A. MARRAS

KIM ON REDUCTION

ABSTRACT. In *Mind in a Physical World* (1998), Jaegwon Kim has recently extended his ongoing critique of 'non-reductive materialist' positions in philosophy of mind by arguing that Nagel's model of reduction is the wrong paradigm in terms of which to contest the issue of psychophysical reduction, and that an altogether different model of scientific reduction – a *functional* model of reduction – is needed. In this paper I argue, first, that Kim's conception of the Nagelian model is substantially impoverished and potentially misleading; second, that his own functional model is problematic in several respects; and, third, that the basic idea underlying his functional model can well be accommodated within a properly reinterpreted Nagelian model. I conclude with some reflections on the issue of psychophysical reduction.

1. INTRODUCTION

A central aim of Jaegwon Kim's work on the mind-body problem over the past two decades has been to undermine various anti-reductionist positions in philosophy of mind and, in particular, to challenge the "anomalist" arguments of Donald Davidson and the functionalist, "multiple realizability" arguments of Putnam, Fodor, and others – arguments whose purpose was to establish some form of "non-reductive materialism" as a physicalist alternative to the type identity theory of J. J. C. Smart and Herbert Feigl.[1] But whereas Kim's earlier stance against the anti-reductionist positions shared with them the common conception of intertheoretic reduction usually associated with the traditional Nagelian model, in his more recent work Kim has actually disputed the very soundness or aptness of the Nagelian model and urged an alternative model of "functional reduction" which he believes to be more appropriate not only for mind-body reduction but also for scientific intertheoretic reduction generally.[2] The Nagelian model, which, as the story goes, requires the derivability of the laws of the target theory from the laws of base (or reducing) theory via "bridge laws", was the model that Davidson presupposed when he denied that intentional psychology, being "anomalous", could be reduced to physical theory – no psychological laws and no psychophysical "bridge" laws being available for the reduction. The Nagelian model was also the

Erkenntnis **57:** 231–257, 2002.

one that functionalists or "multiple realizationists" presupposed when they denied the availability or possibility of biconditional "bridge laws" linking psychological and physical predicates, and, a fortiori, the possibility of psychophysical type-identities. Kim's response to Davidson was, essentially, that mental anomalism makes mental properties "epiphenomenal", and thus can't make sense of intentional explanation as a form of causal explanation. And his response to the multiple realizationists was that a mere "token-physicalism" was too anemic as a theory of the mind-body relation unless it was supplemented by a thesis of "strong" supervenience; but such a thesis, Kim argued, re-opens the possibility of Nagelian reductionism, either by entailing nomological coextensions of psychological properties with disjunctions of their physical realizers, or – if such disjunctions are not permissible – by allowing for species- or structure-specific psychophysical bridge laws adequate to ground "restricted" or "local" reductions. In each case, the Nagelian model was a shared paradigm in terms of which the issue of psycho-physical reduction was being contested. It was, Kim now believes, the wrong paradigm. What is needed is an altogether different model: a *functional* model of reduction. Such a model, Kim holds, will prove appropriate not only for psychophysical reduction, but for intertheoretic, scientific reduction generally.

My aim in this paper is to evaluate these contentions. I shall argue, first, that Kim's conception of the Nagelian model as centered around the notion of bridge laws is substantially impoverished and potentially misleading; second, that his own functional model of reduction is problematic in several respects; and, third, that the basic insight of his functional model can well be accommodated within a properly reinterpreted Nagelian model. I will then consider some of the consequences of my discussion with regard to the issue of psychophysical reductionism.

2. THE NAGELIAN MODEL OF REDUCTION

No doubt the Nagelian model has been under suspicion for a long time.[3] Some of the early criticism of the model was, we may say, "constructive": it was directed against the highly idealized nature of the model, and aimed at bringing it in line with some of the complexities of actual reductions in the history of science, without, however, any intention to restrict its scope or deny the logical/derivational structure of the reductive project (Schaffner, 1967; Hooker, 1981). More radical forms of criticism came from those who, while not wishing to deny the applicability of the model to some of the natural sciences, felt that it is altogether inappropriate to at least some of the "special sciences" – biology, psychology, economics, etc.

– whose reducibility does not ostensibly involve the derivation of laws or the postulation of nomological coextensions across domains, but, at best, the specification of the functional and physical mechanisms in terms of which the workings of higher level, "special science" systems could be understood (Fodor, 1974; Cummins, 1983; Brook, 1994). Kim's recent critique of the Nagalian model is perhaps the most radical. The model is said to be deficient in a number of essential respects that render it incapable of satisfying the basic aims of intertheoretic reduction.

One reason for Kim's dissatisfaction with the Nagelian model seems to be that it presupposes the problematic Hempelian D-N model of scientific explanation:

> ...the Nagel model of reduction is in effect the Hempelian D-N model of scientific explanation applied to intertheoretic contexts. Just as Hempelian explanation consists in the derivation of the statement describing the phenomenon to be explained from laws together with auxiliary premises describing relevant initial conditions, Nagelian reduction is accomplished in the derivation of the target theory from the base theory taken in conjunction with bridge laws as auxiliary premises. It is therefore more than a little surprising that while the D-N model has had few committed adherents for over three decades, Nagel's derivational model of reduction is still serving as the dominant standard in discussions of reduction and reductionism (Kim, 1998, p. 26).[4]

