Nonreductive Materialism and the Nature of Intertheoretical Constraint

ROBERT VAN GULICK

I

During the past two decades philosophers thinking about the mind/body problem have been attracted to one or another version of nonreductive materialism. In accord with our contemporary scientific world view, they have generally accepted the physicalist claim that everything in space or time is physical. However, they have also recognized the failure to date of all attempts to reduce the mental to phenomena specifiable in some physical and nonmental vocabulary. Despite initial optimism, we have not been able to reduce mental states to behavioral dispositions, nor to type-identify mental properties with properties specified in the vocabulary of neuroscience. We have not been able to reduce or identify mental properties such as being in pain or believing that global warming has begun with behavioral dispositions nor with specific neurophysiological properties.

Given the failure to find a reductionistic home for the mental within the realm of the physical, there may seem to be only three options open to the philosopher.

1. She can reject physicalism and become a dualist recognizing the existence of a nonphysical realm (Jackson 1982).

2. She can become an eliminative materialist by denying the reality of the mental, while making some suggestions about physically acceptable replacement theories to take over the explanatory work which we had hoped would be done by our mentalistic theories (Churchland 1986).

3. She can continue the search for a successful way to reduce the mental to the physical (Kim 1989).

The first two options accept the nonreducibility of the mental but draw opposite conclusions. The dualist argues that since the mental surely exists and there is no place for it within the realm of the physical there must be a nonphysical realm. In contrast the eliminativist argues that since the physical is all there is and there is no place for the mental within the physical, the mental does not exist. The only alternative left would seem to be our third option of continuing the reductionist search for a way to find a place for the mental within the physical realm.

However, nonreductionist materialism offers an attractive fourth option; it simply denies that finding a place for the mental within the realm of the physical requires finding a way to reduce mental phenomena to physical phenomena. It offers us the metaphysical result we want without the hard work of showing how to reduce the mental to the physical. It combines a claim of ontological materialism (namely the claim that everything that exists in space or time is physical) with a commitment to theoretical dualism (or pluralism) and the autonomy of mental or psychological theory. Reductionism is criticized as a restrictive and overly demanding view of what we need to make the mental acceptable as part of the physical world. The nonreductionist argues that we must not confuse the plausible claim of ontological physicalism with the implausible claim that the physical sciences provide us with conceptual and representational resources adequate for describing and explaining everything within the physical world. As Jerry Fodor has noted, it is one thing to claim that every economic or monetary transaction is (ontologically speaking) just a complex physical event, but quite another to claim that physical science can provide the theoretical machinery needed to explain such monetary transactions (1975).

Reductionistic physicalism is viewed as a form of intellectual imperialism, which implausibly denies the legitimacy of any theoretical framework that falls outside the conceptual and representational resources of the physical sciences.

Unsurprisingly nonreductionistic materialism has won wide acceptance among philosophers (e.g., Fodor 1981, Davidson 1970, and Boyd 1980); indeed in America it has probably become the majority view, though its critics may attribute this to what Bertrand Russell called the "virtue of theft over honest toil", i.e. to its seductive promise of scientifically respectable materialist status despite its refusal to pay its materialist bills in reductionist coin. Nonreductive materialism also benefits from the intellectual spirit of our era; at a time when doctrinaire ideologies have been generally discredited, nonreductive materialism appeals to our spirit of toleration: "Let a thousand flowers bloom, as long as the all have good materialist roots." Nonreductive materialism offers to satisfy our desire for intellectual pluralism without forcing us to embrace metaphysical dualism or sacrifice our materialist credentials.
But alas as is usually the case, all is not well in paradise. Critics have argued that nonreductive materialism cannot fulfill its promises. On one hand, anti-materialists have claimed that their arguments are equally effective against nonreductive materialism. And on the other hand, reductionists have argued that nonreductive materialism is a basically unstable or perhaps even an incoherent position. Consider three types of objections.

