Schlick, Carnap and Feigl on the Mind-Body Problem

Sean Crawford

“If this be metaphysics, make the least of it!”

Moritz Schlick, Rudolf Carnap and Herbert Feigl are the most prominent of the positivists to formulate views on the mind-body problem. (Carl Gustav Hempel’s one-off treatment [1935] is set aside here; see Kim [2003] and Crawford [2013].) While their views differed from each other and changed over time they were all committed to some form of scientific physicalism, though a linguistic or conceptual rather than ontological form of it. In focus here are their views during the heyday of logical positivism and its immediate aftermath, though some initial scene-setting of Schlick’s and Carnap’s pre-positivist views will help to understand final opposing positions of Carnap and Feigl. All three philosophers are largely — entirely, in the case of Schlick and Feigl — concerned with sensations and sensory consciousness, or what Feigl came to call (after the psychologist E. C. Tolman) “raw feels” and what are now usually called “qualia” or “phenomenally conscious” mental states.

The pre-positivism of Schlick and Carnap

Schlick’s and Carnap’s pre-positivist views on the mind-body problem reflect a view dominant at the time in both psychology and philosophy: psychophysical parallelism — or, more accurately, psychophysical parallelism. In his General Theory of Knowledge (1918, 2nd ed. 1925) Schlick explicitly endorses it, though his aim is to interpret it properly. As already mentioned, the mentality in question, for Schlick, is qualitative or sensory consciousness — “mental qualities” — which for him are “immediately experienced reality,” the “directly given,” the “content of consciousness” (289).
In order to avoid a variety of intractable complications involved in a metaphysically dualistic parallelism (299), especially ones involving causal dependency, Schlick proposed what is often called an “identity theory” but is better described as a kind of “double language” or “double access” monism (Feigl [1975: 14] suggested the term “twofold access”). This theory is part of Schlick’s overall “critical realist” philosophy and especially his structuralist view of science, both of which rely on two crucial Kantian distinctions (Feigl [1975: 25] spoke of it as “Schlick-Russell-Eddington structuralism”): first, that between form and content, and second, that between cognition or conceptual knowledge (Erkennen) and intuitive experience or acquaintance (Kennen, Erleben). The idea is that science provides knowledge only of the structure of reality, which for Schlick takes the form of abstract uninterpreted axiomatic systems, which include scientific quantitative terms, laws and equations, and in which the concepts required for genuine knowledge are defined by Hilbertian implicit definitions that display their formal logico-mathematical relations to each other. The content of this relational structure consists of the intrinsically qualitative natures of the things related by the structure. These qualities are the “things-in-themselves,” the “transcendent” objects of reality, which, contrary to Kant, are knowable indirectly through designation by the (implicitly defined) quantitative scientific concepts. But we can also directly experience a tiny subset of this intrinsically qualitative transcendent reality, namely, the subset containing the qualities that compose parts our individual brains, which makes them accessible to us via introspection.

Metaphysically speaking, then, the intrinsic nature of the universe is ultimately purely qualitative, a web of mostly unexperinceable intrinsic qualities (283-292). We acquire knowledge of the “essence” of any one of these intrinsic qualities by “incorporating [it] into the quantitative conceptual system and thus reducing it to the fundamental intensities selected as a basis” (285). This is done through the “method of coincidences” (§31), part of which involves the coordination of the intrinsic qualities with quantitative physical concepts. A small minority of these intrinsic qualities — the subjective or mental ones — are thus doubly designated: directly by intuitive, often imagistic, psychological concepts (or representations), on the one hand, which are qualitative (e.g., yellow), and
indirectly by unintuitive, physical concepts, which are quantitative (e.g., wavelength of 590-560 angstroms), on the other hand. “The same reality,” he says, “— namely, that which is immediately experienced [i.e., some of the all-pervading qualities that constitute reality] — can be designated both by psychological concepts and by physical ones” (310). Only the latter, however, provide genuine knowledge of reality; the former give us mere acquaintance with some of it.

