Herbert Feigl, *The "Mental" and the "Physical": The Essay and a Postscript* (1967).

E. Arguments Concerning the Identification of Sentience with Neural Events.

I shall now present, as explicitly as I can, the reasons for an *empirical* identification of raw feels with neural processes. I shall also discuss several apparently trenchant arguments that have been advanced against this identity theory of the mental and the physical. It will be advisable first to state my thesis quite succinctly, and to elaborate the arguments for and against it afterwards.

Taking into consideration everything we have said so far about the scientific and the philosophical aspects of the mind-body problem, the following view suggests itself: The raw feels of direct experience as we "have" them, are empirically identifiable with the referents of certain specifiable concepts of molar behavior theory, and these in turn (this was argued in the preceding subsection D) are empirically identifiable with the referents of some neurophysiological concepts. As we have pointed out, the word,"mental" in present day psychology covers, however, not only the events and processes of direct experience (i.e., the raw feels), but also the unconscious events and processes, as well as the "intentional acts" of perception, introspective awareness, expectation, thought, belief, doubt, desire, volition, resolution, etc. I have argued above that since *intentionality* as such is to be analyzed on the one hand in terms of pure semantics (and thus falls under the category of the *logical*, rather than the psychological), it would be a category mistake of the most glaring sort to attempt a neurophysiological identification of this aspect of "mind." But since, on the other hand, intentional acts as occurrents in direct experience are introspectively or phenomenologically describable in something quite like raw-feel terms, a neural identification of *this* aspect of mind is prima facie not excluded on purely logical grounds. Unconscious processes, such as those described in psychoanalytic theory, are methodologically on a par with the concepts of molar behavior theories (as, e.g., instinct, habit strength, expectancy, drive, etc.) and hence offer in principle no greater difficulties for neurophysiological identification than the concepts of molar behavior theory which refer to conscious events or processes (e.g., directly experienced sensations, thoughts, feelings, emotions, etc.). As we have repeatedly pointed out, the crux of the mind-body problem consists in the interpretation of the relation between raw feels and the neural processes. The questions to be discussed are therefore these:

1. What does the identity thesis assert about the relation of raw feels to neural events?
2. What is the difference, if there is a difference, between psycho-physiological parallelism (or epiphenomenalism) and the identity thesis?
3. Can the identity thesis be defended against empirical arguments which support an interactionistic dualism?
4. Can the identity thesis be defended against philosophical arguments which support dualism on the grounds of the alleged fundamental differences between the properties of direct experience and the features of physical

(neurophysiological) processes?

Since I have already paved the way for at least partial replies to question 3, and to some extent also to 4, I shall now primarily concentrate on questions 1 and 2, and discuss the other issues more briefly whenever they will be relevant.

The identity thesis which I wish to clarify and to defend asserts that the states of direct experience which conscious human beings "live through," and those which we confidently ascribe to some of the higher animals, are identical with certain (presumably configuraiional) aspects of the neural processes in those organisms. To put the same idea in the terminology explained previously, we may say, what is *had-in-experience*, and (in the case of human beings) knowable by acquaintance, is identical with the object of knowledge by description provided first by molar behavior theory and this is in turn identical with what the science of neurophysiology describes (or, rather, will describe when sufficient progress has been achieved) as processes in the central nervous system, perhaps especially in the cerebral cortex. In its basic cqre this is the "double knowledge" theory held by many modern monistic critical realists.[1]

This view does not have the disadvantages of the Spinozistic doctrine of the unknown or unknowable third of which the mental and the physical are aspects. The "mental" states or events (in the sense of raw feels) are the referents (the denotata) of the phenomenal terms of the language of introspection, as well as of certain terms of the neuro-physiological language. For this reason I have in previous publications called my view a "double-language theory." But, as I have explained above, this way of phrasing it is possibly misleading in that it suggests a purely analytic (logical) translatability between the statements in the two languages. It may therefore be wiser to speak instead of *twofold access* or *double knowledge*. The identification, I have emphasized, is to be empirically justified, and hence there can be no logical equivalence between the concepts (or statements) in the two languages.

On superficial reflection one may be tempted to regard the identification of phenomenal data with neurophysiological events as a case of the theoretically ascertainable identities of the natural sciences. "Theoretical identity" (explicated in section V D) means the sameness of the referent (universal or particular) of two or more intersubjective descriptions. For example, it is the atomic micro-structure of a crystal which is indicated ("described") by the optical refraction index, the dielectric constant, the magnetic permeability coefficient, and in greater detail evidenced by X-ray diffraction patterns. Similarly, the various behavioral indications for habit strength refer to a certain, as yet not fully specified, neurophysiological structure in a brain, which may ultimately be certified by more direct histological evidence. Logical Behaviorism admits only intersubjectively confirmable statements and hence defines mentalistic (phenomenal) terms explicitly on the basis of molar behavioral theoretical concepts. Thus, to ascribe to a person the experience of, e.g., an after-image amounts, within the intersubjective frame of reference, to the ascription of a hypothetical construct (theoretical concept),

anchored in observable stimulus and response variables. This theoretical concept may then later be identified, i.e., come to be regarded as empirically co-referential with the more detailed and deductively more powerful neurophysiological concept.