1. Anti-materialist arguments based on the special status of our knowledge about the phenomenal or qualitative properties of our own conscious experiences. All such arguments use premises concerning our knowledge about the phenomenal properties of our own conscious states to support conclusions denying ontological physicalism.

A. Frank Jackson (1982), appeals to a hypothetical scientist named Mary stipulated to be the world authority on color perception who knows everything physical there is to know about seeing red. However Mary has spent her entire life in an achromatic environment and thus never herself experienced red. Jackson claims that when Mary is at last released from her isolation and first experiences red she will come to know something about seeing red that she did not know before. But since she already knew everything physical there was to know about seeing red, what she learns must be something nonphysical. Thus he concludes that not all information is physical information, and physicalism is false.

B. A distinct though related argument was earlier given by Thomas Nagel (1974). Nagel appeals to the fact that our ability to understand what it’s like to have a certain type of experience is restricted by the range of experiences that we ourselves are capable of experiencing or imagining. There are aspects of what it’s like to see red that the congenitally color blind person can never understand. Similarly there are facts about what it’s like to perceive the world through echo-location that are knowable by bats but can never be fully understood by us humans who are incapable of having such experiences ourselves. Facts about what it’s like to undergo a given type of phenomenal experience are subjective in the sense that such facts can only be understood from certain types of points of view, i.e. from the points of view of creatures capable of having or imagining such experiences. But, Nagel argues, physical facts are objective facts understandable from many points of view; indeed the objectivity of physical description consists in representing the world in a way which abstracts away from the perspective of any given sort of knower or perceiver of the world. Thus Nagel concludes facts about what it’s like to be a certain sort of conscious creature (or to have a conscious experience with a given sort of phenomenal quality) are not physical facts.

Thus there are nonphysical facts and physicalism is false.

[1've skipped the second objection here.]

3. Reductionist Arguments intended to show that nonreductive materialism cannot provide an adequate explanation of how mental properties are related to or depend upon physical properties. Nonreductionists wish to combine their materialism with a rather loose view of the relation between the mental and physical levels. In particular, they deny that there are any lawlike biconditionals between the two levels or any links allowing one to deduce an organism’s mental properties from its physical description. Yet to be good materialists nonreductionists want to retain the idea that mental properties in some way depend on physical properties. But their critics argue that there is no way to make sense of this dependence once the existence of lawlike biconditionals and deductive links has been denied (Kim 1989). If we cannot deduce an organism’s mental properties from its physical properties, it seems we cannot successfully explain how the former depend on the latter (Levine 1983). It seems we must be left with an explanatory gap and a theory that borders on mystification. The nonreductionist tells us that the mental depends on the physical but denies that he can provide us with any linked theory that would allow us to pass deductively from the physical base to the mental properties that depend upon that base. Once again the reductionist argues that the problem can be solved only by becoming a reductionist; one can close the explanatory gap only by accepting the tight reductive interlevel links the nonreductionist rejects.

Together these three arguments present a serious challenge to nonreductive materialism. Nonetheless I believe the challenge can be met, and in the rest of this paper I will present a particular version of nonreductive materialism and show how it can deal with each of these three criticisms.
Let us now see how our pluralistic teleo-functionalist version of nonreductive materialism can deal with the challenges raised at the outset.

Frank Jackson and Thomas Nagel both appeal to the special status of our knowledge of the phenomenal and qualitative properties of our own conscious states in support of an anti-physicalist conclusion. According to Jackson, Mary gains new factual information when she first experiences red even though she already knew everything physical there was to know about seeing red. Thus he concludes that what she learns is nonphysical and physicalism is false. According to Nagel, there are facts about what it is like to be a bat that we humans can never understand even if we are able to understand everything physical there is to understand about what goes on in the bat’s brain when it echo-locates. Thus he concludes that there are nonphysical facts and physicalism is false.

