But, granting that in the normal case the capacity for using a language is acquired by education, it is not logically inconceivable that a child growing up in complete solitude might devise his own symbolism not only for the objects and events in its environment but also for the raw feels of its direct experience. Such a child might well come to use terms for various aches, pains, itches, tickles, moods, emotions, etc. I do not for a moment deny that the use of such subjective terms, in the usual and normal case, is acquired through trial and error learning, and in this process largely inculcated in the child by other persons who tell him, e.g., "now you are tired," "now you are glad," "you must have an awful pain." Such tellings by others are guided by the facial expressions, vocal emissions, posture, etc., i.e., generally by the observable behavior of the child (and by test condition → test result sequences in its behavior, involving both environmental stimulus situations and a variety of responses).\*

In sum, I believe that there is an indispensable place for "acquaintance" and "knowledge by acquaintance" in a complete and adequate epistemology. A more detailed account and analysis of the meanings of these terms will be given in the two subsequent sections of the present essay.

(ad c) The last epistemological requirement, to be briefly discussed here, is that of a realistic, rather than phenomenalistic or operationalistic, reconstruction of knowledge. With the current liberalization of the criterion of empirical meaningfulness † the narrower positivism of the Vienna Circle has been definitively repudiated, and is being replaced by a ("hypercritical") realism. No longer do we identify the meaning of a statement with its method of verification. Nor do we consider the meaning of a concept as equivalent with the set of operations which in test situations enable us to determine its (more or less likely) applicability. Instead we distinguish the evidential (or confirmatory) basis from the factual content or reference of a knowledge claim. Early and crude forms of behaviorism identified mental states with their (sic) observable symptoms. Embarrassment might then mean nothing but blushing. But refinements and corrections were introduced in due

\* Cf. Carnap (62, 63); Skinner (320, 321); Wittgenstein (357).

† Cf. Carnap (64, 73); Hempel (149, 151); Feigl (105, 106, 109, 110, 112, 114, 116); Ayer (18); A. Pap (243, 246, 248). Also Grünbaum (139); Feyerabend (119).

course. Mental states were considered "logical" constructions based on observable behavior; and statements about mental states were considered logically translatable into statements about actual or possible behavior, or into statements (or sets of statements) about test conditions and ensuing test results concerning behavior. Mental traits were considered as correlation clusters of their (*sic!*) symptoms and manifestations, and so forth.

But even such a refined or "logical" behaviorism is now rejected as an inadequate reconstruction. It was realized that those behavioral test condition → test result conditionals are to be derived from the laws and postulates regarding central states. Such derivations or explanations have been eminently successful in the physical and in some of the biological sciences. In the atomic theory, or in the theory of genes, for example, it is becoming increasingly possible to derive the macro-regularities, regarding, e.g., chemical compounding, or Mendelian heredity from lawlike postulates and existential hypotheses. The central states of molar behavior theory (or the "factors" in the factor analysis of personality traits) are, however, unspecified as regards their neurophysiological basis. This is comparable to the early stages of the atomic theory when nothing was known about the mass and the structure of individual atoms, or to the early stages of the theory of heredity when Mendel's "units" were not as yet identified with the genes, located and spatially ordered in specific ways, within the chromosomes of the germ cells.

There is little doubt in my mind that psychoanalytic theory (or at least some of its components) has genuine explanatory power, even if any precise identification of repression, ego, superego, ego, id, etc. with neural processes and structures is still a very long way off. I am not in the least disputing the value of theories whose basic concepts are not in any way micro-specified. What I am arguing is that even *before* such specifications become possible, the meaning of scientific terms can be explicated by postulates and correspondence rules (cf. Carnap, 73), and that this meaning may later be greatly enriched, i.e. much more fully specified, by the addition of *further* postulates and correspondence rules.\*

\* For a defense of psychological theory without explicit reference to micro-levels, cf. Lindzey (200). The logic of theoretical concepts in psychology has been discussed in some detail by McCorquodale and Meehl (213); Feigl (113); Cronbach and Meehl (79); Ginsberg (133, 134); Maze (212); Seward (317); Rozeboom (283); Scriven (306).

After the recovery from radical behaviorism and operationism, we need no longer hesitate to distinguish between *evidence* and *reference*, i.e., between manifestations or symptoms on the one hand, and central states on the other; no matter whether or not central states are micro-specified (neurophysiologically identified).

The meaning of scientific statements consists indeed in their truth conditions. But "truth conditions" does not mean the same as "confirming evidence". (The only possible exceptions to this are the directly and completely confirmable singular statements regarding immediately observable situations.) A theory is required to tell us which observations form confirming evidence for scientific statements about matters inaccessible to direct observation. It is in the light of such theories that we can then specify how much support a given bit of evidence lends to a specified hypothesis.

In section V, I shall return to the crucial questions of reduction and identification. There I shall discuss the logical nature of the relation between mentalistically, behaviorally, and neurophysiologically characterized central states.

No elaborate arguments should here be required for a realistic interpretation of the statements about the "physical" objects of everyday life or of theoretical physics.\* In the explanatory context (or the "nomological net") concepts pertaining to the unobservables are related to, but not identifiable with, the observables which constitute the evidential data for the confirmation of statements about the unobservables. For example, spectral lines, cloud chamber tracks, scintillations on screens, Geiger counter indications, etc. are the evidential data which, in a complete logical reconstruction, must be conceived as *nomologically* connected with the aspects of atomic and subatomic particles which they confirm. (Less exciting, but logically analogous, is the analysis of statements of common life about ordinary (partly or wholly observable) objects. Here the perceived perspectives of mountains, trees, clouds, etc., or the instrument indications of air pressure, wind currents, air moisture, etc., are to be interpreted as evidence related to what is evidenced, by the geometrical-optical laws underlying the projections in visual perception, or the physical laws which explain the operation of barometers, anemometers, hygrometers, etc.)

\* Cf. B. Russell (288); R. B. Braithwaite (48); Kneale (179); L. W. Beck (24); Feigl (110, 111, 114).

6. The "meat" of an adequate solution of the mind-body problem will consist in a specific analysis of the characteristics and the relations between the attributes of the mental (especially the phenomenal) and the physical (specifically the neurophysiological). It should be clear from the outset that, if a complete solution of these problems is ever going to be achieved, it will arise out of a combination of the results both of scientific research and of philosophical analysis. In all these questions the two components are so intimately bound up with one another, that neglecting either of them seriously jeopardizes the whole endeavor. The philosophical aspects will be given a further analysis in the next section where I shall try to sort out the various meanings and the attached connotations of the terms "mental" and "physical". The most controversial, tangled and perplexing questions concern, of course, the distinctions made rightly or wrongly in the Cartesian and in the subsequent dualistic tradition between the *mental* and the *physical* in terms of the various alleged criteria listed in the accompanying table.

<i>Mental</i>	<i>Physical</i>
subjective (private)	objective (public)
nonspatial	spatial
qualitative	quantitative
purposive	mechanical
mnemic	non-mnemic
holistic	atomistic
emergent	compositional
intentional	"blind"; nonintentional

Practically all the perennial perplexities of the mind-body problem center around the listed contrasts. The dualists make *prima facie* an excellent showing. The more enlightened monists have always realized that any argument in favor of an identification (in some sense!) of the mental and the physical is faced with serious difficulties. Small wonder then that many of the more sophisticated analytic philosophers of the present age either embrace some form of dualism (usually parallelism), or else declare the issue between monism and dualism a pseudoproblem engendered by logical or terminological confusions. I do not share this outlook. In the following section I shall prepare the ground for an "identity" theory, and I shall present my formulation as well as my arguments in section V.