The empirical character of the identification rests upon the exten-sional equivalences, or extensional implications, which hold between statements about the behavioral and the neurophysiological evidence. In our example this means that all persons to whom we ascribe an after-image, as evidenced by certain stimulus and response conditions, also have cerebral processes of a certain kind, and vice versa. In view of the uncertainties and inaccuracies of our experimental techniques we can at present, of course, assert only a statistical correlation between the two domains of evidence. That is to say, the equivalences or implications are, practically speaking, only probabilistic. But in any case, the correlations as well as the theoretical identification of the referents indicated by various items of evidence are formulated in *intersubjectively* confirmable statements.

The identification of raw feels with neural states, however, crosses what in metaphysical phraseology is sometimes called an "ontological barrier." It connects the "subjective" with the "intersubjective." It identifies the referents of subjective terms with the referents of certain objective terms. But in my view of the matter there is here no longer an unbridgeable gulf, and hence no occasion for metaphysical shudders. Taking into account the conclusions of the preceding analyses of "privacy", "acquaintance", "physical", and of "identification", private states known by direct acquaintance and referred to by phenomenal (subjective) terms can be described in a public (at least physicali) language and may thus be empirically identifiable with the referents of certain neurophysiological terms. Privacy is capable of public (intersubjective) description, and the objects of intersubjective science can be evidenced by data of private experience.

The application of phenomenal terms in statements of knowledge by acquaintance is direct, and therefore the verification of such statements (about the present moment of subjective experience) is likewise immediate. Phenomenal terms applied to other persons or organisms are used indirectly, and the confirmation of statements containing phenomenal terms (thus used) is mediated by rules of inference, utilizing various strands in the nomological net as rules of inference. Judging by the structure of one's own experience, there seems to be no reason to assume the existence of absolutely private mental states; i.e., there are presumably no "captive minds" in our world. 'Iliis is of course a basic ontological feature of nature as we have come to conceive it. It is an empirical feature of a very fundamental kind, similar in its "basic frame" character to the 3 + 1 dimensionality of space-time, or to the causal order of the universe. Such frame principles do not differ in kind, although they differ in degree of generality, from the postulates of scientific theories. Their adoption is essentially regulated by the rules of the hypothetico-deductive method.

Logical empiricism as it has come to be formulated in recent years (Carnap, 70, 73; Feigl, 116) recognizes the difference between direct observation (knowledge-by-acquaintance) statements and inferential statements as a contextual difference

between direct and indirect confirmation. It does not matter precisely where, in our epistemological reconstruction, we draw the line between the observable and the inferred entities. But wherever we do draw it, the scope of the directly experienceable or of the directly observable depends on the identity of the experiencing and/or observing subject.[2] What is directly verifiable for one subject is only indirectly confirmable for another. And these very statements (expressed in the preceding two sentences) may be formalized in a pragmatic, intersubjettive metalanguage.

Having formulated and in outline explicated the identity thesis, we now have to attend to several important points of philosophical interpretation. I reject the Spinozistic) double aspect theory because it involves the assumption of an unknown, if not unknowable, neutral ("third") substance or reality-in-itself of which the mental (sentience) and the physical (appearance, properties, structure, etc.) are complementary aspects. If the neutral third is conceived as unknown, then it can be excluded by the principle of parsimony which is an essential ingredient of the normal hypothetico-deductive method of theory construction. If it is defined as in principle unknowable, then it must be repudiated as factually meaningless on even the most liberally interpreted empiricist criterion of significance. But our view does not in the least suggest the need for a neutral third of any sort. This will now be shown more explicitly.

If a brain physiologist were equipped with the knowledge and devices that may be available a thousand years hence, and could investigate my brain processes and describe them in full detail, then he could formulate his findings in neurophysiological language, and might even be able to produce a complete microphysical account in terms of atomic and subatomic concepts. In our logical analysis of the meanings of the word "physical" we have argued that the physical sciences consist of knowledge-claims-by-description. That is to say that the objects (targets, referents) of such knowledge claims are "triangulated" on the basis of various areas of observational (sensory) evidence. What these objects are acquaintancewise is left completely open as long as we remain within the frame of physical concept formation and theory construction. But, since in point of empirical fact, I am directly acquainted with the qualia of my own immediate experience, I happen to know (by acquaintance) what the neurophysiologist refers to when he talks about certain configurational aspects of my cerebral processes.

There is danger at this point of lapsing into the fallacies of the well-known doctrine of structuralism, according to which physical knowledge concerns only the form or structure of the events of the universe, whereas acquaintance concerns the *contents* or *qualia* of existence.[3] This doctrine is to be repudiated on two counts. First, by failing to distinguish acquaintance (the mere having of data, or the capacity for imaging some of them) from knowledge by acquaintance (propositions, e.g., about similarities or dissimilarities, rank-orders, etc., of the qualia of the given), the doctrine fails to recognize that even introspective or phenomenological knowledge claims are structural in the very same sense in which all knowledge is structural, i.e., that it consists in the formulation of relations of one sort or another. Second,

the realistic interpretation of physical knowledge which we have defended implies that whatever we "triangulate" from various bases of sensory observation is to be considered as "qualitative" in a generalized sense of this term. In the vast majority of cases the qualitative content of the referents of physical descriptions is not "given," i.e., it is not part of a phenomenal field. But it is a given content in the case of certain specifiable neurophysiological processes.

