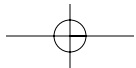
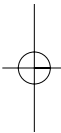
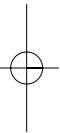


PART I

THE PLACE OF
PHENOMENOLOGY IN
PHILOSOPHY OF MIND



1

Functionalism and Logical Analysis

Paul Livingston

Abstract: Though it is most often deployed in service of naturalist and empirically sensitive explanatory projects, the functionalist theory of mind is essentially a formal theory, drawing its plausibility more from a sophisticated appreciation of the logic and conceptual grammar of terms of psychological description than from any empirical consideration. In this, the functionalist theory of mind exhibits significant methodological continuities with the tradition of phenomenology; but despite its successes, many philosophers believe that functionalism fails in that it leaves out any account of the central explanatory concept of phenomenology, the concept of immediate, subjective experience. In this essay, I analyze the history of the development of functionalism to make perspicuous some of the hidden structural features of the doctrine we know today. Functionalism emerges as a sophisticated response to problems of the meaning and reference of psychological terms left open by its predecessor theories. This shows that the question of the relationship of formally described functional states to empirically described physical states remains open and suggests a new way of viewing the source of functionalism's continued problems with explaining consciousness.

After more than thirty-five years of debate and discussion, versions of the functionalist theory of mind originating in the work of Hilary Putnam, Jerry Fodor, and David Lewis still remain the most popular positions among philosophers of mind on the nature of mental states and processes. Functionalism has enjoyed such popularity owing, at least in part, to its claim to offer a plausible and compelling description of the nature of the mental that is also consistent with an underlying physicalist or materialist ontology. Yet despite its continued popularity, many philosophers now think that functionalism leaves something out, in particular that functional explanations and analyses fail to account for consciousness, qualia, or phenomenal states of experience or awareness.¹ If the objection is correct, then functionalism fails in its inability to capture the central explanatory basis of phenomenological explanation: the phenomena of immediate, first-personal, subjective experience. The apparent failure is all the more striking in view of the close methodological parallels that exist between functionalism and phenomenology; for both projects depend centrally on a program of conceptual investigation of the definitional and explanatory interrelationships of our descriptions of experience and other psychological phenomena. A historical overview

¹ See, e.g., Nagel (1974), Chalmers (1996), and Searle (1992).

of the theoretical pressures that led to the development of functionalism shows that it was in fact problems about the logical form of immediate, subjective experience that led most centrally to its development, and that these problems continue to threaten the coherence of the functionalist theory. This, in turn, suggests that the contemporary problem of explaining consciousness is not a metaphysical or empirical one about the explanation of a particularly puzzling process, but a conceptual one about the logical structure of experience.

The resistance of consciousness to functionalist explanation can initially seem difficult to account for: why should *one* set of mental phenomena—those characterized by phenomenal or qualitative content—so stubbornly resist explanation in the functionalist terms that seem successful elsewhere? But historical investigation shows that the functionalist theory itself emerges from the philosophical pressure put on earlier theories—especially the identity theory of Place, Smart, and Feigl—by more general and recurrent problems with the explanation of the phenomena of subjective experience. For investigative projects in analytic philosophy, beginning with the ‘meaning analysis’ projects of the logical positivists, meaning has been intelligible as a matter of linguistic or logical structure, and the analysis of language as a matter of the description of this structure.² Over against the logical or grammatical structure of language, however, experience has consistently been characterized as immediate, nonrelational content, inaccessible to the structural explanations that the analysis of meaning can offer. Historically based attention to the consistency of this problematic sheds light on the underlying motivations and theoretical contours of various particular versions of the analytic project, and recommends a more explicit and methodologically sensitive discussion of the central question of the relationship of experience to meaning. In the light of historical investigation, the functionalist theory itself, I shall argue, emerges as a particularly sophisticated kind of conceptual analysis, continuous with phenomenology in its aim to give a broadly structural characterization of the logic or grammar of our concepts of psychological description and prediction. And both the underlying motivations for its development and the largest set of problems it continues to face emerge as consequences of the underlying and ongoing resistance of immediate, subjective experience or consciousness to such projects.

1

Functionalism first arose within the analytic tradition as a response to the then-prevalent *identity theory* of mind, the theory that held that mental states, including sensations, mental images, and other phenomenal states, are (as a matter of empirical fact) identical with physical states of the brain. The identity theory of U. T. Place, J. J. C. Smart, and Herbert Feigl improved upon previous behavioristic analyses of mind in that it could construe reports of experience as genuinely referring to genuine inner items, albeit physical ones, rather than (as the logical behaviorist must) simply

² See my ‘Structuralism and Content in the Protocol Sentence Debate’ for the origins of this way of thinking about meaning and experience.

taking them to replace more primitive bits of behavior.³ But though the identity theory solved this decisive problem with the behaviorist's construal of the reference of reports of experience, it omitted any account of the relationship of the ordinary logic of psychological description and explanation—the logical structure of psychological terms, their roles in the prediction and explanation of behavior, and the criteria on the basis of which they are normally ascribed and deployed—to the logical structure of the neurophysiological description of physical states of the brain. Such an account would be needed in order for the identity theory to stand any chance of empirically earning the particular psychophysical identities that would justify its general claim, but such an account would also require all of the resources of a thoroughgoing conceptual investigation of the logical structure of psychology. Justifiably, philosophers wanted to know not only *that* mental states could be physical states, but also *how* they could be and *what* it would tell us if they were.

It was as a response to this question about the meaning of psychological and physical terms that the functionalist theory of mind began to emerge. In 1957, Hilary Putnam began to articulate a new way of looking at the relationship between psychological and physical terms, as this relationship might develop diachronically under the influence of the growth of empirical discovery. This provided the possibility, Putnam thought, for a new kind of defense of the identity theory. Key to the proposal was a distinction between meaning and reference: ordinary-language terms like 'blue sensation' could actually *refer* to brain states, even if there is no sense in which they *mean* the same as any neurophysiological description. When empirical progress reveals the underlying causes for a phenomenon, Putnam reasoned, it will often make sense to construe our language for discussing the phenomenon as referring to the underlying cause rather than the surface phenomenon. Thus, a term like 'polio' will initially be used to describe a characteristic set of symptoms; but when the underlying viral cause is recognized, it will make sense to hold that 'polio' is the underlying virus, and not simply the symptoms.⁴ The relationship between mental states and brain states, Putnam reasoned, might share this structure. Future empirical discovery, if it made good on the identity theory's suggestion, would answer the semantic question about the identity theory by identifying the *actual* referents of our ordinary psychological discourse.

In the 1960 article 'Minds and Machines', Putnam further develops this account of the diachronic change in reference of psychological terms, tying it to a more explicit consideration of the success conditions for the identification of terms within a theory. Essential to Putnam's argument is the observation that, as scientific theories develop, terms are often used in new ways, not because they change their meanings, but because they take on new uses in the new contexts revealed by new pieces of theory. If theoretical identifications represent empirically justified extensions in the uses of terms without implying any great change in the underlying *meaning* of those terms, then the identity theorist can both admit that mental state/physical state identifications are today semantically abnormal and describe conditions under which the very same identifications could become normal and indeed necessary. The theoretical

³ Smart (1959: 144).