Though Kim is no doubt correct in supposing that something is wrong with the D-N model of explanation, this surely can't be the requirement that, at least for a large class of explanations, the explanandum must be logically derivable from laws and suitable auxiliary premises: such derivation may not be sufficient for explanation, but it is surely often a necessary part of it. Kim himself must believe as much in view of his critique of anomalous monism, which calls into question the explanatory relevance of mental types, under anomalous monism, precisely because of the alleged absence of subsumptive mental laws (cf. Kim, 1989, 1993b). Whatever may be wrong with the Nagelian model, then, cannot be the deductive-nomological requirement it shares with the Hempelian model of explanation. (Nagel refers to this requirement as the "condition of derivability"; Nagel, 1961, p. 354). At least for certain types of explanations – reductive or otherwise – their deductive-nomological structure seems to be beyond question; indeed, those critics of the Nagelian model whom I labelled as "constructive", as well as many of those in the more "radical" camp, regard this deductive structure as central to reductive and to other kinds of explanation.[5]

What for Kim is really objectionable about the Nagelian model is not the derivability requirement as such, but the requirement that the derivations are mediated by "bridge laws" (Nagel refers to this requirement as the "condition of connectability"; Nagel, 1961, p. 354.) Thus he regards bridge laws as being "at the heart of Nagel's model", inasmuch as they "provide the essential reductive links between the vocabulary of the theory targetted for reduction and that of the base theory, and thereby enable the derivation of the target theory from its reducer" (p. 90). And although, as Kim reminds us, Nagel did not require that bridge laws have the form of biconditionals, "it has been customary" to assume bridge laws to take this form, so that "the bridge law requirement came to be understood as saying that each property in the domain to be reduced must be provided with a coextensive property ... in the base domain" (ibid.). But, Kim contends, the bridge law requirement is highly problematic, as there are serious questions about the availability of bridge laws and about their ability to provide genuine reductive explanations and ontological simplification – two crucial aims of any (cross domain) reduction. Let us consider these questions in more detail.

The chief problem with the "availability" question is that for any higher-level, "special" science that is a candidate for reduction there is the possibility of "multiple realization", and for any special science property P with multiple realizers it is not possible to provide a single correlate Q in the base theory to obtain a bridge law of the form $P \leftrightarrow Q$.[6] Kim, however, does not seem to regard this as an insurmountable obstacle to Nagelian reduction: for even if one cannot take Q as the disjunction of P's realizers in the base domain (as he once believed one could),[7] one can still settle for "species- or structure-specific bridge laws" of the form $S_i \rightarrow (P \leftrightarrow Q_i)$, and thus for "local" rather than "global" reductions. The possibility of local Nagelian reductions was countenanced by Kim in earlier writings (e.g., 1984, 1989, 1992); there is no reason to suppose that he denies, or is committed to denying, their possibility in his more recent work, given that such reductions on his account merely require the availability of restricted bridge laws. So, for Kim, the "availability question" should not raise a serious objection against the possibility of Nagelian reductions. We shall return to the issue of local reduction later; now let us turn to the "explanatory" and "ontological simplification" questions.

These questions, according to Kim, "pose serious challenges to the whole idea of Nagel reduction and its appeal to bridge laws" (p. 95). These challenges are especially perspicuous in the case of a Nagelian reduction of psychology to neurology. "C-fiber stimulation correlates with pain .... But why? Can we understand why we experience pain when our C-fibers are firing, and not when our A-fibers are firing? Can we explain why pains, not itches or tickles, correlate with C-fiber firings? ... Why is any sensory quality experienced at all when C-fibers fire?" (p. 95). The problem is that bridge laws are themselves "unexplained auxiliary premises" in derivations and are themselves in need of explanation; that pain (rather than itches or tickles) is nomologically correlated with the firing of C-fibers (rather than with the firing of A-fibers), is an unexplained brute fact that "will not advance our understanding of mentality by an inch" (p. 96). So "Nagelian reductions, whether global or local, do not give us reductions that explain" (ibid.). Furthermore, since bridge laws are supposed to be contingent laws, the concepts and properties they relate remain distinct, and so Nagel-reduction gets us neither conceptual nor ontological simplification. To the contrary, Kim continues, since bridge laws must be added to the laws of the base theory to make the derivations possible, there is actually an ideological and ontological price to be paid: "the price paid is the addition of the bridge laws as new basic laws of the base theory, and moreover these laws, by bringing with them new descriptive terms, will expand both the language and the ontology of the base theory" (p. 97). If we want reductions that yield ontological simplification, we should "find a way of enhancing bridge laws, $M \leftrightarrow P$, into identities $M = P$" (p. 97); such identities would "serve to answer the explanatory question as well: M and P are coinstantiated because they are in fact one and the same property. Identity takes away the logical space in which explanatory questions can be formulated" (p. 98). But, Kim concludes, there is no hope of turning bridge laws into identities: "If M and P are both intrinsic properties and the bridge law connecting them is contingent, there is no hope of identifying them. Distinct properties are just distinct, and we can't pretend they are the same .... Moreover, if $M \leftrightarrow P$ is contingent, the identification of M with P must be made consistent with the [Kripkean] thesis ... that identities whose terms are "rigid" are necessary. For if $M = P$ is necessary, $M \leftrightarrow P$ cannot be contingent – unless either M or P is nonrigid" (p. 98).