If one wishes to trace the historical origins of this view, one might find it, if not in Aristotle, then certainly in Kant who came very close to saying that the experienced content is the *Ding-an-sich* which corresponds to the brain process as known in the spatio-temporal-causal concepts of natural science.[4] To put it more picturesquely, in the physical account of the universe as provided in the four-dimensional Minkowski diagram, there are sporadically some very small regions (representing the brains of living and awake organisms) which are "illuminated by the inner light" of direct experience or sentience. This view differs from panpsychism which assumes that the "internal illumination" pervades all of physical reality. But the panpsychists' hypothesis is inconsistent with the very principles of analogy which they claim to use as guides for their reasoning. If one really follows the analogies, then it stands to reason that the enormous differences in behavior (and neural processes) that exist between, e.g., human beings and insects, indicate equally great differences in their corresponding direct experience or sentience. Fancying the qualities of sentience of the lower animals is best left to poetic writers like Fechner, Bergson, or Maeterlinck. As regards the mental life of robots, or of Scriven's, (304) "androids," I cannot believe that they could display all (or even most) of the characteristics of human behavior unless they were made of the proteins that constitute the nervous systems -- and in that case they would present no puzzle.

The identity view here proposed has met with a great deal of resistance, especially on the part of modern analytic philosophers. To be sure, there are identifications which are "above suspicion." For example, it has been suggested that a legitimate form of empirical identification is to be found in such paradigms as the identity of the "visual" with the "tactual" penny (or the visual, tactual, and olfactory rose; or the visual, tactual, and auditory bell). In each of these examples one may distinguish the various domains of sensory evidence from the particular thing (or thing-kind) that the evidence indicates or refers to. Phenomenalists will, of course, be quick to point out that there is no sense in talking of a thing existing over and above the actual and possible "evidential" data and their important correlations. But from my realistic point of view it makes perfectly good sense to explain in terms of physical, psychophysical, and psychophysiological theories how, e.g. a bell by reflecting light, producing sound waves and being a solid, hard body affects our retina, cochlea, and our tactile nerve endings (under specifiable perceptual conditions) and thus produces the visual, tactual, and auditory data in our direct experience. This is indeed the "causal theory of perception" so much maligned by phenomenalists.

We grant that as empiricists we must ultimately justify the causal theory of perception (which is indeed a scientific theory, and not an epistemological analysis)

by reference to the evidential data which confirm it. And this we can do, no matter whether our own perceptions are concerned (in the egocentric perspective) or those of others (in the "side view" or lateral perspective that we obtain by observing the stimuli, central processes and responses pertaining to other persons). The various sensory "aspects" of the bell are thus to be conceived as the effects which the bell, considered either on the common sense level, or on the micro-level of scientific analysis, has upon our sense organs and finally on our awareness (this last effect empirically identifiable with processes in various cortical areas). Since the phenomenalist thesis of the translatability of physical object statements into data statements is untenable, epistemological analysis must "dovetail" with the causal (scientific) theory of perception and render justice to the latter by an explicit reconstruction of the *nomological* (not purely logical!) relations between the data and the illata. This is still conceptual analysis, in that it retraces the relations between the concepts of stimulus objects and the concepts pertaining to the central (cortical-mental) processes in the perceiving organisms.

Our ψ-φ identification, however, cannot be conceived according to the paradigm of the identity of stimulus objects (like the bell, or the rose). The analogy is misleading in that we have, in the case of stimulus objects physical descriptions of them which together with the empirical laws of psychophysics and psychophysiologv enable us (in principle) to derive their various sensory "appearances." Far from requiring an unknown or unknowable "third" or "neutral propertyless substance," ordinary knowledge and especially scientific theory contains a great deal of information about the nature and structure of stimulus objects. The situation in the ψ-φ case is fundamentally different: We don't have two kinds of evidence for one and the same entity (event, process, etc.). In direct acquaintance we have, we *experience* the datum (it is not evidenced, it is *evident*!), and we identify it with a physical process which we posit as an illatum whose existence is asserted on the basis of multifarious data in other evidential domains.

It should now be clear how the view here proposed differs from the Spinozistic double aspect doctrine. The data of experience are the reality which a very narrow class of neurophysiological concepts denotes. I admit this sounds very "metaphysical." And I shall no doubt be accused of illegitimately extending the ordinary meaning of "denotation". I am fully aware that I am extending the meaning. But I plead that this does not involve my view in paradoxes or needless perplexities. It is true that in common parlance, as well as in the widely accepted philosophical usage, we would say that a term like "neural process in the occipital lobe" denotes a pattern of nerve currents, and not a visual experience. But this remark obviously comes down to the true but trivial semantical assertion that a term designates its designatum; (e.g., "neuron" designates neuron!).