⁴ Putnam (1957: 100).

identification of mental states with physiological states, Putnam suggests, will begin to make sense when we understand not only how the two kinds of states are correlated, but also how physiological states themselves *cause* behavior. Were physical science capable of describing the causation of behavior by physiological states, the identification of physical states with mental states would subsequently have two theoretical advantages:

1. It would be possible . . . to derive from physical theory the classical laws (or low-level generalizations) of common-sense 'mentalist' psychology, such as: 'People tend to avoid things with which they have had painful experiences'.
2. It would be possible to predict the cases (and they are legion) in which common-sense 'mentalist' psychology fails.⁵

In contemplating the possibility of reducing psychological theory to physical law, Putnam had also begun to think about the logical structure of common-sense psychological description itself, as well as its relationship to the traditional philosophical problems of mind–body identity. This led him to the most historically significant suggestion of the paper: that a sufficiently complex computational machine with certain abilities of self-description and theory-building—for instance a machine of the abstract kind suggested by Turing a few years earlier—could serve as a rough analogue for a human's psychological organization; and that in so doing, it would develop strict analogues for *all* of the traditional philosophical problems about the relationship of mind–body identity:

In particular, if the machine has electronic 'sense organs' which enable it to 'scan' itself while it is in operation, it may formulate theories concerning its own structure and subject them to test. Suppose the machine is in a given state (say, 'state A') when, and only when, flip-flop 36 is on. Then this statement: 'I am in state A when, and only when, flip-flop 36 is on', may be one of the theoretical principles concerning its own structure accepted by the machine . . . Now all of the usual considerations for and against mind–body identification can be paralleled by considerations for and against saying that state A is in fact identical with flip-flop 36 being on.⁶

Given only the possibility that such a machine can issue reports of its abstract or computational states which do not immediately expose their relation to the *physical* states underlying them, such a machine would be justified in wondering, just as a person might, about the identities between the two kinds of states. The machine could have the same questions that a human might about whether identifying the two kinds of states would unify theory or eliminate unnecessary entities; it could even make the 'dualistic' argument that state A *could not be* identical with the state of having flip-flop 36 on because the one is, while the other is not, an 'immediately observable' or apprehensible state. The possibility of such concerns arises, in the case of the machine at least, from the distinction between two levels on which it might seek to describe itself: in terms of its abstract functional or logical states, on the one hand, and in terms of the underlying physical states that realize these, on the other.⁷

⁵ Putnam (1960: 380).

⁶ Putnam (1960: 363).

⁷ Putnam (1960: 372).

Putnam notes that this situation gives the machine a strict analogue of the distinction between 'mental' and 'physical' as it usually operates in *our* discussions of the mind-brain question. The machine's directly apprehensible and self-evident *logical* states seem, to the machine at least, to be categorically different from its non-obvious and mostly unknown physical states. Putnam even suggests that the distinction between the two levels in the machine case parallels two approaches one can take toward human psychology: the logical-level description of the machine parallels classical psychology's intuitive description of human thoughts as impressions, ideas, and other rationally organized 'mental' states, whereas the physical-level description of the machine parallels the physicalist's description of human behavior in terms that connect it to base-level physical and chemical theories.⁸ Just as in the case of human psychology, the logical-level description can be given entirely independently of the physical-level one; but also as in the case of human psychology, the physical-level description explains such deviations as may appear in the machine's behavior from the laws established by its logical-level description.

2

In a series of papers written over the early 1960s, Putnam would develop the analogy between minds and machines into a full-blown metaphysical description of mind, culminating in the decisive suggestion that our mental states simply *are* abstract states within our total functional organization. In the 1963 article 'Brains and Behavior' Putnam gave a new, and stronger, argument against the logical behaviorist identification of pains and other mental states with behaviors and behavioral dispositions. To show that there is no necessary logical link between mental states and behaviors, Putnam suggested the example of a race of people who, owing to restrictive social conventions, never describe or otherwise express their feelings of pain. These 'super-Spartans' would exhibit no pain behavior; yet it is, Putnam argued, still meaningful to say that they feel pain. For instance, it might well be possible to detect within them a distinctive neurological configuration similar to ours when we are in pain.⁹ Given this, it would make sense to conclude that they are indeed in pain. Even if their neurological states were, in general, different from ours, we could still come to conclude that they are in states enough like ours in relevant respects to be called pains.

This argument's more explicit consideration of the relationship between behavioral evidence and empirical discovery gave Putnam new resources against the logical behaviorist, but still depended on the thought that mental states ultimately are brain states. Putnam still treats states like pains as the *causes* of the behaviors that express them, and he repeats the suggestion that the grammar of pain-ascriptions is controlled by behavioral criteria that function as 'symptoms' of an underlying structure. The Turing Machine analogy makes no appearance in the article, and there is no suggestion that mental states like pains are in any sense functional or logical states distinct from underlying physical states.

⁸ Putnam (1960: 372–3).

⁹ Putnam (1963: 337).

The first impetus for Putnam's development of the Turing machine analogy into functionalism, and indeed much of the theoretical apparatus of functionalism itself, would come, instead, from the articulation of a new anti-reductionist description of psychological explanation by the young philosopher Jerry Fodor. In the 1964 article 'Explanations in Psychology', Fodor argues for the autonomy of psychological explanations from physicalist descriptions on the basis of an extended application of the functionalist model that Putnam had suggested in 1960. Arguing from assumptions strikingly unlike those of Putnam's original reductionist picture of the mental/physical relationship, Fodor suggests that the characterization of psychological states as functional states offers a reasonable model of both the logic of psychological theory and the relationship we can expect to find between it and lower-level physiological and physical descriptions.

Much of Fodor's argument for this depends on a sophisticated consideration of the structure of psychological explanation and prediction, on the basis of which he argues against an oversimple and naïve reductionist view of the relationship of such explanation and prediction to lower-level causal explanations. Psychological theory, Fodor argues, intends to explain and predict behavior; but it is misleading to suppose that this explanation and prediction can be reduced to terms any more basic or primitive than the terms of psychology themselves.¹⁰ Even the simplest notions of psychological description, for instance the behaviorist notion of a 'response', resist reduction to a physicalistic description in terms purely of physical motions. For there is no way even to characterize the set of possible physical movements that can count as a simple behavioral response without using the psychological predicate that characterizes them all as the same 'response' to begin with. The psychological description in terms of responses is not elliptical for an underlying physicalist description, but an autonomous functional description in its own right.¹¹

Even in the simple case of Skinnerian behaviorism, Fodor argues, the grammar of psychological explanation makes ineliminable use of terms that cannot be defined physicalistically; even if explanation on this level is partly causal, what is important in understanding its logic is not definitional reduction of psychological to physical predicates but a functional characterization of the relations of definition and causality *among* psychological terms and their referents.¹²

But what, exactly, is a 'functional' characterization, and what is the relationship between a 'functional' description and a straightforward causal description if one

¹⁰ Fodor's doctrine thus has motivations that parallel, and somewhat overlap, those of Davidson's 'anomalous monism' about the mental, the view that although each (token) mental state is in fact identical with a token physical state, there are no strict psychophysical laws connecting the two types of states. The classic expression of this view is Davidson (1970), and the 'type-token' distinction suggested here would soon give philosophers a natural language in which to express and investigate the insight of functionalism.

¹¹ Fodor (1964: 168).