Even though for him mental/subjective and extra-mental/objective qualities “do not differ fundamentally” (328), Schlick is adamant that his monism is not a form of panpsychism, as the qualities that constitute the world are not all mental (327-29). Only the ones that constitute (some of) our brain, and with which we are directly acquainted by immediate experience, are mental, so that mind is “but a sector of the totality of the natural” (296). But, since “reality is called ‘physical’ in so far as it is designated by means of the spatio-temporal quantitative conceptual system of natural science” (294), and all of reality is so designatable, Schlick’s structuralist qualitative monism is a form of physicalism (somewhat akin to what Davidson [1970] called “nomological monism” but also with similarities to Jackson’s [1998] “Kantian physicalism”). With the formation of the Vienna Circle, and especially under the influence of Wittgenstein and Carnap, Schlick’s double-access quality monism is left behind as his thinking takes an anti-metaphysical turn.

Turning to Carnp, we find that his pre-Vienna Circle Aufbau (1928a; written during 1922-5) embraces both neutral monism (§§162, 183), derived from Russell (1921), and parallelism (§§166-9, 183). No contradiction obtains because he distinguishes between the “problem of mind-body dualism” and the “psychophysical problem.” The former is the question whether there are two separate types of things, be they substances, principles or aspects. Carnap’s answer is No. The conceptual system developed in the Aufbau only recognises one type of object domain and elements thereof are mental when categorized in a certain way and physical when categorized in another way: “the physical and the psychological must not be envisaged as two principles or aspects of the world. They are order forms of the one, unified domain of elements which are propertyless and merely connected through relations” (§162). The psychophysical problem, on the other hand, is the question
of how to interpret the parallelism of psychological and physiological events — or more accurately, how to interpret the fact that for every mental event there is a corresponding physiological event — something which Carnap assumes to have been empirically established. It falls to the science of psychology to establish exactly which types of brain events are correlated with which types of mental events. These are mere empirical correlation problems solved or solvable by science. The philosophical psychophysical problem, which Carnap calls the “essence problem”, is how to interpret the relation between psychological events and physiological events thus correlated. For Carnap, the parallelism can have no scientific explanation: “The quest for an explanation of these findings lies outside the range of science” (§169) and so “the quest for an explanation of that parallelism belongs within metaphysics.” Importantly, Carnap does not say at this point that this means the question is meaningless, only that it does not belong to science (§182). Rather, since the constructional system of the Aufbau is concerned exclusively with the reduction or definition of scientific concepts, such metaphysical problems simply fall outside its purview. (Carnap later stated that “regarding the criticism of traditional metaphysics, in the Aufbau I merely refrained from taking sides” [1963: 18-19; cf. 44-46] and that “in the Aufbau … such [metaphysical] theses are merely excluded from the domain of science” Carnap [1961: xi]).

The positivism of Carnap, Schlick and Feigl
Very shortly thereafter, however, following his participation in the Vienna Circle discussions of Wittgenstein’s Tractatus and its position that metaphysical statements are meaningless (unsinnig), Carnap explicitly commits himself to this view. In Pseudoproblems of Philosophy (1928b; written in 1927-8 after his first year in Vienna) he proposes a criterion of empirial meaningfulness or “factual content,” according to which “Only statements with factual content are theoretically meaningful; (ostensible) statements which cannot, in principle, be supported by experience are meaningless” (Summary, II). While not explicitly declaring the psychophysical (or mind-body) problem to be a meaningless pseudo-problem, that was his new view.
Things get much more familiar to contemporary philosophers of mind with what is standardly but misleadingly called Carnap’s “logical behaviourism,” a term he never used. He called his position “physicalism”, which was was an ambitious program to reduce or translate all scientific theories into a universal physical language that was supposed to serve as an intersubjective confirmation base. On this view, metaphysical statements that cannot be so translated are unconfirmable and hence cognitively meaningless and the traditional metaphysical mind-body problem is deemed a pseudo-problem. Carnap’s “logical behaviourism” (and that of the logical positivists generally) is simply part of this overall project of physicalizing all of empirical science, that is, translating the sentences of all the sciences into the physicalist language (either the language of physics or, later, the physical “thing language”), the only known “universal” language (the only one known to be intersubjective and intersensory). The so-called “logical behaviourism” of Carnap (and Hempel [1935]) is really just physicalism applied to the science of psychology. It is in this light that one must view Carnap’s later general thesis that “all statements of science can be translated into physical language” and the relevant sub-thesis that “all psychological statements can be translated into physical language.” This sub-thesis is no different from the sub-thesis for biology, that “every statement of biology can be translated into physical language” (1932/1934: 70).