A specification of meaning can be attained through semantic designation rules only if the meaning of the translation equivalent of the definiendum is already understood in the metalanguage. Obviously, according to the commonly accepted usage of the word, a "denotatum" is the referent of proper names, and (except for the null cases) also of predicates, relations, etc. A genuine specification of meaning

for empirical terms can be achieved only by a combination of semantical, syntactical and pragmatic rules. The last two types of rules are particularly important. The syntactical rules specify the relations of concepts to one another, and the pragmatic ones make clear which concepts pertain to a basis of direct evidence. The realistic interpretation of empirical concepts depends on an appropriate analysis especially of the roles of proper names (and in scientific languages of coordinates) and of individual-variables (coordinate-variables).[5]

Taking these analyses into account, we can recognize the valid elements in the older critical realistic epistemology of perceptual and conceptual reference. A physical object or process as perceived in common life, or as conceived in science, is the referent of certain symbolic representations. I submit that it is the preoccupation with the confirmatory evidence which has misled positivists and some pragmatists (all of them phenomenalists, radical empiricists, or operationists) to identify the meaning of physical object statements with the actual and/or possible data which, according to our view, merely constitute their evidential bases. Worse still, even sophisticated analytic philosophers tend to confuse the meaning of physical concepts with the perceived or imaged appearance of physical things. No wonder then that we are told that the identity of certain neurophysiological states (or features thereof) with raw feels is a logical blunder. If the denotatum of "brain process (of a specified sort)" is thus confused with the appearance of the gray mass of the brain as one perceives it when looking into an opened skull, then it is indeed logically impossible to identify this appearance with the raw feels, e.g., of greenness or of anxiety.

It would be a similarly bad logical blunder to identify such raw feels with the scientific (heuristic or didactic) tinkertoy models of complex molecular structures (as of amino acids, or proteins) displayed by chemistry instructors in their courses. I don't know whether I should call these blunders "category mistakes." The first one simply consists in the confusion of evidence with the evidenced, or of the indicator with the indicated. What mistake does one make if one confuses smoke with fire, footprints with a man walking, certain darkish spots on an X-ray photograph with tuberculosis? It is strange that of all people it should be the analytic philosophers (who would expose these fallacies with ruthless irony) who do not see that they are making the same sort of mistake in thinking that physical-object concepts denote the perceptual appearance of physical things.

As I have been at pains to point out (in section IV), the only consistent and philosophically fruitful meaning of "physical" (more precisely, of "physical_1") is that of a conceptual system anchored in sensory observation and designed for increasingly comprehensive and coherent explanations of the intersubjectively confirmable facts of observation. This conceptual system or any part of it is in principle non-intuitive (*unanschauhlich* as the Germans call it, i.e., unvisualizable). Hence, an identification of a small subset of its referents with something directly given and knowable by acquaintance is in principle left completely open. In point of fact, the imagery commonly, and sometimes helpfully, employed in the thinking of theoretical physicists, biologists, or neuro-physiologists consists primarily of

pictorial appeals. These are at best intellectual crutches, fruitful only heuristically or didactically, and not to be confused with conceptual meanings. The fallacy of "introjection"[6] which was so vigorously criticized by Avenarius (the empiriocriticist of the last century) consists in the pictorial ascription of raw feels to other organisms. As we have seen, such ascriptions indeed clash with the (equally pictorial) ascriptions of physical-appearance properties to other persons or animals.

In the perceptual awareness of other organisms we are confronted with their behavior, i.e., their responses, facial expressions, tone of voice, gait, posture, linguistic utterances, etc., but never with their raw feels. Raw feels do not and cannot be fitted into the appearance picture. They must therefore be conceived as the subjective counterpart of these appearances. As such they are inferentially attainable but not perceptually accessible. At an earlier point we have already discussed the phenomenology of the alleged intuitive or empathetic apprehension of the mental states of other organisms. Since we must recognize intuitive or empathetic ascriptions as fallible and corrigible, they have to be regarded as inferential from the point of view of logical reconstruction (i.e., in the context of justification), no matter how immediate, "self-evident," compelling, or convincing they may be psychologically.

That "introjection" in this sense leads to absurdities becomes especially clear when we consider the ascription of phenomenal fields, e.g., of visual spatiality to other persons. Unless we are solipsists, there is every good reason in the world to ascribe to others the same sort of "life space" (phenomenal environment) which we find so distinctly within our own experience. But if we think of other persons in terms of their appearance in our own phenomenal environment, then it is impossible to ascribe (pictorially) to them also the particular perspectives that they perceive of their environment (or of parts of their own bodies). The fallacy is just as gross as in the case of expecting to find in the brain of another person looking at a green tree a little picture of that tree. But pictorial thinking is one thing, and conceptual thinking is quite another. For conceptual ascription, however, there is no difficulty. The concepts of neurophysiology are non-intuitive and must not be confused with their logically irrelevant pictorial connotations. These connotations lend, psychologically speaking, a certain "root flavor" to these concepts. But once the pictorial appeals connected with the evidential roots of our physical or neurophysiological concepts are dismissed as irrelevant, they no longer pre-empt those places in the conceptual system of which we may then say that they denote some raw feels.

For these reasons I think that once the proper safeguards are applied, no category mistakes are made if we combine phenomenal and physical terms, as indeed we do quite ingenuously not only in ordinary discourse but also in the language of psychology. There is no reason why we should not say, e.g., "The anticipation of success quickened his pace"; "Morbid and tormenting thoughts caused his loss of appetite"; "Touching the hot stove caused intense pain"; "His repressed hostilities finally produced a gastric ulcer"; etc. Category mistakes do arise from confusions of universals with particulars; or of dispositions with occurrents. The first sort of

category mistake certainly consists in a violation of the Russellian rule of types. I am not sure whether the second sort can always be reduced to the first. But the original diagnosis made especially by Carnap in his early (phenomenalistic) work (60) of the mind-body perplexities as Russellian-type confusions is no longer acceptable. Physical concepts are not logical constructions out of phenomenal concepts.