#### IV. Sorting Out the Various Meanings of "Mental" and "Physical". A Comparative and Critical Analysis

Much of the trouble with the mind-body problem arises out of the ambiguities and vaguenesses of the terms "mental" and "physical". Some of their connotations have been briefly indicated in the juxtapositions listed toward the end of the preceding section. I shall now attempt to analyze these and other meanings more closely, and to point out the merits and demerits of the various actual and possible usages of "mental" and "physical". Philosophers of the modern age clearly differ as to what constitutes the central core or (if there be such clarity!?) the criteria of the mental and the physical. Some philosophers fasten primarily upon one pair of distinctions, others on a different pair as of primary significance.

A. "Subjective" versus "Objective". The juxtaposition of "subjective" and "objective" has been the source of endless and badly confused controversies throughout the ages. There is nevertheless something significant and worth preserving in this distinction. To say that a twinge of pain experienced by person A is "subjective" or "private" to him may simply mean that another person B, observing A's behavior, may *infer* A's pain, but does not have it, i.e. he does not directly experience it. Dentists do not have the toothaches of their patients. In one sense this is clearly analytic (tautological).\* It is analytic for reasons analogous to those which make it self-contradictory to say that I am growing my wife's hair. (Schizophrenics are known to make assertions of this sort.) "I am eating with my wife's teeth" is merely funny, but not self-contradictory. "Dentists always suffer toothaches when their drill comes near the pulpa of their patient's tooth" is synthetic, but empirically false. "I am listening through my wife's ears" if meant literally (not metaphorically) is a border line case, depending on specific detailed interpretation. "I am enjoying Mozart's music exactly as my wife does" is synthetic and may even be rendered as "I have the same musical experience as does my wife." (Remarks about the two meanings of "same" will follow presently.)

The case is a trifle more complex for perception. Two persons sitting next to each other in the concert hall are said to hear the same music,

\* This is now even admitted by Ayer (18) who had earlier (15) held it was synthetic. His earlier position was, however, incisively criticized by Pap (243, 248) and Wating (341).  
*is right about qualities and dispositions*

or at a given moment the same tones or chords, produced by the pianist on the stage. But the facts of the case are really not fundamentally different from the first example. A does not have B's musical experience (or vice versa), even if their auditory discrimination, musical appreciation, etc., does not differ in any discernible way. They may be said to hear the same sounds, to be both equally impressed or thrilled by them; but common sense as well as scientific reasoning clearly indicates that their *experiences* are numerically different. Fundamentally this case does not differ from, e.g., the case of two thermometers immersed next to each other in the same glass of water. It is perfectly proper to say that these instruments indicate the *same* physical condition. It is also perfectly proper to say that the two thermometers not only indicate *but also "have" the "same" temperature.* (This is logically quite like saying that two marbles have the same color.) But it would be most improper and paradoxical to say that the events taking place in the one thermometer are *identical* with those in the other. This is not the place for a discussion of Plato's problem of the "one and the many." Suffice it to point out that the phrases "the same as" and "identical with" are ambiguously used. "Sameness" or "identity" may mean complete similarity, as in the case of the two musical experiences, or in the case of the two thermometric indications. But "sameness" or "identity" in other contexts means the numerical oneness of the individual referent of, e.g., two different names, or of two different unique characterizations (Russellian descriptions). I conclude then that it makes perfectly good sense to speak of the subjectivity or privacy of immediate experience. *Numerically* different but *qualitatively* identical (indistinguishable) experiences may be had by two or more persons, the experiential events being "private" to each of the distinct persons.

Terminological trouble, however, arises immediately when we take a scientific attitude toward direct experience and try to confirm, describe, or explain it "objectively." Is it not an "objective" fact of the world that Eisenhower experienced severe pain when he had his heart attack? Is it not a public item of the world's history that Churchill during a certain speech experienced intense sentiments of indignation and contempt for Hitler? Of course! What is meant here is simply that statements about facts of this sort are in principle *intersubjectively confirmable* and could thus be incorporated in a complete historical account of the events of our universe. To be sure, there are cases in which con-

firmation is *practically* outright impossible. The last thoughts and feelings of a man immediately before his death, especially in a case of complete paralysis, or of death occurring through electrocution, may be inferable only with scant reliability. But this is not different from the case (cf. Carnap, 67, p. 419f) of the confirmation of the electric charge of a specific raindrop that fell into the Pacific Ocean in a place far removed from any observers. Our current liberal formulation of the empiricist meaning criterion countenances all statements of this sort as perfectly meaningful. They do not fundamentally differ from other less difficult-to-confirm statements about, e.g., the "true thoughts" of a liar or play actor. Modern devices, such as the lie detector, and various clinical-psychological techniques enable us to test for such "private" events with increasing (though generally only relatively low) reliability.

The foregoing considerations suggest that the terms "subjective" or "private" at least in one of their commonly proper and serviceable usages are not to be considered as logically incompatible with "objective" or "public" in the sense of "in-principle-intersubjectively confirmable". Private states in this philosophically quite innocuous sense are then simply *central* states. (Whether these are ultimately to be conceived mentalistically or neurophysiologically may be disregarded for the moment; but this will of course be discussed quite fully later.) "Subjective" or "private" in this sense may then designate the referents of direct introspective reports, and it will be understood that these same referents may well be more indirectly characterized by descriptions involving inference from behavioral symptoms or test results of experiments on behavior. In those cases of subhuman animal behavior in which we don't hesitate to speak of experienced pains, gratifications, rage, expectations, etc., there are of course no introspective reports. But other aspects of such behavior are in many respects so similar to the human case that the ascription of raw feels is usually justified on the basis of analogy. Here again, the "private" means the central state which causally effects (or at least affects) the overt and publicly observable behavior.

The terms "subjective" and "objective" are indeed mutually exclusive if they are used in a quite familiar but different way. In designating some impressions, opinions, beliefs, value judgments, etc. as "subjective," we sometimes contrast them with the "objective truth," or "objective reality." If, e.g., my friend maintains that the room is cold,

I am inclined to argue with him by pointing to the thermometer (which reads, say, 74°); and perhaps by explaining his "impression" by the fact that he is too scantily dressed, or that he is sick, or suffers from anxieties, etc. Similarly in the more drastic cases of dreams, illusion, delusion, etc. we criticize some (interpretive) judgments as based on "merely subjective" evidence. And it should go without saying that disagreements in aesthetic value judgments may often be explained on the basis of individual or cultural differences. "De gustibus non est disputandum" is our final resort if no objectively justifiable standard can be agreed upon.\* But wherever beliefs can be criticized as, e.g., "biased," "too optimistic," "too pessimistic," etc., there are standards, such as those of normal inductive inference, which may indeed justify the rejection or correction of such "all too subjective" convictions. Here "subjective" and "objective" are indeed incompatible, although of course there may well again be an "objective" explanation of the genesis of "subjective beliefs."

There is, however, also a philosophical and speculatively extended sense of "subjective" or "private". In this very special and highly problematic sense it is assumed that there may be subjective states which are in principle inaccessible to intersubjective confirmation. Here we had better speak explicitly of "absolute subjectivity" or "absolute privacy." It is this sense which is entertained in some of the more radically interactionistic forms of dualism. And it is this sense which by definition is incompatible with "objectivity" understood as intersubjective confirmability. [As I have indicated before, I no longer insist that a doctrine involving the notion of absolute privacy is entirely devoid of cognitive meaning. But I am inclined to regard it as scientifically meaningless. To recapitulate: if the scientific enterprise is defined as necessarily requiring intersubjective confirmability of knowledge claims, then this follows immediately and quite trivially.