The crucial point to appreciate, however, which has been largely misunderstood by contemporary philosophers — especially contemporary philosophers of mind — is that when it comes to the empirical sciences, such as psychology, the translations in question are not analytic (and therefore not knowable *a priori*). Moreover, they are not even restricted to physical-*behavioural* translations but can also be *physiological* translations (unsurprisingly, given the influence of parallelism in the *Aufbau*). From the early 1930s onwards, Carnap proposed “rules of translation” (or transformation) of the physical language in which the translations are to be carried out. It is clear that not all of these rules are laws of logic and that some of them are intended to be laws of nature (Carnap 1935b). In *The Logical Syntax of Language* (1934) and ‘Testability and Meaning’ (1936-7) Carnap explicitly distinguishes between the L-rules and the P-rules of a
scientific language on the basis of which transformations may be validly carried out: the former are logical laws and the latter physical laws. Both kinds of “translation rules” may be employed in physicalistic analysis or reduction. (Already earlier, he speaks of “the rules of transformation inside the physical language (including the system of natural laws)” (Carnap 1932/1934: 88, emphasis added; cf. 92 and 1933: 171).

In *Philosophy and Logical Syntax*, Carnap claims that “every sentence of any branch of scientific language is equipollent [equivalence] to some sentence of the physical language, and can therefore be translated into the physical language without changing its content” (1935/1963: 455). Carnap is very clear there and in *The Logical Syntax of Language* (1934/2002: §§51 and 82) that there can be two concepts of equivalence in a physical language: logical equivalence (L-equivalence) and physical equivalence (P-equivalence). Two sentences are L-equivalent when they are mutually derivable solely on the basis of logical laws; two sentences are P-equivalent when they are mutually derivable, in addition, on the basis of physical laws. Carnap explicitly allowed a psychological sentence, Q₁, and a physical translation of it, Q₂, to be P-equivalent, as Q₁ could be transformed into Q₂ on the basis of “a scientific law, that is, a universal sentence belonging to the valid sentences of the scientific language-system” (1935a/1963: 456; cf. 1933, 1935b). He made absolutely clear that this universal sentence “need not be analytic; the only assumption is that it is valid. It may be synthetic, in which case it is P-valid” (*ibid.*). Contrary to the received view fostered by Putnam (1965), Carnap never claimed to offer analytically true logical constructions of “mind talk” into either (overt or covert) “behaviour talk” or “physical talk” (see Crawford 2013, 2014).

On the basis of his close association and conversations with Carnap and Wittgenstein, inside and outside the Vienna Circle, Schlick came to reject much of the double-access structuralist qualitative monism he earlier espoused in the *General Theory of Knowledge* and embraced a form of the verifiability theory of meaning and adhered to the new translation physicalism. Although there are strong traces of his previous view lurking in his 1935 positivist paper “On the Relation
between Psychological and Physical Concepts”, he tries to strip these features of any metaphysical trappings, declaring his new adherence: “every psychological proposition can be translated into an expression in which physical concepts alone occur.” With the term “physicalism” defined accordingly, he declared that “I therefore hold the thesis of physicalism to be correct” (1935: 399). 

Contrasting this with his earlier superficially similar position in General Theory of Knowledge, he remarks that “If, as a matter of fact, the physical language is characterised by complete universality, the setting down of this circumstance is in no way the assertion of a metaphysical monism” (407).