A more serious objection to identification comes from reflections upon Leibniz's principle of the identity of indiscernibles. Since we have not only admitted, but repeatedly emphasized the empirical nature of the ψ-φ identification, one may well ask how we can speak of identify if its confirmation requires the observation of empirical regularities. The most direct confirmation conceivable would have to be executed with the help of an autocerebroscope. We may fancy a "compleat autocerebroscopist" who while introspectively attending to, e.g., his increasing feelings of anger (or love, hatred, embarrassment, exultation, or to the experience of a tune-as-heard, etc.) would simultaneously be observing a vastly magnified visual "picture" of his own cerebral nerve currents on a projection screen. (This piece of science fiction is conceived in analogy to the fluoroscope with the help of which a person may watch, e.g., his own heart action.) Along the lines of the proposed realistic interpretation he would take the shifting patterns visible on the screen as evidence for his own brain processes. Assuming the empirical core of parallelism or isomorphism, he would find that a "crescendo" in his anger -- or in the melody he heard -- would correspond to a "crescendo" in the "correlated" cortical processes. (Similarly for "accelerandos," "ritardandos," etc. Adrian's and McCulloch's experiments seem to have demonstrated a surprisingly simple isomorphism of the shapes of geometrical figures in the visual field with the patterns of raised electric potentials in the occipital lobe of the cortex.) According to the identity thesis the directly experienced qualia and configurations are the realities-in-themselves that are denoted by the neurophysiological descriptions. This identification of the denotata is therefore empirical, and the most direct evidence conceivably attainable would be that of the autocerebroscopically observable regularities.

Any detailed account of the ψ-φ identities is a matter for the future progress of psychophysiological research. But in the light of the scanty knowledge available even today, it is plausible that only certain types of cerebral processes in some of their (probably configurational) aspects are identical with the experienced and acquaintancewise knowable raw feels. A "psychological physiology"[7] which frames hypotheses about neural structures and processes on the basis of a knowledge of the characteristics and the regularities in the changes of phenomenal fields must therefore always remain extremely sketchy. Knowledge by acquaintance of phenomenal fields alone cannot possibly yield more than a few strands of the total nomological net of neurophysiological concepts required for the explanation of molar behavior. The identification is therefore restricted to those elements, properties, or relations in the neural processes which (in dualistic parlance) are the "correlates" of the raw feels. In our monistic account this is tantamount to the identity of the denotata directly labeled by phenomenal terms, with the denotata of neural descriptions. These latter denotata are acquaintancewise unknown to the

ncurophysiologist, except if he uses the autocerebroscope himself.

Now it is clear that neural correlates (to speak for the sake of easier exposition once more dualistically) are denoted by concepts which are much richer in meaning than the corresponding phenomenal concepts. The neurophysiological concepts refer to complicated, highly ramified patterns of neuron discharges, whereas their raw-feel correlates may be simple qualities or relations in a phenomenal field. How can, e.g., a uniform patch of greenness, a single musical tone, a stinging pain be identical with a complex set of neural events? Here again it is essential to distinguish between the *scientific* and the *philosophical* components of this question. Our psychophysiological ignorance is still too great to permit anything more than bold guesses on the scientific side.

Tliere has been talk of "thresholds" and "fusion"; i.e., it is assumed that raw feels emerge only if the intensities of the neural patterns have reached a certain degree; and that complex neural patterns may be "fused" so that the emerging quality "appears" simple and uniform. This sort of talk, though dangerously apt to mislead, is not entirely illegitimate. (Talk of thresholds, limens, and fusion is of course quite customary and proper in psychophysics, but its extension to psychophysiology is precarious. It makes perfectly good sense, and is true, to say that the white and black sectors on a swiftly rotating disk phenomenally fuse and yield a uniformly gray appearance. It makes perfectly good sense also, and is equally true, that the intensity of physical stimuli (like light, sound, pressure on one's skin, concentration of chemical substances in the air, etc.) must surpass a certain lower limiting value, if they are to effect a sensation in any of the various modalities (sight, hearing, touch, smell, etc.).

If these facts have any analogies in the intra-cerebral sphere, it would have to be assumed that one area of the cortex "taps" or "scans" other areas and could thus not come to react unless the input reaches a certain intensity. Likewise, one would have to assume that the effect in the second area reflects only certain gross features of the intricate and multifarious process patterns in the first. These would be the analogues of psychophysical thresholds and fusions. Finally, one may assume that the second area (which corresponds to the sensing of the raw feels) is connected with another area corresponding respectively to awareness or judgment (as in introspection) and finally to a motoric area of the cortex which innervates expressive responses or speech.[8] May I say again that I don't for a moment insist on the scientific adequacy of this particular model. I am not trying to do armchair neurophysiology. All I am concerned to point out is that models are conceivable which would enable us to remove the obstacles arising from the apparent disparities of phenomenal unity versus physical multiplicity; phenomenal spatialities and physical space; phenomenal time and physical time; phenomenal purposiveness and physical causality; etc. I am now going to outline these considerations very briefly.