Now, I think it is an essential aspect of the basic working program and of the working hypotheses of science that there is nothing in existence which would in principle escape intersubjective confirmation. Allowances have already been made for the (sometimes) insuperable practical difficulties of even the most incomplete and indirect confirmations. But the optimistic outlook that inspires the advance of

\* On the meaning and the limits of the justification of norms, cf. my essay (109).

science and informs its heuristic principles,\* does not tolerate the (objectively) unknowable or "un-get-at-able." No matter how distant, complicated, or indirect the connection of scientific concepts with some (intersubjective) evidential bases may be, they would not be concepts of empirical science (as contrasted with the concepts of pure logic or mathematics) unless they could in some such fashion be "fixed" by "triangulation in logical space." The "fix" we are able to obtain may be as indefinite as it is when theoretical concepts (like those of the positron, the neutrino, or the meson in physics; that of the unit of heredity; or of memory traces; of the superego, of general paresis, or of schizophrenia in biology, psychology, or psychiatry) were first tentatively introduced by only very sketchily formulated postulates. The concepts of absolutely subjective or completely private data, however, are so conceived that they can be applied only on the basis of the direct experiences contained in a given stream of consciousness. A completely "captive mind" † might experience senselike qualities, thoughts, emotions, volitions, etc., but they would (*ex hypothesi*) not in any way, i.e., not even through weak statistical correlations, be connected with the publicly observable behavior or the neurophysiological processes of an organism.

While it is difficult to spin out this yarn in a consistent (let alone plausible) fashion, I do not think it impossible, in the sense that it would necessarily involve some self-contradictions. There are philosophers who have been concerned with an analysis of the meaning of the "continuance of a pure (immaterial) stream of experience after bodily death"; or with the problem of the "inverted spectrum" (Could pure sensory-like qualia like red and green, blue and yellow, be systematically interchanged for different persons, despite a complete similarity in their discriminatory and linguistic behavior, as well as in their neurophysiological processes?). Speculations of this sort were declared taboo and absolutely meaningless by the early logical positivists. They were compared with assertions about absolute space and time, the (Lorentzian) ether, the "bond" between cause and effect, or the existence of a metaphysical substance, over and above anything that could be verifiably known by science about spatio-temporal relations, coordinate

\* Some philosophers rather speak of them as "metaphysical presuppositions"; for my criticisms of this interpretation of science cf. (110, 114).

† The idea and the phrase are Hilary Putnam's.

transformation, functional relations between observable properties or measurable magnitudes, or relations of compresence of various observable properties. There is no doubt that this positivistic cleansing of the Augean stables of metaphysics had a most salutary effect. But positivists (temperamentally often negativists), in their zeal and eagerness to purge the scientific enterprise of meaningless as well as superfluous elements, have often overshot their goal. A redressing of the balance has become necessary, and we pursue nowadays responsible analyses of, e.g., causal necessity\* which are perfectly compatible with the basic antimetaphysical insights of Hume. Similarly as regards the notion of absolute privacy, it is illuminating to conceive it at least as a logical possibility, and then to state as clearly as feasible the reasons which can be adduced for rejecting the idea for our world as we have come to conceive of it in our science to date.

The notion of absolutely private data of experience, if such data are to be described, would require a purely phenomenal or absolutely private language. Such a language, by definition and *ex hypothesi*, could not serve as an instrument of communication. Even a completely solitary humanlike individual could not engage in audible (or visible, etc.) symbolic activities. Not even soliloquies in this physically expressible form would then be possible. For ordinary soliloquies, amounting to more than the unexpressed thoughts of a private thinker, are expressible, and the very expressions would provide (no matter how unreliable) clues to the "inner" thought processes.†

Now, of course, if by a "language" one means what is customarily meant by it (viz., an instrument or vehicle of intersubjective communication), then an absolutely private language is ruled out by definition. Language as we know it and use it is indeed not absolutely private in the sense explained. But that it is intersubjective reflects a basic empirical feature of our world, or at least a basic feature of our-world-as-we-conceive-it in common life and in science. But I must postpone discussion of the fuller implications of this feature until I present my dénouement of the "world knot" in the final section. For the present I submit that by a "language" one is not compelled to mean an instrument of interpersonal communication. The idea of the soliloquy (*intra-*

\* Cf. Burks (59); W. Sellars (312, 313, 314).

† For an extremely lucid and succinct discussion of this point cf. P. E. Meehl (219).

personal communication) may be restricted and modified in such a manner that it refers to unexpressed and inexpressible thoughts. This preserves a sufficient "family resemblance" with the ordinary notion of language. Such an absolutely private language would still enable the solitary thinker silently to label the qualities of his direct experience and to think silent thoughts which have the logical form of declarative (singular, universal, etc.) statements. I could, for example, with the help of remembrance, think that extreme anger always gradually subsides, that a given tone-as-heard is increasing in intensity, etc. Knowledge thus formulated in a private language may well be called "knowledge by acquaintance." It is true that ordinary discourse entertains a much wider conception of knowledge by acquaintance. There it covers knowledge based on, and not essentially transcending, the observations (amplified by very moderate and limited inferences). Thus we can quite properly say that we know the properties of sticks and stones, of apples and oranges, the manners and mannerisms of our close friends "by acquaintance."

But this ordinary concept of acquaintance is not very sharply defined. Having actually seen Winston Churchill for a few seconds (when on July 10, 1954, he emerged from 10 Downing Street in London and entered his black limousine, holding his cigar and waving to the assembled small crowd), am I entitled to say that I know him "by acquaintance"? Would I know Churchill "by acquaintance" if I had seen him (or rather his image) only in the cinema newsreels? I leave it to the linguistically more sensitive and subtle Oxford analytic philosophers to decide these questions, or else to tell me that "knowledge by acquaintance" is a hazy notion, involving "slippery slopes" in various directions. (Anyway, the latter alternative is what I consider the best analysis of the ordinary usage of the term.)

For a philosophical usage of the term, however, I suggest that "knowledge by acquaintance" be understood as knowledge involving no inferential components—or, if this be chimerical, then knowledge involving only that minimum of inference which is present when only memory is utilized for the recognition of similarities and differences. It is in this sense that I could assert on the basis of acquaintance, "Ah, there is that peculiar smell again; I don't know what causes it, I don't even know how to label it; it is so different from any fragrances of flowers, perfumes, cigar smoke, burnt toast, tangerines, etc. that I can't even

place it in a multidimensional scheme of the rank orders of smells; but I know I have experienced this smell before and I am (subjectively) sure I would recognize it in the future if I were to experience it again."

As I have said earlier, I make no claim for the infallibility of knowledge by acquaintance. Our world, being what it is, is such that corrections of subjective-experience judgments (knowledge claims made on the basis of direct acquaintance) are definitely possible from the vantage point of intersubjective observation. Moreover, it should require no reminder that I quite emphatically want to distinguish *acquaintance* from *knowledge by acquaintance*. "Acquaintance as such" (in the philosophically restricted sense) is to mean simply the direct experience itself, as lived through, enjoyed, or suffered; *knowledge by acquaintance*, however, is propositional. Knowledge claims of any sort may be valid or invalid; the statements which formulate such knowledge claims are either true or false. In the case of practically all knowledge claims which have scientific status, the confirmation of their truth is incomplete and indirect. Knowledge by acquaintance, however, is direct and complete in the following sense: it seems utterly inappropriate to ask someone what his evidence is for asserting that he, e.g., feels at the moment elated, depressed, anxious, dizzy, hot, cold, and so on through the various modalities and qualities.

The philosophically much misused and over-exploited term "self-evident" might well be redefined and restricted to just such reports of immediate introspection or self-observation. With this, possibly unwise, terminological suggestion I do not wish to imply any doctrine of "in-corrigeability" in regard to such protocols of immediate experience. I grant that even such protocol statements may be in error; and not only for the generally admitted reasons such as possible slips of the tongue or the pen; but also because the predicates or relational words used in such statements, if they are what they are intended to be, viz. universals, presuppose for their correct application even in the "absolutely private" language (as fancied above) at least the reliability of memory. This alone would ensure that the same term is applied to an experienced quality of the same kind as before. Otherwise a protocol statement would simply amount to what would in effect be a first introduction of the predicate in question by stipulative-ostensive definition; \* i.e., it

\* The notion of "ostensive definition" is of course highly problematic. In contradistinction to what "definition" (explicit, contextual, recursive, abstractive, condi-

would amount to the resolution to use the same term on future occasions sufficiently similar to the present one. But on the occasion of the first use of a new term, the sentence containing it would be true only in the extremely restricted (very much like analytic) sense that "A," the label which I arbitrarily apply to the completely and incomparably new fragrance that I am just experiencing, designates the quality experienced during each of the moments of its temporary occurrence of finite duration.