¹² A visible influence on Fodor's thinking here is Chomsky's (1959) review of Skinner's *Verbal Behavior*. In it, Chomsky argues that the Skinnerian notions of stimulus, response, and reinforcement, however well defined they may be in the context of particular experiments, resist extension to real-life behavior. Like Fodor, Chomsky argues that there is no helpful reduction of the Skinnerian notions to physicalistic terms. Fodor supplements this realization, however, with the suggestion that the Skinnerian notions *do* characterize the organism under consideration on an autonomous level of functional description.

does not reduce to the other? Picking up on Putnam's suggestion, Fodor argues that psychological explanation has two levels or 'phases'.¹³ On the first phase, corresponding to classical psychology, mental states are characterized in irreducibly psychological terms according to their roles in producing behavior. Importantly, at this level of explanation, the explanatory use of descriptions of mental states requires no reference to the underlying physical mechanisms that correspond to or realize them:

It should be noticed that explanations afforded by phase one theories are not causal explanations, although a fully elaborated phase one theory claims to be able to predict behavior given sufficient information about current sensory stimulations. Phase one explanations purport to account for behaviour in terms of internal states, but they give no information whatever about the mechanisms underlying these states. That is, theory construction proceeds in terms of such functionally characterized notions as memories, motives, needs, drives, desires, strategies, beliefs, etc. with no reference to the physiological structures which may, in some sense correspond to these concepts.¹⁴

By postulating intuitively described inner states like motives and memories, phase-one explanations, Fodor suggests, allow us to predict and explain behavior in a wide variety of situations; all that is required to formulate them is the observations we make of the behavior that people and other organisms produce in response to stimulations. Still, they give us no insight into underlying physiological mechanisms that are literally responsible for causing the behavior in question. For this we need a second phase of explanation, on which we specify the *mechanisms* that actually underlie our functionally defined phase-one states. Applying Putnam's machine analogy again, Fodor notes that any given functional-level explanation corresponds to *indefinitely many* mechanical-level explanations:

In a phase one explanation, we picture the organism as proceeding through a series of internal states that terminate in the production of observable behaviour. But we make no attempt to say what these states are states of: what internal mechanisms correspond to the functionally defined states we have invoked. Now, the set of mechanisms capable of realizing a series of such functionally defined states is indefinitely large. Only our ingenuity limits the number of mechanisms we could devise which, upon exposure to the relevant stimulations, would go through a sequence of internal states each functionally equivalent to a corresponding state of an organism and would then produce behaviour indistinguishable in relevant respects from the behaviour of the organism.¹⁵

The character of the relationship between mechanical-level explanations and functional-level ones has a number of significant consequences for the growth of psychological theory. First, Fodor suggests, mechanical-level explanations may help to suggest new functional-level ones; for instance, speculations about the neurology of memory might lead to new functional-level characterizations of memory in terms of familiar psychological notions. Second, mechanical-level explanations constrain functional-level ones; though each functional system has an indefinite number of mechanical realizations, any functional explanation that is *inconsistent* with the mechanical-level explanation of the same system can be dismissed.¹⁶ Additionally, the one-many relationship between

¹³ Fodor (1964: 171–4).

¹⁴ Fodor (1964: 173–4).

¹⁵ Fodor (1964: 174).

¹⁶ Fodor (1964: 176).

functional-level and mechanical-level explanations implies a non-reductive picture of the relationship of mental to physical states. If psychological explanation really does have the two-phase structure of Fodor's account, then 'reductions,' if there are any such, from the mental to the physical are not mereological decompositions of higher-level entities into their lower-level parts. Instead, they correlate *functions* with *mechanisms*, explaining the functional role played by a mental state by referring to the mechanism enabling it to play that role.

Beginning with considerations of the logic of psychological theory and the unlikelihood of its reduction to physical theory, then, Fodor's article succeeded in defining 'ordinary' or classical psychology as the functional description of internal states of an organism, a description which, in each case, may correlate with any number of mechanical-level descriptions of the same organism couched in the language of neuroscience and physiology. This suggestion led Putnam to define and articulate, over the next five years, the thesis that a mind might simply *be* a system of functional states realized physically. In his articles defining and defending functionalism, Putnam significantly extended and developed Fodor's consideration of the logic of psychological explanation, and drew out its consequences for the philosophical question of the mind-body relation. These consequences would lead Putnam to move decisively beyond the identity theory, as well as to repudiate much of the semantic argument he had formerly deployed in its defense.

Putnam went on to define the functionalist theory of mind in three articles: 'Robots: Machines or Artificially Created Life?' (1964), 'The Mental Life of Some Machines' (1967), and 'The Nature of Mental States' (1967). In these articles, Putnam's arguments for functionalism fall into four main types.

First, there are arguments, akin to Fodor's, from *the logic of psychological terms*. Psychological terms, if they are definable at all, are only *inter* definable; there is no hope of 'unpacking' the definitions of psychological terms into behaviors or behavioral dispositions that are not themselves psychologically described.¹⁷ This suggests that psychological descriptions do *not*, as the identity theory had held, covertly or elliptically refer to physical internal states, and indeed that the hope of defining a physicalist research program culminating in the identification of the physical referents of ordinary psychological description is largely misguided.

A second sort of Putnamian argument for functionalism grew from his earlier arguments against logical behaviorism, particularly the argument that there is no logically necessary link between behavior and mental states.¹⁸ Because the functional states of a Turing machine need not necessarily correspond to or even be determinable on the basis of behavior, it is possible to construct a machine analogue of the 'super-Spartans', a machine which is often in a particular functional state but will not express that it is. Since formal rules govern the transitions between a Turing machine's logical states, it is possible to implement rule-governed 'preference-functions' for the Turing machine. These rules can govern the self-expression of the machine's states; so given an abnormal preference-function (for instance, one that places an infinitely high disvalue on

¹⁷ Putnam (1964: 391).

¹⁸ Putnam (1967a: 421–2); a similar argument is suggested at Putnam (1967b: 438–9).

expressing that it is in the state functionally defined as 'pain') the Turing machine could 'experience' functional states that it does not behaviorally express. Thus, functional states, like our mental states, need not be logically linked or interdefined with behavior. This recommends the functionalist account, and shows that it survives at least one of the objections that doomed logical behaviorism.