W. Köhler (182, 183, 185) and R. Ruyer (290, 292, 293) have convincingly shown that the notorious Cartesian perplexities regarding spatiality can be removed by closer attention to the facts of psycho-physiology combined with a logical

clarification of the distinction between phenomenal space(s) and physical space. (We have laid the groundwork for this in section III B). The surface of objects "physically" outside my skin naturally appears in my visual space as external to the visual appearance of those parts of my body which I can see. There is histological and physiological evidence for a relatively simple projection of the excitation patterns in the retina of the eye, in the area of the occipital lobe of the cerebral cortex. The projection, in its physical and geometrical aspects, is similar to the sort of projection one gets on the screen of a periscope inside a submarine. Not only parts of the surrounding surface of the sea and of other ships, but also parts of the (surfacing) submarine itself are projected upon the screen. Similarly, when I lie on a couch I find not only the appearances of tables, chairs, walls, and windows within my visual field, but I find these object appearances phenomenally outside that part of my phenomenal body (chest, arms, hands, legs, feet) which is also included in my visual field. These simple reflections show that some of the older philosophical puzzles about the outward projection of visual percepts from my mind or brain into the external world are gratuitous, based on confusions, and resolvable by proper attention to the scientific facts on the one hand and to the meanings of spatial terms and phraseologies on the other.

The resolution of the perplexities regarding phenomenal versus physical time, as well as experienced purposiveness versus physical or physiological causality proceeds quite analogously. In the phenomenally temporal "projection" we locate ends-in-view at some distance in the future, and then go about attaining these ends by action, i.e. by the utilization of means. If, e.g., I decide to attend a lecture, I may have to go through a long chain of acts, such as walking to my garage, starting my car, driving to the auditorium, and getting seated there. My actions are clearly goal directed, but there is no need for the myth about the later events (the goal) influencing my antecedent behavior. My behavior is guided, controlled, or modulated by the goal idea which is contemporaneous with my instrumental acts, or possibly precedes them. What in the phenomenal description appears like a future event in my life career determining my current behavior, becomes in the causal account the effect of one part of my cerebral processes upon another. Of course in this case, just as in the case of memory (recollection), our thinking is essentially mediated by symbols; and therefore "intentionality" (cf. section IV F) plays an important role here. But the symbolic representation of past events or of future events is effected by processes occurring now; i.e., these representations are causal factors in the determination of current behavior. Just as there is no need for a curious notion of "final" causes (or, in Lecomte de Noüy's phrase, of "telefinality"), there is no need for the assumption of a literal presence of the past in present recollections. Whatever the adequate and detailed neurophysiological account of memory traces may ultimately turn out to be, it is these memory traces and not some direct and mysterious apprehension of past events which will causally account for the facts of recollection and of the modification of behavior through learning processes.

Similar considerations would seem to apply to the perennial puzzles concerned with the problems of the nature of the "self," i.e., the unity of the ego, or the unity of

consciousness. Here, as in the other puzzles just discussed, the phenomenological descriptions may be correlated with the neurophysiological explanations. Phenomenally there may or may not be a "central core," the "I," in all my experiences. We may admit, following Hume and the later empiricists in the Humean tradition, that there is no distinct element, datum, or impression that could properly be regarded as the self. But it is hard to deny that in the directly given data and in their succession throughout experienced time, there is a certain feature of centralization, coordination, organization, or integration -- the reader may choose whichever term seems most suitable, is unitary organization seems to rest on the ever-present potentialities of recollecting a great many events or sequences of events of one's (sic!) past; the ever present possibility of the occurrence of somatic data (referring to one's own body); the existence of a set of dispositions or behavior tendencies, including those ascribed (psychoanalytically) to the superego (i.e., in plain language our set of values and ideals as incorporated in one's conscience); and finally that conception of one's self which is largely a result of the realization of one's own character and personality, adequately or often very inadequately derived from interpretations of one's own behavior and one's social role as perceived by oneself or by others in the social context.

Whichever of these aspects are in some sense phenomenally "given" -- and I suggest a good many may well be so given -- these aspects very likely "correspond" to (or according to my view, are identical with) certain relatively stable patterns of cerebral structures and functions. In the pathological cases of split or of alternating personalities (of the Sally Beauchamp, or of the Dr. Jekyll and Mr. Hyde varieties), it has often been suggested that we deal with cerebral subsystems, each having "organic unity" in itself, but only one of them dominating in the determination of behavior during certain intervals of time. If according to psychoanalytic theory large parts of the id as well as of the superego are unconscious, this may well be interpreted by assuming that certain portions of the cerebral processes are blocked off (this corresponds to "repressed") from the areas of awareness and of verbal report.

Having rendered plausible the scientific feasibility of at least a parallelistic account of some of the striking and remarkable features of mental life, I return now to the philosophical or logical crux of the identity thesis. We have stressed that the (empirical!) identification of the mental with the physical consists in regarding what is labeled in knowledge by acquaintance as a quale of direct experience as identical with the denotatum of some neurophysiological concept. The scientific evidence for parallelism or isomorphism is then interpreted as the empirical basis for the identification. The step from parallelism to the identity view is essentially a matter of philosophical interpretation. The principle of parsimony as it is employed in the sciences contributes only one reason in favor of monism. If isomorphism is admitted, the dualistic (parallelistic) position may be retained, but no good grounds can be adduced for such a duplication of realities, or even of "aspects" of reality. The principle of parsimony or of inductive (or hypothetico-deductive) simplicity does oppose the operationistic predilection for speaking of two (or more) concepts if the evidential facts, though completely correlated, are qualitatively

heterogeneous.