There are other uncertainties besides the ones mentioned in the use of (available) predicates for the qualia of immediate experience. Am I to describe the way I feel at a given moment as "happy", "joyous", "merry", "gay", "frolicsome", "blithe", "debonair", "light hearted", "buoyant", "bright", "animated", "gleeful", "hilarious", "jolly", or what?

It is time to draw some conclusions from this discussion. There is one meaning of "mental" in which it coincides with one meaning of "subjective". Let us call this meaning "phenomenal". In so calling it we may leave for later the question as to whether what is phenomenally given and phenomenally labeled is always also indirectly characterizable in an intersubjectively meaningful terminology. In any case we have isolated one contrasting (though not necessarily incompatible) pair of meanings for "mental" and "physical": the phenomenal (i.e., the subjectively confirmable) and the intersubjectively confirmable (i.e., the physical<sub>1</sub> in the terminology suggested above). The meaning of "mental" (synonymous with "phenomenal") looms large in introspective and phenomenological psychology. It is also prevalent in Gestalt-psychological descriptions of the configurations in phenomenal fields.

But in the "depth-psychological" statements of the psychoanalytic schools of thought, "mental" includes also subconscious, and some unconscious, states and processes. Since these are described largely with the help of metaphors and similes taken from the phenomenal (disregarding here those from the physical, e.g., mechanical, hydraulic, etc.) sphere, and inasmuch as detailed neurophysiological descriptions are (partial, coordinative, or even implicit) generally means, ostensive definitions cannot be rendered in speech, writing, or printing. "Definition" in its normal use always means specification of the meaning of some symbol by recourse to the meanings of other symbols. "Ostensive definitions" (if this phrase is to be retained at all) had therefore better be regarded as the establishment or acquisition of a linguistic habit, the inculcation of a bit of rule-governed linguistic behavior. In an absolutely private language it may amount to the stipulation of a rule which associates certain thoughts or images with specific other items or aspects of direct experience.

still lacking, it will be well to remember that the word "mental" as commonly employed by present-day psychologists covers both phenomenal and non-phenomenal states and events. The justification for the inclusion of the subconscious ("preconscious") and the unconscious in the realm of mind comes of course from some other attributes traditionally considered as criteria of mentality. We shall turn to those other attributes. The one which (for philosophical-historical reasons) will be taken up first is, however, not as essential in this connection as are some of the others further down the list.

B. *Non-Spatial versus Spatial*. The Cartesian distinction of *res cogitans* and *res extensa* still provides some philosophers of our age with what they consider one of their most powerful arguments in favor of a radical dualism. Mental states and events in contradistinction to physical bodies, so they claim, have neither a location, nor are they characterizable as having shapes or sizes. The apparent plausibility of this doctrine seems to me to derive mainly from (1) a confusion, and (2) inattention to phenomenal spatiality and its relations to physical spatiality. The confusion becomes evident in rhetorical questions asked by dualists, such as "where is the feeling of motherly love located?" "how many inches is it long?" "is it square or pentagonal?" I must confess I have little patience with these silly games. The feeling of motherly love is a universal, an abstract concept, and it makes as little sense to ask about its spatial location as it does in regard to the (physical) concept of temperature. We have here a category mistake of the crudest sort, a confusion between universals and individuals. It makes sense to ask about the location of individual things or events, but it is simply nonsense to ask about the location of a concept (properties or relations in abstracto).

The same sort of nonsense arises if, after hearing the sentence "the mental depression finally left him," someone asks, "Where did it go?" This sort of question can come only from taking the initial (metaphorical) statement as literally as we take "his wife finally left him." Concepts, whether they designate occurrent or dispositional properties, do not as such have spatial location; or rather it makes no sense to ascribe any such to them. But concepts which are constituents of singular (specific descriptive) statements\* are applied to individuals. We

\* I.e., sentences containing proper names or coordinates.

say "Anthony Eden felt depressed after the failure of the Egyptian campaign." In this case there is quite clearly a location for the feeling of depression. It is in the person concerned! The question of location becomes then more sensible, but logically also more delicate, if we ask it of individual mental states.

Using "mental" for the time being in the sense of "phenomenal", we had better—and without too much ado—introduce the indispensable distinction between phenomenal space(s) and physical space. I am perhaps not too acute in matters of phenomenological description but it does seem to me that my feelings and emotions pervade large parts of my body-as-I-experience it. William James has given us some striking illustrations of this. In the phenomenal field of the subject, specific feelings may be located at least vaguely or diffusely in some not very sharply delimited part of the organism. My feelings or sentiments of elation, depression, delight, disgust, enthusiasm, indignation, admiration, contempt, etc. seem to me to be spread roughly through the upper half or two-thirds of my body.

Sounds and smells, at least in the usual situations of "veridical" perceptions seem to be partly outside, partly inside the phenomenal head. Colors are usually perceived as surface qualities of extradermal objects, or in the case of looking at the skin of one's own arms or legs, as surface qualities of those limbs. Colors seen when pressing one's eyelids (closed eyes) are vaguely located either immediately in front of one's eyes, or even inside them. Similarly musical sound images (especially in the eidetic's case) appear either inside one's head or seem to come from the outside as in a concert hall. The taste of an apple is clearly experienced within the mouth. The stars as seen on a cloudless night are tiny bright spots on a fairly distant dark background. These bright spots clearly have spatial relations to one another. A given small portion of the sky-as-perceived is an approximately plane surface with the twinkling stars distributed in certain constellations. If for the moment we may use the names of the stars as proper names—for the bright spots in the visual field, we may well say that, e.g., Sirius is to the left and far below the three stars of Orion's belt. There is no question then that we are "acquainted" with the elements and relations in visual space.

A detailed discussion of the relations of visual, tactual, kinesthetic, and auditory "spaces" among each other is a task of phenomenal psychology. For our purposes it is sufficient to notice that "spatiality"

means qualitatively quite different things for the various sense modalities. But *physical space*, in the sense in which the science of physics (including, of course, astronomy) understands it, is something radically different. The astronomers' measurements and inferential interpretations have provided us with an account of the three-dimensional array of the stars in "objective" space. This three-dimensional order is most properly considered as a conceptual system which can be only inadequately visualized or imaged phenomenally. I don't for a moment deny that in our rooms or in a landscape we perceive directly at least some of this three-dimensional order. (In the case of the stars, we don't.) But what is present in perception at any given moment is always a particular perspective and not the geometrical order which we must assume (together with certain laws of geometrical optics) in order to explain the peculiarities of any (or all) particular perspectives.

I shall not labor the obviously analogous case of time. Phenomenal time and physical time differ from, and are related to, each other very much like phenomenal space and physical space. Experienced durations may seem very long in the case of tiresome waiting, while time packed full with exciting events seems to "pass quickly." But the physically measured durations may be exactly the same. The *psychological* relativity of (phenomenal) time must of course not be confused with the (Einsteinian) *physical* relativity of simultaneity and duration which, in the nature of the case, is not directly observable at all.\*

We conclude then that *mental data* have their own (phenomenal) kinds of spatiality; and that *physical space* is a theoretical construction introduced to explain the features and regularities of phenomenally spatial relations. The exact and detailed derivation, even only of the perspectival aspects of visual spatiality is a quite complex matter, involving geometrical, physical, psychophysical, and psychophysiological laws. Our arguments have so far disproved only the Cartesian contention that the mental is non-spatial. To put it very strongly, mental events as directly experienced and phenomenally described are spatial. Physical bodies geometrically characterized in their measurable positions, orientations, shapes, and sizes are not spatial (in the *visual*, or generally,

\* Except, of course, for such cases as the traveling and returning twin brother, which, though strictly implied by the well-confirmed principles of Einstein's theory, has not been susceptible to direct check thus far (because of obvious practical difficulties).