This shows the logical difference between functionalism and behaviorism; but a third sort of argument Putnam uses for functionalism actually suggests a surprising amount of commonality in philosophical motivation between the two theories. Even if functionalism allows that functional states—and hence mental states—need not be identifiable with or logical constructions from behavior, nevertheless the consideration that our *criteria* for the everyday ascription of mental states are largely behavioral provides an argument in favor of functionalism:

Turning now to the considerations *for* the functional-state theory, let us begin with the fact that we identify organisms as in pain, or hungry, or angry, or in heat, etc., on the basis of their *behavior*. But it is a truism that similarities in the behavior of two systems are at least a reason to suspect similarities in the functional organization of the two systems, and a much *weaker* reason to suspect similarities in the actual physical details. Moreover, we expect the various psychological states—at least the basic ones, such as hunger, thirst, aggression, etc.—to have more or less similar 'transition probabilities' (within wide and ill defined limits, to be sure) with each other and with behavior in the case of different species, because this is an artifact of the way in which we identify these states. Thus, we would not count an animal as *thirsty* if its 'unsatiated' behavior did not seem to be directed toward drinking and was not followed by 'satiation for liquid.' Thus any animal that we count as capable of these various states will at least *seem* to have a certain rough kind of functional organization.¹⁹

Even if mental states are not logically dependent on, or identifiable with, public behavior, it nevertheless remains a philosophically significant feature of the logic and grammar of our common-sense and classical psychological theories that we *ascribe* mental states on the basis of publicly observable behavior. Moreover, the connection between the observation of behavior and the ascription of a mental state is, as Putnam realizes, closer and tighter than the connection between evidence and theory. For as a matter of logical necessity (at least in an extended sense of that term), we will not ordinarily be prepared to *call* an organism 'thirsty,' 'hungry,' 'enraged,' etc., if it does not exhibit *any* of the behavior that is criterial for that particular ascription. Under normal circumstances, the proposition that Jones is angry, if he exhibits *none* of the usual behavioral signs of anger, will at least call for further clarification. As Putnam had earlier argued, the logical behaviorist takes this kind of logical connection between behavioral evidence and the determination of mental states to be stronger than it is, forgetting that there are, after all, *some* conceivable circumstances under which mental states might reasonably be ascribed in the absence of their usual behavioral symptoms. Still, its behavior is *prima facie* good evidence for an organism's having a particular functional organization; and many, if not all, functional states are primarily characterizable in virtue of their logical relationships to publicly observable behaviors.

¹⁹ Putnam (1967b: 437).

Finally, the observation that functional states are partly characterized by their relationship to, and ascribed on the basis of, behavioral evidence suggests what is Putnam's most often cited and characteristic argument for functionalism, what has been called the 'multiple-realization' argument.²⁰ It begins as an argument *against* the identity theory. The identity theorist, Putnam argues, is committed to the *identification* of a particular mental state, say pain, with a particular neurological or neurophysiological structure found in all and only those organisms that are currently feeling pain. Moreover, this identification, if the identity hypothesis has any explanatory force, must be at least nomologically necessary. Whatever state is to be identified with pain must exist, then, in mammalian and molluscan, human and extraterrestrial brains alike, and moreover must be correlated, as a matter of scientific law, with the behavioral manifestations of pain in all of these species. Of course, it is extremely unlikely that any such state exists. What all and only organisms that are in *pain* do share, though, is a certain *functional* state that can be characterized by its logical and causal interrelationships with other functional states (moving away from a particular stimulus, acting as one has acted when physically damaged in the past, etc). Where the identity theory necessarily posits an underlying state that could hardly exist (or, anyway, be theoretically useful; we could, of course, refer to all the biologically distinct states that realize pain in various organisms as a single, wildly disjunctive state), the functional-state theory uses what we already know about the logical criteria on the basis of which mental states are ascribed and discussed to characterize them as functional states that *could* be held in common by a wide variety of possible organisms and systems.

The multiple realization argument has often been considered a decisive argument in favor of functionalism, but it is important to be clear about just what sort of argument it is. Even if the identity theory fails because it requires nomological connections between mental states and (possibly hugely disjunctive) brain states, the functional-state theory improves upon it in this respect only because the specification of a functional state has no particular consequences for the identity of the underlying physical states. The thought that a given functionally characterized system can be realized by any of an indefinite number of possible physical systems had been suggested in passing in Putnam's 1960 article, and Fodor had made it the basis of his anti-reductionist picture of the relationship of phase one to phase two psychological explanations. Following Fodor's suggestion, Putnam clearly thought of the one-many relationship between functional and physical descriptions as one of the most crucial recommendations of the functionalist program. Unlike the nomological identities required by the identity theory, the one-many structure of functionalist explanation allowed that the meaning of ordinary psychological descriptions does not depend, overtly or covertly, on their reference to esoteric neurological or physiological facts. On the level of functional explanation at least, the functional-state theory defines a much more plausible research program: rather than having to determine the underlying physical 'identities' of the entities invoked in our psychological explanations, we treat these entities as well-defined from the outset and simply attempt to characterize further their

²⁰ Its most usually cited version is Putnam (1967b: 436–7); compare Putnam (1964: 392–3) (quoted below) and Putnam (1969: 451).

functional roles, employing only such evidence as is available publicly and prior to the detailed investigations of the brain sciences.

The force of the multiple realization argument, then, does not arise as much from the failure of the identity theory to handle species-specific mental–physical correlations as from the ability of functionalism to define a program of psychological investigation which takes much greater and more sophisticated account of the evidentiary and causal logic of traditional psychological explanation. Were it only the first, defenders of the identity theory could simply respond, as Kim (1972) in fact did, that even if pain is realized in *various* ways in *various* different species, species-specific identity laws are enough to prove the identity theorist's case. Putnam resisted this position not because he thought it would be impossible to identify the species-specific physical 'correlates' of pain in each particular case, but because he thought such identification would have little relevance on the level of traditional psychological explanation and, accordingly, little to do with defining the *identity* of pain.

As Putnam began to define and articulate the view that mental states simply *are* functionally defined states, the one–many character of the functional-state/physical-realization relation became central to his thought about the metaphysical status of the mind, causing him to abandon some of the most important parts of his earlier picture of explanation and reduction. In fact, the thought that a functional description of the psychology of an organism has *no* consequences for the nature of its realization led Putnam to doubt physicalism itself. Since the functional-state hypothesis, as Putnam understood it in 1967, defines a mental state *simply* in terms of an abstract functional description, it has no consequences whatsoever for the nature of the medium realizing it. Functionally defined states are completely logically independent of their realizers. This gives the functionalist reason to doubt not only the identity theorist's 'definition' of mental states in terms of physical states but even materialism itself, as Putnam shows with another argument arising from the possibility of multiple realization:

Indeed, there could be a community of robots that did not all have the same physical constitution, but did all have the same psychology; and such robots could univocally say 'I have the sensation of red', 'you have the sensation of red', 'he has the sensation of red', even if the three robots referred to did not 'physically realize' the 'sensation of red' in the same way. Thus, the attributes having the 'sensation' of red and 'flip-flop 72 being on' are simply not identical in the case of the robots. If Materialism is taken to be the denial of the existence of 'nonphysical' attributes, then Materialism is false even for robots! (pp. 392–3)

As Putnam remarks elsewhere, the functional-state theory is not even incompatible with dualism: even a nonphysical 'soul' could perfectly well 'implement' any given functional organization, as long as it has a number of logically distinct and temporally successive states. And even in the case of an actual, material Turing machine, its functionally defined states are logically distinct from, and not derivable from, *any* of its physical states or attributes. In this respect at least, they are genuinely 'non-physical', defining real and ascertainable attributes above and beyond the set of all of the machine's physical attributes and all of their logical consequences.

Putnam's goal in making these points against materialism, of course, was not to argue for dualism or some new account of the metaphysics of mind, but to suggest the

emptiness, given the functionalist picture, of all traditional philosophical descriptions of the mind–body relation. As in the 1960 article, where Putnam had pointed out how analogues of *all* of the traditional philosophical positions on the mind–body problem would arise for a mechanistic Turing machine, his conclusion is not that functionalism recommends a new and interesting position, but that anyone seriously interested in the metaphysics of a functionally defined system like a Turing machine is wasting his or her time.