Our view of "triangulation" under such conditions of convergence has, I trust, shown the operationist view to be by far too restrictive. But there is still the logical question how concepts with such fundamentally different evidential bases can be interpreted as (empirically) identifiable. In the case of the concept of the electric current (cf. above section V C) as measured by its magnetic, chemical or thermal, etc. effects, the identification of the several operationally introduced concepts is plausible enough. But, it will again be asked, how can we speak of identity in the entirely different psychophysiological case where one of the concepts is characterized by the direct applicability of subjective acquaintance terms and the other (the physiological) is introduced on an intersubjective basis and thus has its evidential roots in the sensory data of any qualified observer? I think the answer is not so difficult any more. If we first consider "acquaintance" in its ordinary usage, we can certainly say that Anthony Eden is acquainted with Queen Elizabeth II, and I am not (never having had the opportunity of meeting her). Nevertheless, I can lay claim to some knowledge about the Queen, based on newspaper reports, pictures, and the like. It is surely the same person that Eden and I know, each in his way. Closer to the point, I know by acquaintance what it is to have an eidetic musical-image experience (I occasionally "hear with my inner ear" entire passages from symphonies, string quartets, etc. in their full tone colors). Someone else lacking this sort of experience does not know it by acquaintance, but he can know about it, especially if he is a skillful experimental psychologist. It would be unparsimonious to assume that the psychologist and I are referring to two different (but correlated) processes.

Now, direct acquaintance with "private" raw feels is describable also in the intersubjective language of science. Its ultimate explanation may again have to refer to various cerebral areas, one of which (speaking for ease of exposition again dualistically) "corresponds" to sensing, another to judging, and possibly another yet corresponds to (introspective) reporting. I conclude that acquaintance statements differ only in the type and domain of evidence, but not in regard to their reference, from certain neurophysiological statements^ Since the neural apparatus of introspection differs most markedly from that of (external) perception, it should not be surprising that knowledge by acquaintance (now taken in its narrow epistemological sense) is so much more crude, undetailed, and imprecise, than knowledge based on sense perception, especially when this is aided by the instruments of science.

Direct awareness, as we have pointed out before, usually furnishes only qualitative or topological orderings of the contents of phenomenal fields. It could not by itself inform us about the cerebral localization of subjective experience. A very crude (but, if taken literally, I fear highly misleading) analogy might help illuminate this point. A man lost in a jungle perceives the trees and undergrowth in his immediate environment. But the location of this very same part of the jungle can be determined in a much more accurate and encompassing manner by a cartographer making his measurements from the vantage point of an airplane or balloon high

above the jungle. This simile is misleading, of course, in that both the lost wanderer and the cartographer use sensory perception as evidential bases for their knowledge claims. This clearly differs from the case in which I report (or "avow" as Ryle puts it), e.g., a feeling of anxiety and a behavioral psychologist infers my anxiety from the "symptoms," or a neurophysiologist recognizes it in the "corresponding" cerebral processes. Nevertheless, I fail to see that the difference, important though it is in many ways, affects the argument for the identification of the referents of the introspective avowal, with those of the two scientific descriptions.

I conclude that ψ-ϕ identity as I conceive it is then still an identity of indiscernibles as defined by Leibniz and Russell. But as the clarification of the "paradox of analysis" (cf. Feyerabend, 120) and of related puzzles about belief sentences should by now have made amply clear, mutual substitutivity even of logically synonymous expressions holds only in non-pragmatic contexts. The empirical synonymy of ψ and ϕ terms (or, more cautiously perhaps, their empirical co-reference) a fortiori does not allow for substitutivity in pragmatic contexts. By this I mean that the "salva veritate" condition is fulfilled only in contexts of substitution which do not depend on what we know, or what evidence we have for our knowledge claims. As we pointed out before, there are or were many people (primitive, ancient, etc.) who have no idea of the association of mental life with cerebral processes. But it is nevertheless as justifiable to speak of identity here as it is in the case of "Walter Scott = the author of the Waverley novels," regardless of whether this fact is known or unknown to a given person. In this particular and well-worn example the identity concerns an individual. But, not being a nominalist, I see no difficulties in the identity of a universal, named or described in various ways. Psychophysiological identity may be identity of particulars (this twinge of pain with a specific cerebral event at a certain time), or of universals (pain of a certain kind, and a type of cerebral process).

I am finally going to tackle more specifically and pointedly the question: What is the difference that makes a difference between the parallelism and the identity doctrines? The pragmatist-positivist flavor of this question suggests that it concerns empirically testable differences. But I have already admitted that there are no such differences and that there could not be any, as far as conceivable empirical evidence is concerned. Is the identity thesis then a piece of otiose metaphysics? Whether it is metaphysics depends of course on what one means by "metaphysics". As I see it, the question is not only similar, but indeed intimately related, to such "metaphysical" issues as realism versus phenomenalism, or the modality versus the regularity view of causality. As most philosophers nowadays realize, these issues unlike disputes regarding scientific theories cannot be decided by empirical tests. These questions concern the explication of the meaning of concepts and assumptions. They are a subject matter for logical analysis.