Jackson and Nagel view their rejections of physicalism as denials of ontological physicalism, i.e., as denials of OP1–OP3. But the epistemic facts to which they appeal support no more than a denial of reductionism. We can accommodate both Mary’s gain of knowledge and our inability to understand what it’s like to be a bat by embracing the nonreductionist view that there are a plurality of theoretical frameworks many of which cannot be captured within the resources of physical theory but that nonetheless accurately represent aspects of the world that obtain in virtue of underlying physical facts. The representational system associated with reflexive phenomenal self-consciousness is one such nonreducible but valid framework. As we noted above, this framework mediates very special intra-mental causal connections; the causal avenues for self-modification it provides the mind with are not accessible through the use of the theoretical representations of physical science. Self-conscious minds are self-monitoring and dynamically self-modifying systems. The first person representations associated with self-consciousness acquire their content at least in part on the basis of the causal role they play in direct intra-mental causal relations especially those that constitute the mind’s organizing control of itself as an evolving dynamical system.

Given these facts, we should not expect to be able to establish exact correspondences between the concepts of our first-person self-conscious representational framework and the concepts of the physical sciences. They are too different with respect to how the representation-user causally interacts with the representations, too different with respect to the goals which the framework can be used to achieve, and too different with respect to the role played by indexical and self-referential factors in determining content. Some measure of incommensurability is what we should expect.

If we accept a fine-grained individuation of propositions, according to which propositions are distinguished in part on the basis of internal conceptual structure (e.g., such that the proposition that \[5 + 7 = 12\] is not the same as the proposition that \[8 + 9 = 17\] even though they are both necessarily true in every possible world) then the nonreductive physicalist can concede that Mary comes to know a new proposition when she first experiences red. The concept of phenomenal red that she is able to apply from the first person perspective need not be capturable within the resources of physical theory; it need have no exact physical theory equivalent. Only from the causal and pragmatic perspective internal to the framework of self-consciousness is Mary able to capture the concept of experiencing phenomenal red; she is able to delimit the set of underlying physical properties that satisfy the relevant functional role only from within that framework.

Yet the phenomenal property about which Mary gains knowledge can still be an ontologically physical property in so far as every instantiation of that property is dependent upon underlying physical-theory properties that realize it.

It is worth noting that unpacking the mental/physical relation in terms of instantiation provides for a much more robustly physicalistic notion of dependence than the sort of supervenience that Jackson is willing to allow. He concedes that it may be impossible to have a mental difference without a physical difference. But supervenience so construed requires nothing more than systematic correlation or nomic covariance between mental and physical properties which may remain quite separate and distinct. On the functionalist notion of instantiation, by contrast, mental properties are not merely correlated with or caused by physical properties; they are realized by them. Every mental property instance has an underlying physical realization.

We can reply to Nagel’s worries about understanding what it’s like to be a bat much as we did to Jackson. There is no representation of physical theory that we can use to achieve a causal interactive relation with the world that duplicates the relation that the bat bears to his own experiences. That sort of understanding is available only from within the self-interactive system of the bat’s conscious mind. Nonetheless, the particular fact that the bat understands is an ontologically physical fact in so far as it obtains
in virtue of underlying physical facts about the bar's brain, namely those facts that underlie the bar's brain having the sort of causal structure that allows the bar to understand itself and interact with itself as it does (Van Gulick 1985). A commitment to the principle of compositionality is again all that is required. To be good physicalists we need not suppose that the representations of physical science provide us with a universal language we can use to comprehend every property realized within the physical world, nor need we suppose that physical theory is a universal tool affording every possible avenue of causal access to that world.

Thus contrary to what Jackson and Nagel may believe, their arguments do not refute ontological physicalism. They refute only reductionist claims about the universality of physical theory as a system of representation and understanding, claims which are of course rejected by nonreductionists.

Our third and final challenge involved the claim that nonreductive materialism is unable to provide an explanatorily adequate account of the relation of the mental and the physical and of how the mental depends on the physical. The reductionist argues that unless we can identify mental properties with physical properties or at least provide a means of deducing an organism's mental properties from its physical properties, we are left with an unacceptable explanatory gap in our allegedly materialist view of mind which fails to explain how the mental depends on the physical.