In sum, then, the functional-state identity theory, as developed by Fodor and Putnam through 1967, offered a theoretically innovative account of the mind–body relation that exhibited much greater sensitivity to the logic of psychological explanation than had its predecessors, abandoned the oversimple reductionism of previous versions of physicalism, and at least suggested that traditional pictures of the mind–body relation should be foregone or radically overhauled. Its respect for the criteria ordinarily employed in psychological description allowed it to characterize ordinary talk about mental states as straightforwardly meaningful, without involving it in the semantic contortions suggested by Putnam's original investigation of diachronic theory change and the logic of theoretical claims of identity. But while the multiple realization argument allowed the liberation of autonomous psychological description from the demand of physicalistic reduction, it also left deeply unclear the metaphysical nature of the relationship between physically and functionally defined systems. This led to the admitted possibility that functionalism might not be 'materialist' after all, as well as to the reality that functionalist description failed to define any program lending insight into the physical and physiological causation of behavior.

Functionalism earned its plausibility from its closeness to the logic of ordinary and classical psychological explanation. But, at least as it stood so far, it bought this closeness at the price of the kind of metaphysical specificity that would have been needed to fundamentally clarify the relationship of philosophical description of mental states to empirical discovery of their physical correlates. Methodologically, it rewrote what was common to the traditional analytic project of logical analysis of mental states and the phenomenological project of formal analysis in an idiom that avoided the excesses of logical behaviorism, but in so doing lost the empirical-mindedness of the identity theory. Despite years of concerted thought on the part of functionalists and their predecessors, the semantic and conceptual analysis of the logic of mental states still threatened to float free of any clear application to the newly developing cognitive sciences of mind and brain.

Though functionalism is often taken to be a naturalist theory inviting a purely physicalist ontology, the historical overview shows that its chief methodological motivation was a practice of logical analysis or clarification that is in fact far removed from empirical work. Attention to functionalism's method of logically analyzing our terms of mentalistic description therefore suggests its closeness to the central methods of the phenomenological tradition's conceptual and logical analysis of the structure of

experience. Phenomenology's analysis of experience is *logical* in that it is grounded in ideal, formal structures of meaning; and it is *conceptual* in that it depends crucially on reasoning about the definitional and explanatory interrelationships among our concepts of experience, the same concepts whose logical interrelationships Putnam and Fodor sought to capture with the idea of a functional state description. The substantial historical and conceptual continuity between functionalism and phenomenology in this respect casts light on hidden features of functionalism's conceptual structure, and suggests a new way of viewing the contemporary problem of the recalcitrance of conscious experience to functionalist description.

Throughout the development of his phenomenological project, Husserl looked to analyses of the formal structure and meaning of our psychological concepts to underwrite the explanation of the nature of consciousness and the constitution of objects of awareness. Beginning in the *Logical Investigations*, he envisioned phenomenological investigation as the investigation of the ideal 'laws of essence' that characterize both the structure of our conscious experience and the structure of our concepts of it.²¹ Phenomenology, as the investigation of the formal structure of experience, was to be grounded in 'pure logic', a universal system of rules governing the actual and possible combinations of objects, meanings, and phenomenal properties to comprise our experience of the world. Here, Husserl connected the project of phenomenological analysis, on the one hand, to the establishment of 'pure categories of meaning' that govern the semantic possibilities for using linguistic terms, and on the other to his vision of a pure phenomenology purged of empirical psychology by the logical purity of its terms of description.²² In *Ideas I*, Husserl further expanded on the nature of phenomenology's investigation of laws of essence and their grounding in the ideal synthetic acts of the transcendental consciousness, suggesting at one point that the phenomenological investigation of transcendental constitution is essentially the investigation of a functionally defined structure.²³ In this work and elsewhere, Husserl argues that the formal, essential laws governing possible perceptual and cognitive structures as well as defining the ontology of material and other objects result from immediate, given experiences by way of a complicated process of formal abstraction. Indeed, since our access to empirical objects of perception and observation always depends upon their being given in intuition, every concrete act of perception involves the synthesis of unformed 'hyletic' or phenomenal data into a conceptually formed complex. It follows that the phenomenological analysis of the structure of experience is at the same time the formal analysis of our concepts of experience, and that the conceptual or logical interrelationships among these concepts are a main source of evidence for the analysis. Though the ideal laws of essence that govern the structure of experience are conceived as abstracted from both concrete experience and language, phenomenology's primary method of insight into them is logical-level reflection on the structure of the concepts with which we characterize it.²⁴

²¹ Husserl (1900).

²² Husserl (1900); see, e.g., 'Prolegomena', section 67, and Investigation I, section 2.

²³ Husserl (1913), section 86.

²⁴ 'The objects which pure logic seeks to examine are, in the first instance, therefore given to it in grammatical clothing' (Husserl 1900, section I.2). See also *Ideas I*, section 11.

The continuities between phenomenology's logical investigations of experience and Putnam and Fodor's functionalism, moreover, are not only conceptual; for a significant line of historical influence on Putnam and Fodor's method also originates with Husserl's phenomenology. As early as the 1920s, logical empiricists like Carnap had sought the conceptual reduction of sentences involving psychological terms of description to logically prior terms describing immediate experience, citing Husserl's analysis of experience explicitly as an antecedent of their own project.²⁵ When it became clear to these philosophers that they could not give a unified description of the structure of propositions describing immediate experience itself, they suggested instead the analysis of psychological descriptions into logically prior descriptions of behavior, yielding the project of 'logical behaviorism.'²⁶ In the late 1940s, Gilbert Ryle's *The Concept of Mind* propounded an influential program for the investigation and clarification of our ordinary concepts of mentality. Though Ryle's main aim was to dispel the sources of the Cartesian dualist's theory of mind as a non-physical entity causally connected to the purely physical, mechanistic body, he also described his own work of 'conceptual geography' as a 'sustained essay in phenomenology' aiming to reveal the logical categories of our ordinary language of psychological description.²⁷ As we have seen, Putnam and Fodor essentially drew upon the logical behaviorist method of analysis of mentalistic terms in their own development of functionalism, even while repudiating the suggestion of any possibility of reducing the reference of these terms to patterns of behavior. In making use of this practice, they consistently applied much the same method of conceptual or logical analysis that Husserl had made the central methodological innovation of phenomenology. Putnam's central thought that a conceptual analysis of our concepts of experience and mentalistic explanation could yield a purely structural, functional description of the mind, abstracted from the particular mechanical operation of any specific mind or brain, was thus already a central component of Husserl's project. This basic phenomenological thought would continue to guide the development of functionalism, subsuming its appeal to a level of formal description of minds largely independent of, and only problematically related to, the level of physical and causal description of the brain.

Its continuities with phenomenology cast light, as well, on another important aspect of the conceptual structure of functionalism. As we have seen, much of the supposed advantage of functionalism over logical behaviorism rested on its ability to capture the results of the *empirical* investigation of the mind and brain and thereby ensure the possibility of a purely naturalistic method and a physicalist picture of the world. This at first seems to stand in stark contrast with the method of phenomenology; for Husserl famously and strenuously opposed any 'naturalization' of phenomenology's inquiry into essence, holding that it could be no part of pure phenomenology to invite or depend on empirical results.²⁸ But historical reflection on the development of functionalism shows that this difference is in fact more apparent than real. For throughout its development, the main source of evidence for

²⁵ Carnap (1928). ²⁶ See, e.g., Carnap (1931) and Hempel (1935).