Like his teacher Schlick’s 1935 paper, but even more so, Feigl’s first paper on the mind-body problem, ‘Logical Analysis of the Psychophysical Problem: A Contribution to the New Positivism’ (1934) presents a rather confusing hybrid of Schlick’s old qualitative monism and the new positivist translation physicalism. But the main message, at least according to the later Feigl, was his version of translation physicalism: “the only non-metaphorical non-metaphysical formulation of the double aspect or double knowledge theory [is]: the formulation and detailed analysis of the mutual translatability of two universal languages. This, then, is the new positivist view of the psycho-physical relation” (437). Moreover, “we deem … Materialism strictly meaningless if [it] pretends to express [a] factual hypothesis” (443).

Unfortunately, unlike the ever rigorous Carnap, neither Feigl nor Schlick were at all clear about what they meant by “translation.” Schlick simply does not say. But his silence is less confusing than what Feigl does say. For Feigl talks about the “logical” nature of the psychophysical problem and its “logical” solution. At one point he says that “Logical mutual translatability, isomorphism, means simple identity of the two propositions” (436). He does not say what he means here by “logical” or by the “identity” of the “two” propositions. The natural interpretation would be that the “identity” of the two propositions arrived at by “logical mutual translatability” means that “they” are at least synonymous in the strong sense of analytically equivalent (and hence knowable a priori) — to use Carnap’s terms, L-equivalent. But this is of course completely implausible and
indeed absurd, as it would amount to an armchair natural science of psycho-physiology. On the other hand, “isomorphism” suggests, much more plausibly, a merely contingent or synthetic correlation between the two propositions.

**The aftermath: Feigl vs Carnap**

By 1950 Feig had clearly thought through the matter with greater care and now realized that there can be no “logical” (in the sense of analytic) connection between psychological statements and their physical translations but only a synthetic one. (Feigl must have actually realized this even before the publication, though presumably not the writing, of Feigl [1934]). For in a letter to Carnap in 1933 he suggested that the relation between psychological sentences and physical sentences will be synthetic, something Carnap in reply explicitly agreed with — unsurprisingly, as this was his position all along. See Feigl [1963: 255n28] and Crawford [2013].) Feigl now returned to what he seems to have always felt was the right view to take of the mind-body problem all along, but for the positivist peer pressure: that it was not a pseudo-problem (Feigl 1961) and his teacher Schlick had basically got it right with his structuralist qualitative monism earlier in *General Theory of Knowledge*.

In “The Mind-Body Problem in the Development of Logical Empiricism” (1950) Feigl declares that “The precipitous assertion of logical equivalence [which he rejects] was of course based on the phenomenalistic claims of the explicit definability of the entities in one realm in terms of the entities of the corresponding realm” (620). But who exactly made this “precipitous assertion of logical equivalence”? In his later classic long essay ‘The “Mental and the “Physical”’ (1958) he repeats the point more clearly and explicitly:

A most important *logical* requirement for the analysis of the mind-body problem is the recognition of the *synthetic or empirical* character of the statements regarding the correlation
of psychological to neuro-physiological states. … I was tempted to identify, in the sense of logical identity, the mental with the neurophysiological …

But if this theory is understood as holding a logical translatability (analytic transformability) of statements in the one language into statements in the other, this will certainly not do. …

[T]he question which mental states correspond to which cerebral states is in some sense … an empirical question. If this were not so, the intriguing and very unfinished science of psychophysiology could be pursued and completed by purely a priori reasoning. …

… Subjective experience … cannot be logically identical with states of the organism; i.e., phenomenal terms could not explicitly be defined on the basis of physical\(_1\) or physical\(_2\) terms. (1958: 389-90; cf. 391).

According to Feigl, then, he adhered to an analytic translatability thesis back in 1934 (or 1933) and so he himself made the “precipitous assertion of logical equivalence”. Did anyone else? Feigl gives the impression that this was the general consensus during the early phase of logical positivism. But as we have seen, Carnap certainly did not hold any such view at any time. Whether anyone other than Feigl held this misbegotten view for more than a minute is not completely clear (on Hempel [1935] see Crawford [2014: 717n8]).