These challenges are important and need to be addressed if one is to resist Kim's indictment of Nagel-style reductions. Let us begin by raising the following preliminary questions. (1) Is it really the case, as Kim thinks it is, that the use of bridge laws in the derivation of a law L of a reduced theory requires "the addition of bridge laws as *new basic laws of the base theory* ... [and the consequent expansion of *both the language and the ontology of the base theory*" (p. 97, my italics)? It is true that in order to derive (the laws of) a reduced theory (T) from the base theory (T^*) with the help of bridge laws (BL) we need the *conjunction* (T^* & BL) as premise; but it does not follow from this that T^* itself has changed, or that its content or integrity have been altered. If it had, we could no longer say that the target theory has been reduced to the *base* theory; we would have to say instead that the target theory has been reduced to a hybrid theory containing both the base theory and the bridge laws. But that seems incorrect; for while the bridge laws may play a role in the derivation as auxiliary premises, they are not part of what the target theory is reduced to. (Thermodynamics, it is generally claimed, is reduced to statistical mechanics, not to a theory comprising statistical mechanics and bridge laws.)[8] It is true that sometimes a reduction may require the expansion of the base theory in order to "incorporate" the laws of the target theory (reformulated in the vocabulary of the base theory); but this does not mean that what gets incorporated are the bridge laws as well: these have served their purpose in bridging the two heterogeneous vocabularies, as needed for the derivation, but they do not become part of the base theory (indeed, if they did, so too would the laws of the target theory in their original form, since they would be derivable within the expanded theory, and there would be no reduction at all). So the use of bridge laws implies no expansion of the ideology or the ontology of the base theory.

(2) Is it really the case that the (biconditional) bridge laws deployed in a Nagelian reduction cannot be "enhanced" into identities? The common view on the matter is rather that successful reductions via bridge laws often provide excellent inductive warrant for supposing that the descriptive terms appearing in the relevant biconditionals are not merely coextensive but actually refer to one and the same property.[9] In other words, successful reductions often warrant re-interpreting the bridge laws as expressing "theoretical identities", thereby leading to ontological *simplification* rather than to ontological *expansion*, as Kim seems to suppose. Why then does Kim rule out that bridge laws can be reinterpreted as identities? As the quotation at the end of the next-to-the-last paragraph indicates, he seems to provide two related arguments. The first makes use of the following general conditional as major premise: "If M and P are intrinsic properties and $M \leftrightarrow P$ is contingent, then M and P must be distinct" (p. 98). The (tacit) minor premise is that the properties related by Nagelian bridge laws are intrinsic, and the bridge laws themselves contingent. The second argument is like the first except that it makes use of the Kripkean thesis that identities both of whose terms are rigid designators must be necessary. (The link between the two arguments, presumably, is that intrinsic properties correspond to rigidly designated properties.) The argument employs the following major premise: "If M and P are rigid and $M \leftrightarrow P$ is contingent, then $M = P$ can't be necessary, and thus M and P must be distinct". The (tacit) minor premise is that the terms related by Nagelian bridge laws are rigid, and the laws themselves contingent. Given the above link, the major premise of each argument assumes that intrinsic properties, or rigidly designated properties, can only be identified if their coinstantiation is non-contingent. Presumably by 'non-contingent' Kim means '*logically*

(or *analytically*) non-contingent', and thus 'assertible a priori'. This assumes a very stringent criterion of property identity, something as strong as the logical equivalence or synomymy of the relevant property designators. This seems like an implausibly stringent criterion, one that assumes a rather 'abundant' conception of properties, and that rules out such theoretical, a posteriori property identifications as the identification of water with H_2O, temperature with kinetic energy, light with electromagnetic radiation, etc.[10] But even if one accepts the major premise of either argument, it remains unclear that one must accept the minor premise as well: need the properties figuring in bridge laws be regarded as intrinsic properties, or as properties designated by 'rigid' terms? Are the mean kinetic energy of a collection of molecules, the rate of change of molecular momentum per unit area, the capacity to transmit genetic information, and so on, intrinsic properties? Or are their corresponding designators rigid? I confess that I have no clear intuitions on the matter, but I suspect there are no clear intuitions to be had.[11] I think it's fair to say that Kim's above arguments are at best inconclusive: the possibility that bridge laws can be "elevated" to identities, thereby sharing in the explanatory benefits of the latter, remains open.

3. REINTERPRETING THE NAGELIAN MODEL

I have been supposing, with Kim, that bridge laws are indeed "at the heart" of Nagelian reduction. This may seem uncontroversial, given Nagel's "connectability" requirement as one of the formal conditions for intertheoretic reduction. However, there are questions that can be raised about the alleged centrality of the role of bridge laws in reduction, even in the context of Nagel's own articulation of the model. It is instructive to briefly review Nagel's own (admittedly highly simplified) account of a stock example of intertheoretic reduction: the "incorporation" of (a fragment of) thermodynamics within mechanics (Nagel, 1961, pp. 342–345).[12]

Nagel begins by reminding us of a few historical facts about the evolution and systematization of thermodynamics, noting its affinities with mechanics (in the use of notions like volume, weight and pressure, and of such laws as Hook's law and the law of the lever), as well as its distinctive character as a "relatively autonomous physical theory". He then notes how "experimental work early in the nineteenth century on the mechanical equivalent of heat stimulated theoretical inquiry to find a more intimate connection between thermal and mechanical phenomena ..." – thereby exhibiting the desire for unification as a motivational factor for the reduction. He then proceeds to outline how the reduction was effected, limiting himself to a simplified account of the derivation of the Boyle-Charles' law from the assumptions of the kinetic theory of matter. These assumptions, statable in terms of the fundamental notions of mechanics, include such familiar idealizing assumptions as that the gas is composed of a multitude of perfectly elastic molecules of equal mass and volume, that they occupy a container with a fixed volume V and perfectly elastic walls, that their positions and momenta can be statistically computed, that the pressure p they exert on the walls of the container is equal to the average of their momenta, etc. From these assumptions, Nagel concludes,