²⁷ Ryle (1949); Ryle (1962: 188). ²⁸ See, e.g., Husserl (1911).

functionalism was non-empirical reflection on the structure of our concepts, a kind of reflection which is equally at home in phenomenology and in logical analysis. As we have seen, this conceptual work sat, initially at least, in some tension with functionalism's claim to be a thoroughly empirical theory; despite prognostications to the contrary, it was not clear how a functional description funded by the logical interrelationships of concepts could fit with a physical-level description of our physiological states. The tension reflects many of Husserl's own concerns about the prospects of providing a naturalistic basis for phenomenology, concerns, in particular, that no natural or empirical basis could preserve the necessity and *a priori* character of the phenomenological laws of essence and meaning. Though the subsequent development of the analytic tradition witnessed the widespread denial of any principled analytic/synthetic distinction, and accordingly of the kind of necessity that had formerly been thought to accrue to the results of pure conceptual analysis, the underlying methodological tension between conceptual description and empirical discovery remained unresolved. This tension, as we have seen, continued to affect the coherence of the theory of functionalism; and although it would be superficially healed on the level of doctrine a few years later, its effects continue to problematize functionalism's account of experience even today.

Of course, missing from functionalism's explicit theoretical vocabulary was any analogue to phenomenology's central appeal to subjective experience or intuition as the foundation and ground for all of our intentional acts and processes of cognition. For Husserl, the analysis of our concepts of experience could only be the analysis of structures founded on a basis of immediate, lived experience and abstracted from it; without this basis, no theory of mind or indeed of reality could be considered complete. Functionalism's chief divergence from phenomenology might, then, seem to consist in its apparent refusal to characterize subjective experience *itself* independently of our conceptualization of it; and this might easily be thought responsible for functionalism's apparent inability to provide an explanation of consciousness. But as the historical overview shows, concerns about the nature of subjective experience in fact played a central role in the development of functionalism, and decisively influenced its theoretical shape. It was, after all, the inability of logical behaviorism to explain our use of language referring directly to experience that most centrally inspired Putnam to reject it; and it was the problem of the meaningfulness and reference of our descriptions of experience that led him to propose the notion of a functional-state description as an alternative. In its very structure, therefore, functionalism retains a determinative concern with the nature of subjectivity and the special problems of theorizing it by means of conceptual, structural, or broadly logical analysis. If these underlying problems indeed explain functionalism's apparent inadequacy for explaining consciousness, then it seems likely that they arise equally, and as prominently, for phenomenology's parallel project. For both projects, the nature of immediate, unconceptualized experience poses a continuing problem for forms of explanation that are inherently conceptual and structural in nature. The underlying problem can be concealed, but not removed, by the foundationalist rhetoric of a basis of conceptual thought in subjective experience; Husserl's project, as much as Putnam and Fodor's, remains open to the same underlying tension.

In 1965, in a brief and crisply argued *Journal of Philosophy* article, David Lewis proposed a philosophical innovation that, when added to the functional-state theory as Putnam had defined it, completed the theory from a logical (if not a chronological) point of view, effectively ended the further metaphysical speculation that might otherwise have been engendered by the unclarity of Putnam's account, and defined much more specifically the kind of relationship between philosophical analysis and empirical discovery that could be expected on a functionalist theory. Despite its functionalist motivation, Lewis called his article 'An Argument for the Identity Theory'. But its central innovation was essentially a semantic one: that mental states, and in particular 'experiences', are *defined* by their *causal roles*, their pattern of typical causes and effects. With this innovation, Lewis made it possible to maintain that the functional roles definitive, according to functionalism, of mental states are *at the same time* causal roles, and therefore that the place of a mental state in our ordinary and classical psychological descriptions adverts to, and locates it in the total theory by means of, the *same* properties and features that locate it in the total causal web of physicalistically described nature.

If the suggestion is accepted, the logical analysis of the grammar of the ordinary description of mental states will henceforth be an *integral part* of the empirical analysis of the underlying physical states, for the semantic features of mental-state terms will mirror the causal roles in virtue of which their bearers can be identified with physical states. What had seemed to be purely 'logical', 'grammatical', or 'phenomenological' analysis will then have a new richness of empirical relevance; the structure of the traditional philosophical investigation of the relational logic of mental states will be mirrored as the empirical investigation of the causal relations of functionally defined states. Methodologically, the innovation of Lewis' account would add to Putnam's functionalism the most philosophically compelling features of Smart's physicalism: its explanatory and metaphysical economy, its sensitivity to the possible philosophical relevance of new discoveries in the reductive brain sciences, and its congeniality to an uncompromisingly physicalist picture of the world with no suggestion of esoteric non-physical or mental facts, properties, or entities.

Together, Lewis' arguments for his suggestion recommend a position that recognizably combines the two distinct levels of analysis that Fodor had originally suggested: experiences are defined, Lewis suggests, as causal roles, and particular physical states, as a matter of contingent fact, are the *occupiers* of those causal roles. Experiences are defined, and spoken of, as the patterns of what causes them and what they cause, but it is ultimately particular physical states of the brain that are doing the causing. Thus, the contingent identities of the identity theory fit right alongside the analytic or near-analytic analyses of functional description; one side of the account constrains the other in that only something that really can do the causal work of a particular experience is a candidate for contingent, species-specific identity with that experience.²⁹

²⁹ Lewis (1969).

In 1968, David Armstrong would make much the same suggestion the centerpiece of his influential *A Materialist Theory of Mind*. Like Lewis, Armstrong aims to defend a sophisticated version of Smart's identity theory. But he argues that the identity theorist's identification of mental states with brain states ought to be augmented with specific analyses of our mental concepts much like the analyses suggested by logical behaviorists. The two strands of theory can be joined, Armstrong suggests, by recognizing that 'the concept of a mental state is primarily the concept of a state of the person apt for bringing about a certain sort of behaviour' (p. 82). On Armstrong's suggestion, then, mental states are identified in terms of the types of behavior that, under ordinary or appropriate circumstances, they normally cause.³⁰ As on Lewis' view, this allows the proponent of the identity theory to accept much of the logical behaviorist's analyses of mental concepts into behavioral facts, without denying that the objects of mental concepts are brain states:

I have emphasized that the argument put forward for a Materialist theory of mind involves two steps. In the first place, it is argued that a mental state is a state of a person apt for the bringing about of behaviour of a certain sort. This is intended to be a piece of logical analysis. In the second place, it is argued on general scientific grounds that this inner cause is, as a matter of fact, the brain. (p. 116)

To motivate the first component of the argument, Armstrong goes on to offer logical analyses of the concepts of willing, knowledge, perception, and mental images into the kinds of behavioral and public facts apt to cause them and be caused by them.³¹

With the Lewis/Armstrong suggestion, logical-level functional analysis and empirical-level discovery of psychophysical identities fall cleanly into their relative places in a comprehensive program of jointly functional and causal analysis. In addition to defining a realistic research program combining logical and causal analysis, moreover, the suggestion effectively quells any remaining doubts about the extent of functionalism's compatibility with physicalism. If the relationship between functionally defined states and their physical realizers is the relationship between a causal role and its occupant, then ontologically we need not countenance anything more than objects and causally interrelated events. From this perspective, there is no danger that functionally defined states, because *logically* distinct from their physical realizers, will be in any interesting or relevant sense 'non-physical' or represent any obstacle to a materialist description of the world. But at the same time, Lewis' suggestion, because it depends only on the

³⁰ Armstrong (1968: 83).