At any rate, Feigl and others influenced by him seem to have confused analytic translation with explicit definability. But for Carnap explicit definitions need not be analytic (the constructional definitions of the \textit{Aufbau}, e.g., are extensional [Carnap 1928a: §§35, 43, 47-49; cf. 1961]). The point is that if a non-primitive expression, the \textit{definiendum}, is explicitly definable in terms of primitive expressions, then it can be eliminated and replaced by its \textit{definiens}, by the primitive expressions. Such explicit definitions may be extensional, offering merely materially necessary and sufficient conditions for the \textit{definiendum}, that is, the construction of a material bi-conditional whose right-hand side, the \textit{definiens}, contains only undefined primitive terms. The defined
expressions will be so-called “theoretical” terms (e.g., mental ones) and the primitive expressions
the “observation” terms (e.g., behavioural and physiological ones). To his credit, Carnap (1936-
37) very early on saw that the search for explicit definitions of all empirical scientific terms in the
extensional physical-thing language was badly misconceived, especially since many such
definitions were supposed to be empirical laws. Instead of invoking a stronger intensional
language, he weakened the project to that of providing (extensional) “reduction sentences” that
linked the empirical term in question to observable physical conditions only under certain test
circumstances. Since these physical reduction sentences were not definitions of the terms they
were reducing — they were only partial “conditional definitions” — they did not allow the terms
to be eliminated and replaced and hence they could not form the basis for translations. The crucial
point to notice about this move from definition to reduction is that, with respect to the
physicalization of psychology and other empirical sciences, it is not a shift from the category of
analytic truths (knowable only a priori) to the category of synthetic truths (knowable only a
posteriori). Rather, it is a shift within the single category of synthetic truths knowable only a
posterior from explicit (complete) definability, which permits elimination of the defined term, to
conditional (incomplete) definability, which does not permit elimination of the partially defined
term.

In “The Mental and the Physical”, operating under the misapprehension (seemingly shared by
Putnam [1957]) that he has discovered that the relation between mental statements and physical
statements is not analytic as the early logical positivists allegedly thought, but synthetic, Feigl
returns to what is in all essentials a version of Schlick’s pre-positivist position (updated by
replacing Schlick’s original distinction between conceptual knowledge and intuitive acquaintance
with Russell’s distinction between knowledge by description and knowledge by acquaintance):

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 (450).

Although he describes his own theory as an identity theory, he also, like Schlick, describes it as a kind monism, which seems much more appropriate given that they both view the referents of both mental and physiological terms — that is, the “realities in themselves” — as intrinsic qualities. Feigl, for instance, says that “The ‘mental’ states or events (in the sense of raw feels) are the referents (denotata) of both the phenomenal terms of the language of introspection, as well as of certain terms of the neurophysiological language” (1958: 447). And in a later paper, he writes that “I take these referents [of mental and neurophysiological terms] to be the immediately experienced qualities” (1960: 38; cf. 1958: 457, 474; 1963: 257, 262.). Strikingly, for Feigl, as for Schlick, “sentience (qualities experienced, and in human beings knowable by acquaintance) and other qualities (unexperienced and knowable only by description) [are] the basic reality” (1967: 107). The concepts of theoretical physics “denote realities which are unknown by acquaintance, but which may in some way nevertheless be not entirely discontinuous with the qualities of direct experience” (ibid.: 40; cf. 1971: 308).

The issue of whether this is a form of panpsychism arises, just as it did for Schlick. Feigl responds pretty much like Schlick: the world “as it is in itself” is a field of connected qualities some of which are mental and others which are not. But all these qualities can be known indirectly by description through physical science, hence we again have a kind of structuralist physicalism: everything that exists is structurally describable physically and unfolds according to physical law. A small subset of the universe of qualities that is the content of this physical structure, and that the laws govern, is also known by acquaintance and therefore mental, which is to say that it, unlike
the other qualities, can be referred to or described in psychological terms. While Feigl, like Schlick, explicitly rejects panpsychism, he is prepared to apply the term “pan-quality-ism” to his view that intrinsic qualities constitute all of nature — not just parts of our brains and nervous systems (1961: 39). (Compare Feigl’s own exploration of the similarities between Russell [1927] — who also endorsed “pan-quality-ism” — and Schlick.) Unfortunately, neither Feigl nor Schlick satisfactorily explained exactly what the difference between the mental and the non-mental qualities that constitute reality is, and so neither really managed successfully to shake off what seemed like the panpsychistic implications of their view. Feigl himself says that “one is tempted, with the panpsychists, to assume some unknown-by-acquaintance qualities quite cognate with those actually experienced” (1958: 108). But no clear account is ever given of what this “cognate” nature could be that falls short of being mental — calling forth Feigl’s (1961) witty remark which forms my epigraph: “If this be metaphysics, make the least of it!”