> . . . it is possible to deduce that the pressure p is related in a very definite way to the mean kinetic energy E of the molecules, and that in fact $p = 2E/3V$, or $pV = 2E/3$. But a comparison of this equation with the Boyle-Charles' law (according to which $pV = kT$, where k is a constant for a given mass of gas, and T its absolute temperature) suggests that the law could be deduced from the assumptions mentioned *if* the temperature were in some way related to the mean kinetic energy of the molecular motions. Let us therefore introduce the postulate that $2E/3 = kT$, that is, that the absolute temperature of an ideal gas is proportional to the mean kinetic energy of the molecules assumed to constitute it. (p. 344)

Nagel's account of the derivation of the Boyle-Charles' law, as part of the reduction of thermodynamics (T) to statistical mechanics (T^*), can be seen to involve the following three basic steps:

1. The formulation of a number of limiting assumptions and initial conditions (*LA/IC*) centering around the identification of a fixed volume of an ideal gas with a fixed number of molecules.
2. The derivation, from the principles of statistical mechanics (T^*) together with *LA/IC*, of a law L^*, namely $pV = 2E/3$, which is the mechanical counterpart (an "image" or "close analogue")[13] of the Boyle-Charles' law $pV = kT$ (call this L). L^* is of course entirely in the vocabulary of T^*.[14]
3. The postulation of a bridge law (*BL*), $2E/3 = kT$, consequent upon a "comparison" of L^* with L, enabling the formal derivation of L from L^*.

It is evident that bridge laws play a role only in step 3, and that step, arguably, is not as central to the reduction as popular belief has it. Indeed, as Beckermann (1992) explains,

> . . . the discussion [on reduction] over the last decades has shown that bridge laws do not play such an essential role in reduction as the classical account supposed . . . [I]t is not essential for the reduction of one theory T_1 to another T_2 that the laws of T_1 themselves can be deduced from the laws of T_2. What is really essential is that one can deduce from T_2 something that is commonly called an image of T_1. In other words, it is essential to show that there are properties which we can refer to by terms definable in terms of T_2 which play

(almost) the same role as the properties which are referred to by the terms of T_1. And it is just this that is shown by the possibility of deducing an image of T_1 from T_2. (p. 108)

So it appears that it is the first two steps above that are really crucial to the reduction, for they alone are both necessary and sufficient for the derivation of the respective "images" of the laws of the theory to be reduced; it is of course just these images, not their counterparts in the reduced theory, that are retained or "incorporated" within the reducing theory. Thus, in the present case, it is the deduction of L^*, not of L, that is essential to the reduction of thermodynamics; and the deduction of L^* from (T^* and *LA/IC*) can be effected without any bridge laws. On this account, bridge laws are quite inessential. Of course, once bridge laws *are* available, they can directly be employed to deduce the actual laws of the reduced theory from their image – e.g., we can use *BL* (i.e., $2E/3 = kT$) to deduce L from L^*; but recall that *BL* was itself derived from (a "comparison" of) L^* with L,[15] so the deducibility of L from L^* & *BL* is quite trivial – it merely serve the logical/expository function of formally exhibiting a *result* of the reduction and certifying that it has been successful, rather than the scientific/methodological function of actually *effecting* the reduction. The essential part of the reduction – the real scientific achievement – has been accomplished once the steps corresponding to 1 and 2 above have been completed for (the images of) all the laws of T; the deduction of T-laws from T^*-laws via bridge laws is, to put it bluntly, a house-keeping chore for the logician of science.[16]

It might be objected that although this account does reflect (some of the) recent wisdom on the nature of reduction, it does not faithfully represent the Nagelian "classic" account. This objection raises a difficult question of interpretation that, however, we need not go into; suffice it to say that steps 1–3 above closely reflect Nagel's own classic account, and that the important first two steps (as well as the first part of the third, which explains where bridge laws come from) are generally omitted from popular accounts of Nagelian reduction. In any case, even if the account I have given does not give a totally faithful interpretation of Nagel, I believe it is sufficiently *Nagelian* in capturing the idea of a type of intertheoretic reduction that essentially involves, in Nagel's own words, the "incorporation" (1961, p. 342) of the reduced theory within the base theory via the derivation of laws like L^* cast in the vocabulary of the base theory; how central Nagel himself thought the derivation of such laws as L itself to be to the reduction, is a question we can well leave open. My chief concern has been to establish that a Nagelian need only regard bridge laws as central in a merely relative sense: they are central to the derivation of laws like L (i.e., the very laws of the reduced theory); but the derivation of such laws (as distinct from the derivation of their "images"), is not central to the reduction itself.[17]

Let us now return to Kim, and compare his account of the role of bridge laws in Nagelian reductions with the one just given:

> ... it is obvious that the availability of bridge laws is the critical factor for questions about Nagel-reducibility of theories... If each predicate or property, M, in the target domain can be correlated with a coextension, P, in the base domain, that in itself would guarantee Nagelian reduction .... For let L be any law in the theory to be reduced: either L is derivable from the base theory via biconditional bridge laws or it is not. If it is, Nagel reduction goes through for L. If it is not, rewrite L in the vocabulary of the base theory using bridge laws as definitions, and add this rewrite as an additional law of the base theory. L would then be derivable from the laws of the enhanced base theory via the bridge laws, again satisfying Nagel's derivibility condition. The rewrite of L is a true lawlike generalization expressed entirely in the vocabulary of the base theory, and the original theory was incomplete in that it missed a true generalization within its domain. (p. 91)