³¹ It is worth noting, though, one slight difference of emphasis between Armstrong's and Lewis' ways of putting the point. Lewis draws more clearly than Armstrong does the distinction between causal roles and their contingent occupiers. This allows him to envision a program comprising two clearly distinct levels of analysis: first, the logical description of causal roles, and second, the empirical identification of their occupiers. Armstrong, by contrast, does not draw the role/occupier distinction and therefore often seems to consider the causal relationship between a physicalistically described brain state and a mental state to be logically on a par with the causal relationships among mental states. Although this leads to a less well-defined distinction between the two components of the suggested analysis, it also allows Armstrong's analysis more room to exploit the suggestion that the logic of many of our concepts of mental phenomena—for instance the concepts of sensation—already implies that they are caused by internal brain states, even before empirical results are available to verify this implication.

physical explicability *in principle* of every physical event, does not obviously demand or imply the oversimple reductionist picture of psychological explanation that Fodor had originally resisted. The relationship between a causal role and its contingent occupier, unlike the compositional relationship between a macro-level object or process and its micro-level constituents, is plausibly a relationship characterized by some degree of explanatory autonomy. Because various structures may accomplish one and the same causal role, the explicability in principle of each physical event does not demand, on this picture, that there be, in general, any univocal or nomological relationship of explanatory reduction between an experience and the physical state with which it is (contingently) identical.³²

Much of the subsequent discussion of functionalism over the last thirty-five years can be traced to issues left open in the final configuration comprised of the combination of Putnam's functional-state theory and Lewis' suggestion. But functionalism has encountered its greatest obstacles in its description of the nature of *consciousness*. In 1972, together with Ned Block, Fodor first expressed cautious doubts about the ability of functionalist description to explain subjective, phenomenal, or conscious states, and in recent years these doubts have grown into a widespread position of resistance among philosophers who doubt that a functionalist explanation of consciousness can be correct.³³ This situation cannot be viewed without a certain level of historical irony, in that for Putnam and Lewis alike it was the facility of functionalism in describing the nature of subjective states like pains and other experiences that first, and most primarily, recommended it as a systematic description of the mind. Here, as at other moments in the history of twentieth-century thought about consciousness, what began as a systematic way of capturing the logic of the terms with which we describe experience became a structural pattern of explanation that seemed to deny the unstructured immediacy of experience itself. The very virtue that had originally recommended the functionalist theory over logical behaviorism—its ability to make reference to genuine inner states—involved functionalism in a complete structuralist explanation of these states that seems inadequate to capturing their immediacy, intuitiveness, and spontaneity.

5

Historically viewed, Lewis' suggestion saved (or prevented) Putnam's functional-state theory from inaugurating what might otherwise have become a far-reaching investigation, inspired by the special question of the relationship of physical to functional states, of the metaphysical relationship of minds to machines, rules to causes, and the logic of psychology to the logic of physics. By showing how functional analysis could at the same time be causal analysis, Lewis allowed the analytic program of logical analysis to continue in a new—and newly empirically respectable—form, while

³² For more on the logic of Lewis' suggestion for the nature of theoretical identifications and comments on its relation to Putnam's developing account, see Lewis (1972).

³³ See Block and Fodor (1972).

guaranteeing the amenability of functionalist description to the prevailing physicalist picture of explanation and ontology. In this respect, Lewis' suggestion ameliorated the fundamental unclarity of Putnam and Fodor's unaugmented picture on the relationship of functionally defined states to their physically defined realizers. But viewed historically, his suggestion has something of the character of a solution by fiat, a pragmatic suggestion that allowed philosophical discussion to continue in an empirical domain but left many outstanding, and important, philosophical issues internal to its doctrine unresolved. Despite the physicalism of Lewis' account, the underlying suggestion that experiences are defined by their causal roles improves little, as Lewis himself recognizes, over the logical behaviorist's claim that mental states are logical constructions from publicly observable behaviors or dispositions to behave.³⁴

Methodologically, then, the historical investigation reveals functionalism as a hybrid doctrine, born of the competing demands of conceptual, logical, and phenomenological analysis, on the one hand, and allowance for specialized empirical discovery, on the other. But the most philosophically significant suggestions made by a perspicuous representation of the history of the development of functionalism concern its relationship to consciousness. The historical overview reveals that the functionalist description of the mind arises primarily from logical-level attention to the explanatory structure of psychological theory and description. Whatever their causal implications, functionalist analyses of particular mental states often remain essentially semantic or conceptual analyses, characterized by a descriptive and explanatory logic largely independent of the logic of causal description. This suggests that the origin of the problem of functionalizing consciousness may *not* lie in some special feature or property of conscious states in their causal interrelationships with other physical events. Instead, the underlying problem may arise from the special *logical* features of conscious states that make them uncongenial to functional-level description. If this is the case, then the specially problematic features of consciousness are not special 'properties' of consciousness as an empirical phenomenon, but logical and conceptual features of the relationship of our descriptions of consciousness to our descriptions of other kinds of mental states and physical events.

Recognition of the substantial continuity between functionalism and earlier projects of logical, phenomenological and conceptual analysis allows the underlying problem with the explanation of consciousness to emerge in its full generality. In the light of historical interpretation, the complaint that consciousness resists functionalist explanation emerges as one instance of a more general and perennial phenomenon: the resistance of subjective experience to broadly *structuralist* practices of conceptual and logical analysis. Consistently throughout the history of twentieth-century attempts to theorize the mind, the structural form of conceptual or logical analyses of our concepts of experience has seemed to run counter to the demands of accounting for the immediacy of experience itself. This tension, the present investigation suggests, points to a deep and unresolved problem about the relationship of conceptual structure

³⁴ Lewis writes: 'Yet the principle that experiences are defined by their causal roles is itself behaviorist in origin, in that it inherits the behaviorist discovery that the (ostensibly) causal connections between an experience and its typical occasions and manifestations somehow contain a component of analytic necessity' (pp. 20–1).

to the immediate matter or content of subjective experience, a problem that vexes every systematic attempt to define or characterize experience in formal or structural terms. Phenomenology's 'ideal laws of essence' are themselves defined structurally by attending to the logical connections among our concepts of experience; in this respect, the form and method of the analytic tradition's logical analyses of our psychological concepts incur, despite phenomenology's central appeal to subjective experience, much the same tension. Against the backdrop of the deep tension, revealed by the historical analysis, between experience and structural and conceptual forms of explanation, the characteristic Husserlian appeal to an abstractive foundation of our concepts in immediate, unconceptualized experience emerges as deeply problematic.

In this way, the historical investigation of the development of functionalism clarifies one kind of parallel between the characteristic methods of analytic philosophy and one of the chief projects of the continental tradition. The parallel suggests that the problems about the nature and adequacy of our concepts of experience that led to the development of functionalism are substantially continuous with the problems that led Husserl increasingly to distance himself, in his last work, from any foundational appeal to conscious experience, preferring to speak of the foundation of theoretical practice in the untheorized everyday 'life-world'.³⁵ Much the same set of problems, moreover, presumably underwrote the tendency of Husserl's phenomenological successors, Maurice Merleau-Ponty and Martin Heidegger, to move away from Husserl's characteristic appeal to immediate experience and subjectivity as the foundation for all abstract conceptualization, and toward versions of the phenomenological project that no longer rely on individual experience or subjectivity as a theoretical basis.