As to whether there is a tenable meaning of "causal necessity" related to regularity, but not reducible to it, this is a controversial issue today. My own reflections favor a view of causal modalities (possibility, necessity, impossibility) which explicates the use of these terms meta-linguistically, and nevertheless does not conflict with

Hume's basic, and in my opinion irrefutable, contention; viz., that (if I may put it in my own way) the only evidence we can ever have for the assertion of causal connections must be observed regularities. There is, as I see it, no test for causal necessity over and above the tests for regularity. But this does not preclude meaning from the distinction between accidental and necessary universal synthetic statements. A world is conceivable in which a certain metal with a high melting point (say, e.g., platinum) everywhere and always in the infinite history of that world occurs in the solid state, simply because the temperature in that world "happens" never anywhere to surpass a certain upper limit. In such a world the universal statement "$(x,y,z;t)(Pt_{xyzt} \supset S_{xyzt})$," i.e., "platinum is everywhere and always solid" would be a true universal statement. But the counterfac-tual conditional "if the temperature were ever to reach or surpass a certain value, platinum would melt" might even be deducible from the basic laws of physics of that world. The universal statement in question is accidentally true. It is not a consequence of a basic law of nature; its truth depends on certain contingent features of the initial and boundary conditions of the fancied world. This shows that there are meaningful distinctions for which no conceivable empirical test could be designed.

Even closer to our problem is the issue between realism and phenomenalism. As I have shown elsewhere (110), there is again no testable difference between these two interpretations of factual knowledge, but there are excellent reasons for the repudiation of phenomenalism and hence for the acceptance of a realistic epistemology. To relegate the issue to the limbo of metaphysics is a lazy man's way of saving himself the troubles of careful analysis. But close attention to the logic of evidence and reference shows that phenomenalism, even in its most liberal forms does not and cannot substantiate its translatability doctrine; and that only a view which relates phenomenal evidence synthetically to statements about physical objects is ultimately tenable.

It is precisely because realists locate both the evidence and the evidenced within the nomological net, that they can give a more adequate account of the relation between "the knower and the known" than positivists, pragmatists, or operationists have ever been able to provide. And it is for this very same reason, that our view of the nature of physical concepts enables us to identify some (of course very few only!) of their referents with the referents of raw feel terms. Dazzled by the admittedly tremendous importance of the evidential basis for our knowledge claims, positivists have regrettably neglected the very objects of those knowledge claims. They have myopically flattened them into the surface of evidence, and thus prevented themselves from giving a viable account of the concepts of physics; and they have merely evaded or repressed the mind-body problem which they thought would vanish if their "reductions" -- phenomenalistic or behavioristic -- were accepted. Ingenious and tempting though their more sophisticated endeavors of reduction have been, they did not succeed. This is why I felt that an explicit reinstatement and defense of a realistic solution of the mind-body problem would be timely and worthwhile.

Notes

1. Especially Alois Riehl, Moritz Schlick, Richard Gatschenberger, H. Reichenbach, Günther Jacoby, Bertrand Russell, Roy W. Sellars, Durant Drake, and C. A. Strong. To be sure, there are very significant differences among these thinkers. Russell has never quite freed himself from the neutral monism (phenomenalism) of his earlier neorealistic phase. R. W. Sellars and, following him on a higher level of logical sophistication, his son, Wilfrid, have combined their realistic, double-knowledge view with a doctrine of evolutionary emergence. Opposing the emergence view, Strong and Drake, originally influenced by F. Paulsen, adopted a panpsychistic metaphysics. My own view is a development in more modem terms of the epistemological outlook common to Riehl, Schlick. Russell, and to some extent of that of the erratic but brilliant Gätschenberger. The French philosopher Raymond Ruyer (289, 290) especially before he turned to a speculative and questionable neovitalism (293) held a similar view. Among psychologists W. Köhler (182, 183), E. G. Boring (40), and D. K. Adams (1), again differing in many important respects, hold similar monistic positions. Personally, I consider sections 22-35 in Schlick (298) as the first genuinely perspicacious, lucid and convincing formulation of the realistic-monistic point of view here defended. It is to be hoped that an English translation of this classic in modem epistemology will eventually become available.

2. As I understand Dewey and other pragmatists, as well as contextualists like S. C. Pepper (254, 255), this point has been explicitly recognized by them. Cf. also the discussions by analytic philosophers, such as Hampshire (141), Watling (341), and Ayer (18). An exact logical account of the linguistic reflection of direct versus indirect verifiability has been given in the analysis of egocentric particulars (token-reflexive, indexical terms) by B. Russell (286), Reichenbach (274), Burks (58), W. Sellars (308, 312),and Bar-Hillel (20).

3. This doctrine has been espoused in various forms by Poincare1 (257), Eddington (93), C. I. Lewis (195), Schlick (299), et al.

4. Cf. I. Kant, *Critique of Pure Reason*, section on "The Paralogisms of Pure Reason."

5. Cf. especially W. Sellars (308); H. Feigl (110, 111); Bar-Hillel (20).

6. The term "introjection" as used by R. Avenarius has nothing to do with the well-known homonymous psychoanalytic concept.

7. Advocated by W. Köhler (184, 185) and critically discussed by C. C. Pratt (260).

8. I am indebted to R. Carnap for suggesting (in conversations) this sort of brain model.