Kim is indeed right in supposing that *if* each M in the target domain can be correlated with a coextension P in the base domain (that is, if the relevant bridge laws are available), then "that in itself would guarantee Nagelian reduction". But, as indicated in the foregoing, bridge laws are generally not available (or at least not required) *prior* to the reduction; they become available after the essential part of the scientific reduction, as outlined in steps 1 and 2, has already been accomplished – in particular, after an image of L (in our case, L^*) has been derived from the base theory, and, further, after that image has been "compared" with L itself. And since the biconditional bridge law has thus been derived from L and L^*, it is true but uninteresting that L itself can be derived from L^* (and thus from the base theory) via the bridge law. The trivial role of bridge laws in Kim's account of Nagelian reduction is especially evident in the second case Kim considers – the case where L is not derivable from the base theory via bridge laws. In that case, we are told, just "rewrite L in the vocabulary of the base theory, and add this rewrite as an additional law of the base theory". (The rewrite of L is, of course L^*, which in this case, ex hypothesi, is not derivable from the base theory.) Then of course L would be derivable from the "enhanced base theory" via bridge laws: how could it fail to be so derivable, since now the enhanced base theory contains L's rewrite (i.e., L^*), which was itself derived from L via bridge laws! Surely Nagelian reduction can't be this trivial. What's missing from Kim's account, I suggest, is precisely the condition that L^* (L's "rewrite") has to be derivable from the base theory (and of course the limiting assumptions mentioned in step 1) *independently* of any bridge laws.

4. KIM'S FUNCTIONAL MODEL OF REDUCTION

Let's now turn to Kim's "functional model" of reduction. Whereas on the Nagelian model the focus is on the reduction of theories via the derivation of laws, here the focus is on the reduction of properties via their identification with base properties:

> ... to reduce a property M to a domain of base properties, we must first "prime" M for reduction by construing, or reconstruing, it *relationally* or *extrinsically*. This turns M into a relational/extrinsic property. For functional reduction we construe M as a second-order property defined by its causal role – that is, by a causal specification H describing its (typical) causes and effects. So M is now the property of having a property with such-and-such causal potentials, and it turns out that property P is exactly the property that fits the causal specification. And this grounds the identification of M with P. (p. 98)

This conception of reduction, Kim remarks,

> ... accords well with the paradigm of reduction in science. To reduce a property, or phenomenon, we first construe it – or reconstrue it – functionally, in terms of its causal/nomic relations to other properties and phenomena. To reduce temperature, we must first stop thinking of it as an intrinsic property but construe it as an extrinsic property characterized relationally, in terms of causal/nomic relations, perhaps something like this: it is that magnitude of an object that increases (or is caused to increase) when the object is in contact with another with a higher degree of it, that, when high, causes a ball of wax in the vicinity to melt ... [etc.]. Here is another example: the gene is that mechanism in a biological organism causally responsible for the transmission of heritable characteristics from parents to offsprings. To be transparent is to have the kind of molecular structure that causes light to pass through intact. And so on. We then find properties or mechanisms, often at the microlevel, that satisfy these causal/nomic specifications and thereby fill the specified causal roles (pp. 24–25).[18]

Before we proceed to compare Kim's account of reduction with the Nagelian account, the following comments are in order. First, since Kim's model of functional reduction involves the *identification* of the property to be reduced – a property defined by its causal role – with a role-filler property in the base domain, the model, if successful, clearly accounts for what Kim holds as paramount aims of a reduction, namely, ontological simplification and explanation: if M and P are the same property, then their respective domains do not after all involve different ontologies, and it's no longer a mystery why they should be lawfully coinstantiated. We saw earlier, however, that, *pace* Kim, such intertheoretic property identifications and the explanatory benefits that come with them are not ruled out by Nagelian reductions; so reductive property identifications should not be regarded as a virtue for which Kimian reductions should be privileged over Nagelian reductions.[19]

Second, the identification of a "functionalized" property – a second-order property defined by its causal role – with a first-order filler-property, is problematic.[20] This is an issue that Kim himself raises (p. 99, note 11): "If *M* is an extrinsic/relational property and *P* (presumably) isn't, how can they be one and the same property? ... If *M* is a causal role and *P* its occupant, how could *M* and *P* be the same property? How could roles be identical with their occupants?" Indeed, how can second-order properties be identified with first order properties? If they cannot, then Kimian reductions no longer have the virtue of yielding property identifications and ontological simplification. If they can, then the role/filler distinction collapses, and we are left with no clear idea of what a "functionalized property" and "reduction by functionalization" amount to. Confronted with the apparent incoherence involved in the above identifications, Kim seeks remedy in semantic ascent: replace talk of second-order *properties* with talk of second-order *property designators*, and let these (non-rigidly) designate whatever first-order, role-filler base properties, (rigidly) designated by first-order designators, happen to satisfy the role-concept expressed by the second-order designators.[21] It's unclear how this strategy removes the incoherence, for now we have different-order designators – one expressing a role-concept, the other expressing a role-filler concept – designating one and the same property. Perhaps this logico-semantic problem can somehow be resolved, since conceptually distinct designators can still be coreferential (though it's not clear they can still be so if they belong to distinct logical orders). In any case we are now confronted with a further problem, for the ascent from second-order functional properties to second-order functional designators raises the threat of eliminativism with respect to those very properties that were candidates for reduction by the method of functionalization: if there are no second-order, functional properties, then such properties as temperature, transparency, pain, etc., which on Kim's account were to be (re)construed as functional properties, do not exist after all, and if they don't, it's difficult to see how they can be *reduced* to base properties at all – except, perhaps in the sense of "reduction by elimination", in which case the reduction would no longer enable us to answer such explanatory questions as "Why is the temperature of an ideal gas always coinstantiated with the mean kinetic energy of its molecules?" or "Why is pain always coinstantiated with the firing of C-fibers?". There would be no property identities to which we could appeal in order to answer, or forestall, these questions.[22]