More generally, identifying the real underlying form of the problem of explaining consciousness allows us to perceive the problem's true significance for our ongoing attempts to understand our own nature. For the recurrent problem with the explanation of consciousness is an enduring problem about the relationship of subjective experience to forms of explanation and analysis that are otherwise comprehensive. Viewed in the general light that the historical investigation makes possible, the problem with consciousness that both led to the development of functionalism and continues to trouble it is just the problem of our own relationship to structural forms of formal and scientific explanation. These forms of explanation—logical, structural, and causal—subsume a great amount of knowledge within a single, unified framework of objectivity. The protest against the functionalization of consciousness manifests the underlying thought that subjectivity itself cannot be captured within this structure, that our immediate experience systematically resists inclusion in its web. With this clarified, the complaint of contemporary philosophers who hold that consciousness resists functional explanation can emerge as the protest that it is: a protest, in the name of the distinctiveness of our own inmost nature, against the inclusion of this nature within an abstracted, total picture of the world in terms of its logical, conceptual, or causal structure. Against this totalizing picture, the complaint gestures toward the immediacy and irreducibility of subjective experience; but the historical investigation provides the beginning of the conceptual resources needed to identify

³⁵ Husserl (1937).

the complaint's genuine ground, rather, as the logical nature of the subject, a nature whose logical peculiarity will constantly tempt us to describe it, even as it continues to resist any such description.

REFERENCES

- Armstrong, D. M. (1968) *A Materialist Theory of the Mind* (London: Routledge).
- Block, N. and Fodor, J. A. [1972] (1980). 'What Psychological States are Not', Ned Block (ed.), *Readings in the Philosophy of Psychology*, vol. 1 (Cambridge, MA: Harvard University Press), 237–50.
- Brandt, R. and Kim J. (1967) 'The Logic of the Identity Theory', *Journal of Philosophy* 64/17: 515–37.
- Carnap, R. [1928] (1967). *The Logical Structure of the World*, trans. Rolf A. George (Berkeley: University of California Press).
- Carnap, R. [1931] (1959) 'Psychology in Physical Language', in A. J. Ayer (ed.), *Logical Positivism* (New York: The Free Press).
- Chalmers, D. (1996) *The Conscious Mind* (New York: Oxford University Press).
- Chomsky, N. [1959] (1980) A review of B. F. Skinner's *Verbal Behavior*, in Ned Block (ed.), *Readings in the Philosophy of Psychology*, vol. 1 (Cambridge, Mass: Harvard University Press), 48–63.
- Davidson, D. (1963) 'Actions, Reasons, and Causes', *Journal of Philosophy* 60/23: 685–700.
- Davidson, D. [1970] (1980) 'Mental Events', in Ned Block (ed.), *Readings in the Philosophy of Psychology*, vol. 1 (Cambridge, Mass.: Harvard University Press), 107–19.
- Fodor, J. A. (1964) 'Explanations in Psychology', in Max Black (ed.), *Philosophy in America* (Ithaca: Cornell University Press).
- Fodor, J. A. (1968) *Psychological Explanation: An Introduction to the Philosophy of Psychology* (New York: Random House).
- Fodor, J. A. [1974] (1980) 'Special Sciences, or the Disunity of Science as a Working Hypothesis', in Ned Block (ed.), *Readings in the Philosophy of Psychology*, vol. 1 (Cambridge, Mass.: Harvard University Press), 120–33.
- Hempel, C. G. [1935] (1980) 'The Logical Analysis of Psychology', reprinted, with revisions, in Ned Block (ed.), *Readings in the Philosophy of Psychology*, vol. 1 (Cambridge, Mass.: Harvard University Press), 14–23.
- Husserl, E. [1900] (1977) *Logical Investigations*, trans. J. N. Findlay (London: Routledge & Kegan Paul).
- [1911] (1981) 'Philosophy as a Rigorous Science', in Q. Lauer, (ed.), *Phenomenology and the Crisis of Philosophy* (New York: Harper and Row).
- [1913] (1983) *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy, First Book*, trans. F. Kersten (The Hague: Martinus Nijhoff Publishers).
- [1937] (1970) *The Crisis of European Sciences and Transcendental Phenomenology*, trans. D. Carr (Evanston, Ill.: Northwestern University Press).
- Kim, J. [1972] (1980) 'Physicalism and the Multiple Realizability of Mental States', in Ned Block (ed.), *Readings in the Philosophy of Psychology*, vol. 1 (Cambridge, Mass.: Harvard University Press), 234–6.
- Lewis, D. K. (1966) 'An Argument for the Identity Theory', *Journal of Philosophy* 63/1: 17–25.
- [1969] (1980) 'Review of Putnam' in Ned Block (ed.), *Readings in the Philosophy of Psychology*, vol. 1 (Cambridge, Mass.: Harvard University Press), 232–3.

- Lewis, D. K. (1972) 'Psychophysical and Theoretical Identifications', *Australasian Journal of Philosophy* 50: 249–58.
- Nagel, T. [1974] (1997) 'What is it like to be a Bat?' in N. Block, O. Flanagan, and G. Guzeldere (eds.), *The Nature of Consciousness: Philosophical Debates* (Cambridge, Mass.: MIT Press).
- Oppenheim, P. and Putnam, H. (1958) 'Unity of Science as a Working Hypothesis', in Herbert Feigl, Michael Scriven, and Grover Maxwell (eds.), *Minnesota Studies in the Philosophy of Science, Vol. 2* (Minneapolis: University of Minnesota Press).
- Putnam, Hilary (1957) 'Psychological Concepts, Explication, and Ordinary Language', *Journal of Philosophy* 54: 94–9.
- [1960] (1975) 'Minds and Machines', in Putnam (1975: 362–85).
- [1963] (1975) 'Brains and Behavior', in Putnam (1975: 325–41).
- [1964] (1975) 'Robots: Machines or Artificially Created Life?', in Putnam (1975: 386–407).
- [1967a] (1975) 'The Mental Life of Some Machines', in Putnam (1975: 408–28).
- [1967b] (1975) 'The Nature of Mental States', in Putnam (1975: 429–40).
- [1969] (1975) 'Logical Positivism and the Philosophy of Mind', in Putnam (1975: 441–51).
- [1973] (1975) 'Philosophy and our Mental Life', in Putnam (1975: 291–303).
- (1975) *Mind, Language, and Reality: Philosophical Papers, Vol. 2* (Cambridge: Cambridge University Press).
- Ryle, G. (1949) *The Concept of Mind* (Chicago: University of Chicago Press).
- Ryle, G. [1962] (1971) 'Phenomenology versus *The Concept of Mind*', in *Collected Papers, volume 1: Critical Essays* (Bristol: Thoemmes).
- Searle, J. (1992) *The Rediscovery of the Mind* (Cambridge, Mass.: MIT Press).
- Smart, J. J. C. (1959) 'Sensations and Brain Processes', *Philosophical Review* 68: 141–56.