In order to maintain his status as a materialist, the nonreductionist must hold that mental properties depend on physical properties — that in every instance in which a mental property applies to some part of the world it does so in virtue of physical properties that apply to the world. For the functionalist the in virtue of relation is to be unpacked by the idea of realization; any given mental property M is realized by underlying physical properties satisfying the functional relation R_M associated with M. According to the functionalist, the relation between physical properties and mental properties is not one of identity but rather one of realization or instantiation. Thus we will not get any true identity statements of the form (Physical property P = Mental property M)

First the reductionist's charge overstates the extent of what has not been explained. Deducibility is not a sine qua non for explanation. The nonreductionist functionalist admits that there is no algorithm that will take you from a physical description of an organism to a correct description of its mental properties, but the functionalist has a great deal to say about what sorts of causal functional roles are relevant to having those mental properties and about what sorts of physical organizations could contribute in the relevant sorts of ways. Consider the tracking conditions we discussed above and the role they play in the teleo-functionalist's story about how representational content gets fixed. A good physicalist story about how certain central nervous system states of an organism systematically covary with changes in environmental conditions obviously has a lot of value in explaining how these physically described brain states help to realize perceptual or other contentful states of the organism. The nonreductionist recognizes the explanatory value of such connections, but has a modest and qualified view about how tightly the links between the mental and physical frameworks can be forged. Given the differences in the range of pragmatic features to which the two frameworks must be sensitive, the nonreductionist denies that deducibility is the appropriate way to think about the explanatory correspondences we should expect to establish between them. But to deny deducibility is not to leave the connection between the mental and the physical as a mystery or a totally unexplained brute fact. Thus to say the nonreductionist leaves a "huge explanatory gap at the center of his theory" is simply not true.

Secondly there is a sense in which the nonreductionist's theory is explanatorily complete. There are things it does not explain and gaps it leaves unfulfilled, but it provides meta-explanations of why those matters cannot be explained and of why those gaps cannot be filled.

For example, consider our limited ability to explain the phenomenal qualities of experience. In contrast with our functionalist accounts of intentional states like beliefs and desires, we have as yet no very illuminating theories about the functional role that a physical pattern of organization must satisfy to instantiate a phenomenal experience. Yet with time, I suspect we will make progress and arrive at some general account of the functional roles that distinguish conscious phenomenal processes from those lacking phenomenal properties. Yet I doubt that we will ever be able to produce a theory that fully explains why a state's having certain nonqualitative properties, be they physical science properties P_1, ..., P_n or functional properties F_1, ..., F_m, makes it a state with a specific qualitative property Q_i (e.g. why it is a visual experience of phenomenal red). Larry Hardin (1988) has argued that we may be able to explain certain aspects of the phenomenal quality, e.g. why it involves a unary hue (such as red) rather than a binary hue (such as orange). But it seems to me that there will always be some nonrelational aspects of the phenomenal quality that will be left unexplained.
But this is just what the teleo-functionalists would expect. As we saw above in answering the Jackson and Nagel objections, the concepts associated with our phenomenal first-person representational framework are to some extent incommensurable with those associated with our physical science framework. Our first-person phenomenal concepts provide us with a way of understanding ourselves and of entering into adaptive intra-mental causal relationships with our own minds that are not accessible to us through any representational structures of physical theory. No physical science representation can give us the knowledge of what it’s like to have a specific sort of phenomenal experience because we can not use any such representation to structure our intra-mental relations in the way we can through our use of phenomenal representations and concepts. Thus the subjectivity of facts about experience and their inaccessibility from the objective framework of physical science are predictable consequences of our teleo-functionalists version of nonreductive materialism is explanatorily complete. It leaves gaps in our ability to explain how specific mental properties derive from the physical properties of their realizations, but it predicts the existence of those very gaps and explains why they cannot be filled (Van Gulick 1985).

Thus I hope to have shown that our teleo-functionalists version of nonreductive materialism can successfully turn back each of the three challenges we raised against it. And those of us who want to reject reductionism but keep our materialist credentials can continue to “eat our cake and have it, too”.

References