Third, and most importantly, how are we to account for multiple realizability on Kim's model of reduction? As soon as we acknowledge that the same "role" property may have different realizers, or fillers, for different types of individuals or for the same individual on different occasions, the question arises: how can that role-property be identical with any one of its realizers, unless each realizer is identical with each of the others? Indeed, how can we account for, or reconstruct, the asymmetrical realization-relation in terms of a symmetrical and reflexive relation like identity? Kim's strategy for dealing with multiple realization is just to acknowledge that a functional property M with multiple realizers, say P_1 and P_2, is "a property we will have to learn to live without" (p. 106). For we can be serious about M as a property "only if we are willing to countenance M as a disjunctive property, $P_1 v P_2$, but there are weighty reasons for rejecting disjunctive properties of this kind" (ibid.); *ergo*, there are weighty reasons against taking M seriously as a property. The eliminativist threat is now turned into a virtue. There is no unitary property M over and above each of its realizers: "the fact that something has M amounts to the fact that it has P_1 or it has P_2" (p. 107). Since P_1 and P_2 are distinct properties (each realizing M for different species or structure types), the unrestricted identification of M with either P_1 or P_2 is of course "out of the question" (p. 110); the functional reduction of M "consists in identifying [each instance of] M with its realizer P_i relative to the species or structure under consideration" (ibid.). As long as there is variation in what property occupies the M-role, the relevant identifications need to be restricted: not simply $M = P_i$, but M-for-$S_i = P_i$. Such kind-restricted property identifications are Kim's way of dealing with multiple realization. The result, however, is that we have lost M as a single, unified property: "multiply realized properties are sundered into diverse realizers in different species and structures" (p. 105).[23] So too are their causal powers sundered, for these get identified with the heterogeneous causal powers of their diverse realizers on each occasion, with the consequence that multiply realizable functional properties lack "the kind of causal homogeneity and projectibility that we demand from kinds and properties useful in formulating laws and explanations" (p. 110). What lends unity to talk of functional properties, Kim concludes, is "conceptual unity, not the unity of some underlying property": functional talk "may well serve important conceptual and epistemic needs, by grouping properties that share features of interest to us in a given context of inquiry" (ibid.), but it does not bring any new properties into existence.

It should be clear that Kim's functionalism in his model of reduction runs against the grain of classical functionalism, for the central idea behind classical functionalism (the functionalism of Putnam and Fodor that was intended to replace the Smart-Feigl "identity theory") was that a functional property retains its identity and projectibility across heterogeneous physical realizations. The search for unification and nomological homogeneity in spite of physical diversity is what was supposed to drive functionalism; if we "sunder" the multiply realized properties into their diverse realizers, then functional-level homogeneity, generality, and projectibility are lost. It's not clear how this unity and homogeneity can be restored by Kim's "nominalization" of functional properties, that is, by his replacing functional properties by functional *concepts* or *designators*: how can using a functional designator 'M' bring us "conceptual unity" if 'M' turns out to be semantically *equivocal*, referring not to a unitary property across instantiations, but now to this and now to that member of a heterogeneous class of properties? The fact is that, for Kim, whatever unity is to be found at the functional level derives from unity at the realization level: for example, what makes psychology possible as a unitary science, he tells us, is that "conspecifics share largely similar neural systems... [The] uniformity of human psychology, to the degree that it obtains, is due to similarity in our neural systems – that is, the uniformity of human physiology" (p. 94). That may well be so; but if it is, it goes against the spirit of orthodox functionalism, which, in embracing the multiple realizability thesis, asserts the possibility of a unitary psychology *despite* physiological diversity. To infer physiological uniformity from psychological uniformity is, in effect, to renounce classical functionalism.

The aim of the foregoing comments was not to undermine Kim's model of reduction, but only to raise some questions about the ostensibly functionalist features of it. The gist of my worry has been that to the extent that we strive to assimilate the functionalist features of Kim's model to the classic, "multiple-realization" functionalism of the Putnam-Fodor variety, it becomes doubtful that the model can serve its professed aim of providing reductions of "functionalizable" properties by reductive property-identifications. Either a functional property is multiply realizable – in which case it can't be reductively identified with any of its realizers; or if it is to be so reductively identifiable, then it must be given up as a unitary property and "sundered" into its structure-relative realizers. The latter is admittedly a coherent form of reductionism – indeed, of classical type-type reductionism – but of course it has little to do with classical functionalism.[24] Despite Kim's retreat from unrestricted type-identities, we are still left with type-identities; for restricted identities of the form M-for-$S = P$ are still type-identities – precisely the sort of identities that classical reductionism aimed to establish via Nagelian intertheoretic reduction. Indeed, as I shall now go on to argue, Kim's functional model is not substantially different from the Nagelian model which it was intended to replace.