*phenomenal*) sense at all. "Space" in the physical sense is an abstract theoretical ordering system. The reader who accepts my arguments may nevertheless maintain that the emphasized distinction between phenomenal and physical spatiality (and temporality) reaffirms all the more convincingly the dualism of the mental and the physical. My rebuttal of this contention will be given in the concluding sections. Suffice it here to suggest that if by "physical" we do not understand a kind, type, part, or aspect of reality, but rather a method, language, or conceptual system, then there is no room for a dualistic opposition of mental and physical events or processes, let alone substances.

C. *Quality versus Quantity*. Another time-honored distinction between the mental and the physical is made in terms of the qualitative and the quantitative. This distinction also is fraught with the danger of various confusions. A prima facie plausible argument maintains that, e.g., the qualities of colors-as-experienced, sounds-as-heard, odors-as-sensed, heat-intensities-as-felt, etc. are undeniably and fundamentally different from the quantitatively measurable wave lengths of light radiation, the frequencies and energies of sound waves, the chemical compositions of odorous substances, the mean kinetic energies of the molecules, etc. Of course, they are. But the argument misses the essential point. What the physicist measures are quantitative aspects of stimuli or stimulus patterns. These stimuli produce, under certain ("normal") circumstances, certain qualitatively characterizable sensations within the phenomenal fields. The familiar freshman's question, "Is there a sound when on a lonely island, with neither men nor beasts present, a tree falls to the ground?" is quickly clarified by the distinction between the sound waves (vibrations in the air) and sounds-as-heard. The dualistic argument would, however, be strictly to the point if it concerned the distinction between the sense-qualities-as-experienced and the "correlated" cortical processes in the brain of the experiencing subject. These cortical processes could be quantitatively described in a completed neurophysiology. Various more or less localized patterns of nerve currents ("firings" of neurons, etc.) would be the object of a "physical" description. Just which phenomenal qualities correspond to which cortical-process patterns has to be determined by empirical investigation. In our previous discussion of "conservanda" and "explicanda" we have not only admitted, but insisted upon, the synthetic character of the statements which formulate these correlations. Reserv-

ing fuller arguments for monism again for the final sections, a few preliminary critical observations are in order at this point:

(a) Purely phenomenal descriptions are generally not restricted to a merely qualitative form. Semiquantitative or rank-ordering ("topological") descriptions are possible at least among the qualities within each modality of experience. "My pain is increasing"; "this (sensed) blue is darker than that"; "my embarrassment was worse than any I had ever felt before"—these examples illustrate semiquantitative singular statements. Universal statements of this form can also be made, e.g., "Purple is more bluish than scarlet." "D is higher in pitch than C." Universal statements of this sort can be organized in topological arrays of one, two, three (or more) dimensions, as in the tone scale, the color pyramid, the prism of odors, etc. Moreover, there are cases of remarkable intersubjective agreement even in purely introspective judgments of the *metrical* relations of given qualities or intensities among each other. S. S. Stevens,\* for example, found by careful experimentation that subjects agreed on what was the mid-point in a series of sounds of varying intensities. Shapes, sizes, distances, durations—all-as-directly-experienced are often susceptible to metrical estimates far surpassing in accuracy anything the uninformed might ever expect.

As regards the differences among such experiential modalities as colors, sounds, and smells, or between larger classes such as the sense qualities and the emotions, it must of course be recognized that they differ qualitatively from one another; and no merely quantitative distinction will serve as a criterion to characterize their different generic features. Dualists have tried to utilize this as an argument by asking, Why should there be more than one basic quality (or modality, for the matter of that), if all of the manifold phenomenal data are to be nothing but the subjective aspects of basically homogeneous brain processes? But the answer may well be that there are sufficient topographical, configurational, and quantitative differences even among those "homogeneous" neural processes.

(b) The magnitudes determined by physical measurement, and syntactically represented in scientific language by *functors*,† differ among each other in a way that can hardly be called anything but "qualita-

\* Cf. his article in the *Handbook of Experimental Psychology* (S. S. Stevens, ed.). New York: Wiley, 1951.

† Cf. Carnap (65, 68); Reichenbach (274).

tive". What else can we say about the differences between, e.g., mass, temperature, pressure, electric current intensity, electromotoric force, gravitational field intensity, etc.? What is it that is, respectively, indicated by thermometers, manometers, ammeters, voltmeters, etc.? I think it is entirely justifiable to speak of these scientific variables as *qualitatively* different. To be sure, they are not directly experienced qualities. But is there any good reason for restricting the term "quality" to the phenomenally given?

I conclude that the attempt to define "mental" and "physical" in terms of the distinction qualitative-quantitative begs the question. It makes perfectly good sense to speak of mental quantities and of physical qualities.

D. "Purposive" versus "Mechanical". Along with *direct experience*, it is perhaps *intelligence* which makes up the most important characteristic of the commonsense concept of mentality. And intelligence is usually and most basically characterized as the capacity of utilizing means toward the attainment of ends. One trouble with this characteristic is that common language is apt to describe as "intelligent" even the instinctive behavior of many animals. In the case of, e.g., social insects (termites, ants, bees, etc.) the behavior is stunningly purposive, highly organized, and intricate; and yet we hesitate to ascribe sentience or subjective experience (raw feels) even only remotely resembling our own to these entirely different organisms. Moreover, the current scientific use of the word "intelligence" tends to be restricted to those evolutionary levels and species in which learning combined with ingenious (inventive) and symbolic behavior plays a dominant role. Pigeons, rats, cats, dogs—those favorite laboratory animals of the behavioristic psychologists—show (in each species) marked individual differences in the speed and the scope of their learning. Anthropoid apes, like the chimpanzees, are famous (ever since W. Köhler's original experiments) for their inventiveness—in addition to their commonly known capacities for imitation. Genuinely *linguistic* behavior, involving syntactical, semantical, and pragmatic features, seems to be restricted to *homo sapiens*; the so-called language of the bees (which is apparently instinctive and lacking in syntactical and semantical flexibility) does not seem to be an exception.

If intelligence or just purposiveness were chosen as the sole criterion of mentality, then it would be hard to draw a sharp line anywhere

within the realm of organic life. Even in the kingdom of plants we find processes whose teleological characteristics are not fundamentally different from the features of purposive behavior in the lower animals. Of course, if one deliberately makes the (often suggested and no doubt helpful) distinction between two types of teleology, one of them involving *conscious* aims, and the other excluding them, and designates only the former as "purposive," then the empirical evidence suggests (but does not force upon us) the decision to call "intelligent" only the behavior of the higher animals, or perhaps to restrict the label "intelligence" to human beings (i.e., if and when they behave in a genuinely sapient manner).

It becomes clear then that the scope of the two criteria (*sentient* and *sapient*) is not necessarily the same. The two concepts are not co-extensive. The situation has been further complicated in our age by the construction of "intelligent" machines. Logical reasoning, mathematical proofs and computations, forecasting, game playing, etc. are all being performed by various and usually highly complex electronic devices. Here the temptation to ascribe "raw feels" becomes even weaker than in the case of the lower animals.\* Inductively it is plausible that sentience requires complex organic processes.