It appears, then, that Schlick’s and Feigl’s (1967: 142; 1975) structuralism about natural science combined with what Hatfield (2004) has called “respect for the phenomenal” (“raw feels”) forms the basis for their physicalist-monistic solution to the mind-body problem. Natural science, especially the foundational and comprehensive science of physics, is limited to providing indirect, quantitative and purely structural knowledge by description of the entire world in physical language. This structure has a content, however, that is purely qualitative, part of which can be accessed or known directly by acquaintance. Feigl, I think, is right, in holding that the name “double-language theory” (or “double-knowledge theory”) is a better term than “identity theory” for this kind of monism, for neither he nor Schlick are really identifying what, pre-theoretically, seem like two different kinds of properties, qualitative raw feels and quantitively measurable physiological processes. Rather, when it comes to the mental part of intrinsically qualitative existence, the raw feels are the true existents — the things or realities “in themselves” — and they are referred to in two different ways.

Now what was Carnap’s reaction to all this? Well, this is what he says in response to Feigl:
The identity statement mentioned [that a certain psychological process $P$ is identical with a certain neurophysiological process $N$] is a sentence of the object language; this fact may mislead the reader into believing that the controversy about the identity view concerns a question of fact. … It seems preferable to me to formulate the question in the metalanguage, not as a factual question about the world, but as a question concerning the choice of language. … Those facts Feigl proposes as evidence for the identity view are perhaps better regarded as reasons for preferring a monistic language … in this language the predicates “$P$” and “$N$”, though not L-equivalent, are P-equivalent … I am willing to call my position an identity conception in the following sense: in agreement with Feigl I prefer the monistic language, and like him I believe that the evidence available today provides good reasons for the assumption that this language will also function well in the future (Carnap 1963: 885-6).

Unlike Feigl, Carnap continued to maintain that the mind-body problem, like all traditional metaphysical problems, was a pseudo-problem. In 1935 he had said that the “pseudo-object” identity-sentence of the material mode, ‘The evening-star and the morning-star are identical’ is to be replaced by the syntactical formal-mode sentence ‘The words “evening-star” and “morning-star” are synonymous’ (1935a/63: 447); given the earlier more rigorous treatment (1934/2002: §75), it is clear that he means P-synonymous. But even in his later semantic period, I think he would grant that his position on the mind-body problem could be expressed — albeit highly misleadingly — in the material mode as a synthetic identity theory. But that is as close to the Schlick-Feigl “identity theory” as he is willing to get.

Carnap makes no remarks about Feigl’s Schlick-inspired structuralist pan-quality monism. Obviously, he would flatly reject it as a metaphysical pseudo-doctrine. It is interesting to note, however, that around the time of his reply to Feigl, Carnap too, like Schlick and Feigl, explored a kind of structuralism. (See Carnap [1966] which makes essential use of the Ramsey sentence of a
scientific theory — a key ingredient in some forms of structuralism — in his explication of the division between its analytic and synthetic components.) The difference is that Carnap’s structuralism — unlike Schlick’s and Feigl’s — was, unsurprisingly, anti-metaphysical, eschewing questions about the intrinsic ontological nature of the content of the structure as pseudo-questions. That said, despite their various changing and divergent attitudes towards metaphysics, in particular whether the mind-body problem is a pseudo-problem, all three of our scientific philosophers never wavered from some form of linguistic-conceptual physicalism.

References


Feigl, Herbert. 1958. ‘The “Mental” and the “Physical”’, Minnesota Studies in the Philosophy of Science II.