5. COMPARING KIM'S FUNCTIONAL MODEL WITH THE NAGELIAN MODEL

On Kim's account, for a system of kind S to have a functional property M is for it to have a (first order) property P that "fills" or "occupies" the causal role definitive of M. How is such a causal role to be defined? Presumably, by articulating the set of causal/nomic relations into which M enters with other states $M_1, M_2 \ldots M_n$ (including input and output states), according to some theory T about S. The causal role definitive of pain for humans is specified by articulating the way pain interrelates with other states under appropriate conditions according to "folk psychology" (or according to some preferred psychological theory). The causal role definitive of temperature for an ideal gas (if such a property is functionalizable as Kim holds) is specified by articulating the way temperature nomically interrelates with other states of the gas according to thermodynamics under appropriate conditions. And so on for other functionalizable states or properties. What this amounts to is that, in general, the causal role of a functional or functionalizable property is given by the *set of nomic connections* satisfied by the property or state in question, i.e., by the set of causal *laws* in which the property figures, according to theory T. We can represent this as follows:

The role of M_i for structure type S is R according to theory T iff: There are properties $M_1 \ldots M_n$, and laws L_1, L_2, etc. of T about S such that $[L_1(M_1 \ldots M_i \ldots M_n)$; $L_2(M_1 \ldots M_i \ldots M_n)$, etc.], where $L_k(M_1 \ldots M_i \ldots M_n)$ is to be read as: M_i interrelates with $M_1 \ldots M_n$ according to law L_k of T.

To use a stock example, mental state M has the role of pain for S according to folk psychology iff there are folk-psychological laws according to which (under appropriate conditions) any system s of type S is in M if s's tissues are injured while s is not under anaesthetic; and if s is in M then s will wince, believe she is in pain, seek help, etc. in the presence of other M-states, etc. And if temperature is functionalizable, in Kim's sense, then a state (magnitude) T of an ideal gas g has the role that classical thermodynamics assigns to temperature iff kT is directly proportional to the product of its pressure p and volume V (i.e., if it obeys the Boyle-Charles' law $kT = pV$), and so on for other pertinent laws of thermodynamics in which T figures.

Let us now ask what is meant by saying that a given property P_i "fills the role"R assigned to M_i by a theory T for structure type S. Expressions like "filling a role", "occupying a role", etc. are strictly speaking metaphorical and need explicating. I suggest the explication goes something like this: For any s of structure type S, if s possesses M_i, then s possesses

P_i, and P_i belongs to a set of (possibly micro-level) properties $P_1 - P_n$ assigned to s by a theory T^* about S, which nomically interrelate with one another in such a way that, under specified boundary conditions, for every law $L_k(M_1 \ldots M_i \ldots M_n)$ of T there is a law L_k^* $(P_1 \ldots P_i \ldots P_n)$ of T^* (or derivable from T^*), such that there is a structure-preserving mapping from L_k to L_k^*. When this is the case we can say that P_i occupies the M_i role, in virtue of the fact that P_i interrelates with $P_1 - P_n$ the way M_i interrelates with $M_1 - M_n$. (This is of course a highly idealized account; for some of the laws of T there may only be an approximate mapping into laws of T^* holding under specific boundary conditions; or L_k^* may contain terms which it shares with L_k., as in Nagel's example of thermodynamics' reduction). Thus for example we may say that C-fiber firing fills the role of pain inasmuch as corresponding to the folk-psychological laws that under appropriate conditions tissue damage causes pain and that pain causes wincing, there are neurophysiological laws to the effect that under those conditions tissue damage causes C-fiber firing and C-fiber firing causes wincing. And we may say that the mean kinetic energy of the molecules of an ideal gas fills the role of temperature in thermodynamics inasmuch as corresponding to the Boyle-Charles' law $pV = kT$ there is the statistical-mechanical law $pV = 2E/3$. I think the general idea is clear: we can say that P in T^* fills the role definitive of M in T if T^* together with appropriate boundary conditions enables the derivation of laws which are "analogues" or "images" of the T-laws in which M figures, and P occupies in T^*-laws the "position" that M occupies in T-laws.

It will have been noticed that this account of "role filling" essentially respects the steps outlined above in connection with the Nagelian reduction of thermodynamics to statistical mechanics. As we have seen, in order for P_i to play the M_i role, there must be a law $L^*(P_1 \ldots P_i \ldots P_n)$ isomorphic with $L(M_1 \ldots M_i \ldots M_n)$.[25] On the Nagelian account, the L-law will be a law of the theory T targetted for reduction, and the L^*-law will be the "image" of the L-law, derivable from the base theory T^* together with pertinent assumptions about the composition of the structure type to which L applies. On Kim's account of functional reduction, the (structure-specific) role filler P_i of the functionalized property M_i gets identified with M_i itself in virtue of its occupying in L^* the position that M_i occupies in L. But an analogous move occurs in the third step of a Nagelian reduction: for once an image L^* is available for a given law L, then, as Nagel pointed out, we can postulate biconditional bridge laws linking certain predicates of L^* with corresponding predicates of L, and these bridge laws, as we argued earlier *contra* Kim, can inductively be interpreted as expressing property identities once the intertheoretic reduction has been carried out in a sufficiently "smooth" and comprehensive way. The actual derivation of L from L^* (and, more generally, of T-laws from T^*-laws) via bridge laws has an obvious counterpart in Kim's functional model: for given the isomorphism between L and L^*, P_i fills the M_i role only if each of the other properties $P_1 - P_n$ in L^* can be linked biconditionally to (or even identified with)[26] a corresponding M-property in L; and so we have the equivalent of bridge laws enabling the derivation of L from L^* under Kim's model.