Descartes was perhaps not completely wrong in restricting mentality to human beings. If "mind" is understood as the capacity for reflective thought, then indeed we may have reason to deny minds (in this sense!) to animals (and perhaps even to electronic computers!). The issue is difficult to decide, because the connotations of "reflective thought" are numerous and indefinite. But if it connotes a *conjunction* of sentience, learning capacity, spontaneity (free choice), purposiveness (in the sense of goal directedness), original inventiveness, intentionality (in the sense of symbolic reference), and the ability to *formulate* rules of behavior (practical, moral, linguistic, etc.), then *mind* (in this sense) is clearly the prerogative of man.

All the foregoing considerations need not disturb us. They merely lead to the scarcely surprising conclusion that the term "mental" in ordinary and even scientific usage represents a whole family of concepts; and that special distinctions like "mental<sub>1</sub>", "mental<sub>2</sub>", "mental<sub>3</sub>",

\* Cf. however, the remarkable and stimulating discussion of the robot problem by Scriven (304). We shall return to this issue in connection with the scrutiny of the analogy argument for "other minds" in section V.

etc. are needed in order to prevent confusions. (We shall return to a brief discussion of "intentionality" in subsection F.)

As far as the original distinction of purposive versus mechanical is concerned, it scarcely helps in the definition of the mental versus physical distinction. If "purposive", despite our warnings, is taken as synonymous with "teleological", then we have a distinction, which, though it becomes rather irrelevant to the mental-physical issue, is not useless in the natural sciences and in technology. But then it can no longer be considered as either sharply exclusive, nor as particularly enlightening. The flow of a river toward the sea is a mechanical and non-teleological phenomenon, but the functioning of servomechanisms is mechanical as well as teleological, and the functioning of the heart is teleological and presumably "mechanical" in the same (wider) sense in which complex servomechanisms operating by negative feedback are regulative physical devices. In short, the phrase "teleological mechanisms", in our age of cybernetics is no longer a contradiction in terms.

E. "Mnemonic", "Holistic", "Emergent" versus "Non-Mnemonic", "Atomistic", "Compositional". This bundle of contrasts has often been associated with the distinction of the mental and the physical. Fortunately, except for one facet of the emergence issue, discussion can be quite brief. The *mnemonic* as a criterion of mind was stressed especially by Bertrand Russell. But long before him, the physiologist Ewald Hering (and his disciple Semon) considered the mnemonic as a general property of all organic matter. Even in inorganic matter there are more or less permanent modifications of dispositional properties which can be effected by various influences. Certain features of elasticity and of magnetic hysteresis are "mnemonic" in this sense. And of course the storage of information in present-day computing machines clearly shows that mnemonic features, just as the "purposive-intelligent" features, need not coincide with mentality in the sense of sentience or awareness.

The holistic aspects of the phenomenal fields were brought to the fore by the Gestalt psychologists. But almost from the beginning, this school of thought (especially ever since W. Köhler's book on *Physical Gestalten*, 1920) emphasized the idea of the isomorphism of phenomenal with neurophysiological configurations. Thus again, without the addition of the criterion of immediate experience we do not obtain a distinction between the mental and the physical configurations or "organic wholes" or "dynamic Gestalten."

Inseparably connected with holism and the Gestalt philosophy is the doctrine of *emergence*. The old slogan "the whole is greater than the sum of its parts" has of course no very clear meaning. Much of its obscurity is due to the lack of a definition of the phrase "the sum of the parts". Recent analyses \* of the still controversial significance of "organic wholeness" and of "emergent novelty" have contributed a great deal to the clarification of the issues. There is no imperative need for us to enter into details here. It will be sufficient for our concerns to realize that in modern natural science no sharp distinction can be made between *resultants* (as in the composition, i.e. vectorial addition of forces or velocities) and *emergents*. In the explanation of the properties and the behavior of complexes and wholes we always need laws of composition—be they as simple as the straightforward arithmetical addition of volumes, masses, electric charges, etc., or slightly more complicated as is vector addition, (or just a trifle more involved as is the relativistic "addition" formula for velocities), or extremely complex as are the so far not fully formulated composition laws which would be required for the prediction of the behavior of organisms on the basis of a complete knowledge of their microstructure and the dynamic laws interrelating their component micro-constituents.

Modern quantum physics, on a very basic level, employs laws which have "organismic" character, as for instance the exclusion principle of W. Pauli † which holds even for single atoms. It is conceivable that much of what is called "emergent novelty" on the chemical and biological levels of complexity may ultimately be explained in terms of the organismic or holistic features of the laws of atomic and molecular dynamics; and that, given those basic micro-laws, the only composition laws (which scientists often take for granted like "silent partners") are simply the postulates and theorems of geometry and kinematics. This is indeed my own, admittedly risky and speculative, guess; that is to say, I believe that once quantum dynamics is able to explain the facts and regularities of organic chemistry (i.e. of non-living, but complex compounds) it will in principle also be capable of explaining the facts and regularities of organic life. But no matter whether these conjectures

\* Schlick (299); Nagel (232, 235); Henle (153); Bergmann (28, 34); Hempel and Oppenheim (152); Rescher and Oppenheim (277); Pap (244).

† Cf. the clarifying discussion by Margenau (208); and the stimulating, but perhaps somewhat speculative, ideas of Kaila (169).

prove correct or incorrect, emergent novelty from a logical or methodological point of view simply means the impossibility of the derivation of the laws of complexes ("wholes") from the laws that are sufficient to predict and explain the behavior of their constituents in relative isolation. Thus, the laws that are sufficient to account for the motion of free electrons (as in cathode rays, and traversing electric or magnetic fields) are clearly insufficient to account for the behavior of electrons when they are constituents of atoms.

It stands to reason, that in order to "glean" (i.e., to ascertain) the laws of nature, scientists can't afford to stop their investigations on a very low level of complexity. In some cases we are lucky in that from such a very low level of complexity upwards to higher complexities of any degree, no new physical laws (but only geometrical composition laws) are required. This holds, for example, for the law of the lever which remains applicable even for the most complex system of pulleys. It also holds for the law of gravitation and the laws of motion (both in their Newtonian form). The "many bodies problem" is unsolved only in the mathematical sense that no single set of simultaneous equations has as yet been found for the prediction of the motions in complex star systems. But successive approximations can be computed to any desired degree of accuracy. In other cases (as with the behavior of electrons) we could never glean all the relevant laws below a certain level of complexity. And I have admitted (in section II) that it is always *logically* conceivable that our scientific theories may have to be amended and enriched by the introduction of new basic concepts (variables), and this is of course tantamount to the introduction of new (lawlike) postulates and/or existential hypotheses.

We have seen that the mnemonic, teleological, holistic, and emergent features are not adequate as criteria of mentality, because these features characterize even inorganic structures and processes. Emergence as conceived by most dualists, however, refers to the evolutionary novelty and the (physical<sub>2</sub>) underderivability of *sentience* or *raw feels*. The whole issue therefore turns again upon the criterion of *subjective experience*. The issue can be brought out by questions such as the following: Suppose we could predict the detailed chemical structure of an entirely new perfume which will be manufactured in Paris in the year 1995. Suppose, furthermore, that we could equally exactly predict the neurophysiological effects of this perfume on the mucous membranes of a human nose,

as well as the resulting cortical processes in the person thus smelling the perfume. Could we then also predict the quality of the experienced fragrance? The usual answer to this question is in the negative, because it is assumed that the fragrance in question will be an "emergent novelty." But behaviorists, and physicalists generally, need not take such a pessimistic view. For given the presuppositions of our questions it should also be possible to predict the answers to questionnaire items like "Is the fragrance more similar to Chanel 5 or to Nuit d'Amour?" That is to say, we should be able to predict the location of the quality in the topological space of odors, provided we have a sufficiency of psychophysiological correlation laws to make this particular case one of interpolation or (limited) extrapolation.

The issue can however be made more poignant if we are concerned with the prediction of qualities within an entirely new modality. In the case of the congenitally blind who by a cataract operation suddenly attain eyesight, the experience of colors and (visual) shapes is a complete novelty. Suppose that all of mankind had been completely blind up to a certain point in history, and then acquired vision. Presupposing physical<sub>2</sub> determinism we should (according to my basic conjecture) in principle be able to predict the relevant neural and behavioral processes, and thus to foretell all the discriminatory and linguistic behavior which depends upon the new cortical processes (which correspond to the emergent, novel qualities of experience). What is it then that we would not or could not know at the time of the original prediction? I think the answer is obvious. We would not and could not know (then) the color experiences *by acquaintance*; i.e., (1) we would not *have* them; (2) we could not *imagine* them; (3) we could not *recognize* (or *label*) them as "red", "green", etc., even if by some miracle we suddenly *had* them, except by completely new stipulations of designation rules.\*

I conclude that the central puzzle of the mind-body problem is the logical nature of the correlation laws connecting raw feel qualities with neurophysiological processes. But before we tackle this difficult question, a glance at one more issue is required.

F. "Intentional" versus "Non-intentional". The mental life of (at least) the adult *homo sapiens* is characterized by the capacity for awareness—in addition to the occurrence of mere raw feels. (We credit some

\* Cf. Pap's discussion of absolute emergence (244).

animals and certainly young children with the latter in any case.) To have an experience, and to be aware of having it, is a distinction which I think cannot be avoided, even if in a given case it may be very difficult to decide whether awareness actually supervened. This is one of the notoriously difficult questions of phenomenological description. But assuming the distinction, it is fairly plausible that awareness is impossible without some sort of symbolism, even if it be the "silent" symbolism of imagery or (if there be such) of imageless thought. It is here where the idea of "intention" (not in the sense of purpose, end-in-view, or resolution, but) in the sense of *reference* becomes essential.

I shall try to show that the scientifically relevant issues regarding interactionism versus parallelism (or epiphenomenalism) should be carefully separated from the philosophical issues which stem from the "intentional" features of mind, stressed by Brentano and the phenomenological schools of thought. According to this point of view the most fundamental difference between the mental and the physical consists in the fact that the mental life consists of *acts directed upon objects*, no matter whether these objects exist in the world, or are pure concepts, or figments of the imaginations. It is true that dualism in the Cartesian tradition has emphasized the intentional as well as the raw feel features of mind. The mind-body problems in the larger sense therefore have customarily included such questions as, Can we give a physical (1 or 2, in this case) account of how thoughts, beliefs, desires, sentiments, etc. can be *about something*? Can we give a naturalistic translation of the language of reasoning as it occurs in arguments, i.e., discourse in which we give reasons intended to support knowledge claims, or value judgments? I think it has become increasingly clear\* that the answer must be in the negative; but not because human behavior involving "higher thought processes" is not in principle capable of physical (at least physical<sub>1</sub>) explanation and prediction; but rather because the problem is one of the logical reducibility or irreducibility of discourse involving *aboutness* (i.e., intentional terms), to the language of behavioral or neurophysiological description. Now it seems fairly obvious that such discourse, just like discourse involving *oughtness* (i.e., normative discourse) is not logically translatable into purely factual statements. The relation of designation (formalized in pure

\* Cf. especially Wilfrid Sellars (310, 311).

semantics) is not an empirical relation, but a construct of semantical discourse.

Personally, I therefore consider the problem of intentionality not as part of the psycho-physical but rather as a part of the psycho-logical problem, i.e., as part of the relation of psychological to the logical forms of discourse. This becomes even more evident because, assuming the ultimate possibility of a full neurophysiological account of behavior (including linguistic behavior), we should then have the problem of relating the physiological to the logical forms of discourse. If many writers permit themselves nowadays to speak of "thinking machines" (electronic computers, chess playing machines, etc.), then it is equally justified to pursue the problem of the relation between the mechanical (or the electrical) and the logical. In the case of the machines, it is ourselves who have built them in such a way that in their functioning they conform to certain rules of logical, mathematical, or semantical operations. In the case of human beings we have nervous systems which through education and training acquire the dispositions toward certain types of symbolic behavior which in actual operation then is more or less in conformity with certain rules.

But the abstract statement of a rule is not to be confused with the formulation of the (statistical) empirical regularity of the symbolic behavior. An illicit inference or a computation mistake is a violation of a rule, it is not an instance which would disconfirm a law of behavior. The recent phase of the clarification of these issues was in essence initiated by Husserl and Frege in their critique of *psychologism*, i.e., of the confusion of logical with psychological discourse. The pan-empiricist position of, e.g., John Stuart Mill who regarded logical truths as on a par with the truths of the natural sciences, was thus effectively and definitively refuted. Later, very much needed refinements of the anti-psychologistic position were added by Carnap (65, 68, 69, 71, 72), and a full study of the logical status of rules and rule-governed behavior has been contributed by W. Sellars (*loc. cit.*).

No matter what the most clarifying analysis of rule-governed symbolic behavior in its relation to the rules as such may turn out to be, there can be no doubt that if physical (at least physical<sub>1</sub>) determinism is to be maintained, the following will have to hold: A person's brain state when thinking, e.g., about Napoleon's defeat at Waterloo must qualitatively or structurally differ from the brain state of the same person (or,

for that matter, of other persons) when thinking about Caesar's crossing of the Rubicon. This aspect of the psychology and physiology of thought is definitely relevant for our problem.

#### V. Mind-Body Identity. Explications and Supporting Arguments

With due trepidation I shall now proceed to draw the conclusions from the preceding discussions, and to present the dénouement of the philosophical tangles. There are many points on which I have sincere and serious doubts. There is yet a great deal of analytic work to be done on several puzzling aspects for which I can at present only sketch the sort of solution which seems to me especially plausible.

A. *Review of the More Basic Meanings and Connotations of "Mental" and "Physical". Conclusions regarding their Respective Merits and Demerits.* The surveys and discussions of the preceding sections have paved the way for a summary and systematic appraisal of various characteristics which have been proposed as defining criteria of the mental and the physical. Outstanding candidates among the criteria of mind are (1) direct experience and (2) intelligence. "Direct experience" is synonymous with one sense of "subjectivity", viz. sentience, raw feels, or phenomenal givenness. "Intelligence" connotes learning capacity, purposive (goal directed) behavior and—on the human level—intentionality (symbolic behavior). Although the two criteria have in fact a certain area of coincidence, this coincidence (or overlap) is not a matter of logical necessity. By and large then, the two criteria of mentality define two entirely different concepts.

"Mind" as we have come to suspect all along, is an ambiguous term, or at best a group of concepts with family resemblances (in Wittgenstein's sense). The major components of the connotation of "intelligence" may be attributed not only to the higher animals but also to the "thinking machines" which we generally consider not only as lifeless but also as devoid of sentience. Direct experience, on the other hand, may well be attributed to some of the lower animals, babies, idiots, and to the severely insane; but in each of these classes at least some, if not all, of the marks of intelligence are lacking. Furthermore, it is customary in contemporary psychology to classify the unconscious (deeply repressed) traumata, anxieties, wishes, conflicts, etc. as *mental*. This again indicates that direct experience is not the criterion here, even if—according to the psychoanalytic doctrine—deeply repressed

matters are *potentially* conscious, in that they can be brought to the fore of awareness by special techniques. Hypnotic and posthypnotic phenomena also often involve deeply unconscious processes, which because of their other similarities with the conscious processes are unhesitatingly classified as mental.

One might suppose that the term "physical" (to which we have paid thus far only sporadic attention) is much more definite in meaning than the term "mental". Unfortunately, the contrary is the case. There are some superficial and entirely inadequate definitions of "physical" which need only be mentioned in order to be promptly dismissed. For example, to define "physical" as the "outer" aspect (in contradistinction to the "inner" mental life) is to use misleading metaphors. "Inside" and "outside", "internal" and "external" have a good clear meaning in ordinary usage. What is literally inside, e.g., the skin of a person is most of his body (i.e., the body minus the skin) and that's "physical" in at least one very good sense of the term. After all, anatomy and physiology are concerned with the physical structure and the functions of organisms. *Inside* the skull is the brain of man, and that is "physical" in the same well understood sense.

Similarly unhelpful is the definition of the "physical" as the *mechanical-compositional*, as contrasted with the *purposive-holistic*. We have already repudiated this sort of definition-by-contrast, by pointing out that "mechanical" in the strict sense of "characterizable by the concepts and laws of Newtonian mechanics" designates only a narrow subclass of the class of *physical* events or processes, using "physical" (comprising also electrodynamic, relativistic, and quantum-theoretical characteristics) in the sense of modern physics. And if by "purposive" we mean no more than by "teleological" and "holistic", then there are innumerable teleological mechanisms, many of them with typical features of organic wholeness, both in nature and among the artifacts of technology. If "purposive" is understood in the narrower and more fruitful sense, then it involves intelligence (and this, on the human level, includes intentionality).

But the fact that there are (human) organisms functioning intelligently and displaying (symbolic) behavior which indicates intentional acts is describable in an intersubjective ("physical<sub>1</sub>") manner and therefore again does not support a definition-by-contrast between the physical (in this case physical<sub>1</sub>) and the mental. It remains true, however,

that among the objects and processes describable in physical<sub>1</sub> terms, there are differences at least of degree (often of very considerable degree) if not of a fundamental, evolutionary-emergent type, as between the structure and the dynamics of electrons, atoms, molecules, genes, viruses, and unicellular and multicellular organisms. The tremendous differences between, e.g., a simple inorganic structure and a human being are therefore not in the least denied. As Castell (74) puts it, the solar system and an astronomer thinking about it, are in many essential respects very dissimilar indeed. (But the dualistic conclusions drawn by Castell seem to me nevertheless *non sequiturs*.)

The foregoing considerations suggest some of the more fruitful definitions of "physical". "Physical<sub>1</sub> may be defined as the sort of objects or processes which can be described (and possibly explained or predicted) in the concepts of a language with an intersubjective observation basis." This language or conceptual system is—in our sort of world—characterized by its spatio-temporal-causal structure. This is so fundamental a feature of our world that it is extremely difficult to imagine an alternative kind of world in which intersubjectivity is not connected with this feature. One can understand, but need not concede, Kant's contentions regarding the synthetic a priori character of this "presupposition." \* The concept of "physical<sub>1</sub>" is closely related to but by no means equivalent with one of the primary meanings of "physical" in ordinary language, viz. observable by sense perception. In its most natural usage "observable by sense perception", clearly comprises the solid and liquid objects of our environment; it includes of course our own bodies; it includes a trifle less clearly the air (which can be felt if it moves with sufficient speed; or other gases if they can be smelled); it includes less obviously some of the dispositional properties of various sorts of matter (such as their hardness, elasticity, solubility, fusibility, etc.); and it scarcely includes electric or magnetic fields, atoms and electrons, or the secret thoughts of other persons.

But in one usage "observable by sense perception" does comprise the feelings, emotions, and even some of the (dispositional) personality traits of other persons. For example, we say, "I could see how disappointed he was." "I can see that he is a depressive person," etc. But these are usages, which from the point of view of logical analysis are perhaps

\* For a critique of this rationalistic position, cf. Pap (242); Nagel (233); Reichenbach (275); Feigl (114).

not fundamentally different from the case of a physicist, who (looking at a cloud chamber photograph of condensation tracks) says, "Here I see the collision of an electron with a photon." Such (extended) "observation statements" urgently demand a logical analysis into their *directly* verifiable, as contrasted with interpretive and inferential components. Logical analysis, pursuing as it should, an epistemological reconstruction, must therefore be distinguished from phenomenological description.

From the point of view of a phenomenological description, the "pre-analytic data" of the clinical psychologist contain his direct impression of (some of) the personality traits of his clients; just as the experienced physician's judgments may be based on his direct impression of the disease (diabetes, multiple sclerosis, Parkinson's disease, etc.) of his patient. Phenomenological description is a subtle and interesting matter, but philosophically much less relevant than it is often supposed to be. By a little exercise of our analytic abilities we can, and for epistemological purposes we must, separate the directly verifiable situation (the patient is very slow in all his movements, hangs his head, speaks with a very low voice; or: he has dry skin; his breath has a fruity smell; his hands tremble; etc.) from the inferential interpretations, i.e., the conclusions regarding his mental or physical illness.

Inasmuch as the use of terms like "psychoneurosis" is established, and diagnoses of psychoneuroses can hence be confirmed, on an intersubjective basis, the concept of psychoneurosis is evidently a physical<sub>1</sub> concept. At least partial *explanations* of the behavior and the subjective experience of psychoneurotics have also been given on a physical<sub>1</sub> (roughly: behavioristic) basis. We can plausibly explain neurotic dispositions by tracing them causally to the childhood situations of the patient (not necessarily neglecting some of his biologically inherited constitutional traits). And we can predict his anxieties, depressed moods, etc. on the basis of such intersubjectively confirmable information as, e.g., about a preceding period of highly "id-indulgent," overbearing, or hostile behavior. These "physical<sub>1</sub>" explanations do not differ fundamentally from explanations of, e.g., the growth of plants or the behavior of lower animals. That a plant grows poorly may be explained by the sandy soil in which it is rooted, the lack of rainfall, etc. The behavior (or some aspects of it) of an amoeba may be explained by the thermal and chemical conditions of its immediate environment.

The distinction between psychoneuroses and "physical" nervous disorders originates from the same commonsense considerations that have traditionally led to the contrast of "states of mind" and states of the body. No matter whether normal or abnormal processes are concerned, whenever scientifically or philosophically innocent people speak of something as being "in the mind" or "merely in the mind," this means apparently that it is not directly accessible to sensory observation. But, it is also positively characterized by the fact that these "states of mind" can (usually) be reported by those who *have* them, and that they can (sometimes) be influenced by talking. Sticks and stones cannot be made to move by merely talking to them.\* Persons (having minds!) can be made to do things by suggestions, propaganda, requests, commands, etc., often by just giving them certain bits of information.

But important and interesting as is this sort of difference, in its scientific aspects it no longer establishes a fundamental difference between *inanimate things* and *minded persons*. Modern robots have been constructed which emit information about their "inner" (physical!) states, and they can be made to do things by speaking to them. But if intellectually acute and learned men † discuss seriously the problem as to whether robots really have a mental life (involving thoughts and/or feelings), there must be a question here that clearly transcends the obviously scientific and technological issue as to whether robots can be constructed which in their behavior duplicate all essential features (of course, one must ask: which ones and how completely?) of human behavior. If by "thinking" one means a kind of performance which, starting with "input" premises yield "output" conclusions of deductive or inductive inference, and consists (at least) in certain observable relations between *input* and *output*, then there is no doubt that certain types of robots or computers do think. If one means by "feeling" what the logical (or illogical?) behaviorists mean, then it is at least conceivable (cf. Scriven, 304) that there might be machinelike structures (artificially made, or even naturally existing on some other stars) which behave (respond, etc.) in every way as if they had feelings and emotions. ‡

\* This still seems safe to assert even in view of the alleged but highly questionable "facts" of psychokinesis.

† Cf. Turing (338); MacKay (216); Spilsbury (326); Scriven (304).

‡ The question in this form is by no means new. William James discussed it in his *Principles of Psychology* (Vol. I) by means of the example of the "automatic sweetheart." He was severely criticized by E. A. Singer (319) who, ironically enough,