6. PSYCHOPHYSICAL REDUCTION: TAKING STOCK

Where does this leave us with respect to the issue of psychophysical reduction? We have seen that Kim's response to the anti-reductionist, multiple-realization argument is to opt for kind-restricted reductions: where M has diverse realizers in different species or structure types, we get restricted reductive identifications of the form M-for-$S = P$ for each kind S in which M may be (uniquely) realized. As noted, these kind-restricted identities are still type-identities, the sort of identities that, as we saw, may well be available on the Nagelian model as a result of the isomorphism between the two levels of law involved in the reduction and the consequent postulation of biconditional bridge laws linking the two levels. Whether we cast the reduction in the Nagelian or in the Kimian form, multiple realizability does not seem to constitute a serious obstacle to psychophysical reduction, for – the reductionist will argue – kind-restricted or domain-specific reductions "are the rule rather than the exception in all of science" (Kim, 1989, p. 39). After all, it is not temperature *as such* that gets reduced to mean kinetic energy, but temperature *for ideal gases* – temperature for solids, or for a plasma, or in a vacuum, has different physical realizers, and is identified with different physical kinds in each case. If reduction across kinds or structure types is not demanded in science generally, why demand it in the psychological domain?

An answer to this question might be that whereas physics has no need for a generic notion of "temperature as such", that is, of temperature as attributable across diverse physical states, psychology often needs concepts that apply across species or physical structure kinds: the very possibility of a comparative psychology, and the very relevance of experimental work on non-human animals for human psychology (e.g., in connection with learning, problem solving, perception, etc.), seem to indicate that psychology needs concepts that are projectible across species and physical/biological structure kinds. But there is no need to pursue this line of response to the restricted-kind reductionist. For there is of course nothing implausible in the idea of relativizing psychological types to species or to other theoretically interesting physical/biological structures (so long as these are so identifiable independently of their satisfying the psychological types in question): human psychology is obviously worth pursuing even if does not significantly overlaps with canine (or "Martian") psychology. However, the multiple realizability argument is supposed to cut deeper, according to its proponents: psychological types are multiply realizable even *within* species or structure types – indeed even within the same individual at different times (as Kim readily acknowledges; p. 94; and 1989, p. 273), and so the restricted-kind reductionist is again faced with essentially the same problem: how can we get kind-restricted reductions if a psychological type is multiply realized within that very kind or individual?[27] Further narrowing the kinds into subkinds or into ever more restricted physical/biological structures may be self defeating, as then, with the increasing loss of generality, the reductions run the risk of being purely ad hoc and theoretically uninteresting.[28]

The problem that multiple realizability poses for the reductionist can then be stated, in its most general form, as follows: If psychological types, like functional types in general, "cross classify" physical types even when relativized to species, structure types, or individuals across times, how can they be reduced to physical types? Whether we construe reduction in the Nagelian form or in Kim's "functionalist" form, diversity of physical realization precludes the identification of a psychological property, or type, with a physical type, no matter whether we base such identifications on the postulation of biconditional bridge laws, as on the Nagelian model, or on the satisfaction of the "role-filler" relation, as on Kim's model. But it is just such identifications that reduction, presumably, is supposed to give us.

Perhaps the most promising reductionist response at this point is simply to question the just mentioned supposition: *must* scientific reduction give us property identifications? Kim certainly thinks so, and so too does the popular view on Nagelian reduction. But perhaps this assumption should be questioned, and, indeed, it *has* been questioned by some. In a well known paper, Richardson (1979) denied that even Nagel's own account of intertheoretic reduction requires such identifications, for Nagel did not require that bridge laws have a biconditional form; indeed Nagel is clear on this, as Richardson points out: "the linkage between *A* [a term in the secondary science] and *B* [a term in the primary science] is not necessarily biconditional in form, and may for example be only a one-way conditional: If *B*, then *A*" (Nagel, 1961, 355n; quoted in Richardson, 1979, pp. 548–549). So Richardson concludes:

Reduction demands only that there be a functional relation between the physiological and the psychological domains: each physiological type, within specified boundary conditions, should map onto a psychological type. The "suitable relations" demanded by the condition of connectability need not be biconditional. Derivability, with its explanatory parsimony, is adequately accounted for, in turn, if only we find sufficient conditions at a lower level of organization capable of accounting for the phenomena initially dealt with at a higher level; and this too requires no more than a mapping *from* lower *to* higher level types and *not* a mapping from higher level to lower level types. (Richardson, 1979, p. 548)

I think Richardson is absolutely right in holding that the reduction of psychology requires no more than a *mapping* from the physiological base level to the higher psychological level – a mapping effected by means of one-way conditionals linking each physiological type with a psychological type as a sufficient condition for the latter.[29] But Richardson is mistaken in thinking that this is consistent with Nagel's conception of reduction, *if* the derivability constraint on reduction is indeed to be satisfied. Having quoted the above passage from Nagel, Richardson should also have quoted the sentence immediately following it: "But in this eventuality [i.e., when the linkage between A and B is not biconditional in form] 'A' is not replaceable by 'B', *and hence the secondary science will not in general be deducible from a theory of the primary discipline*" (Nagel, 1961, p. 355n; italics added). The reason is not difficult to see: take any secondary science law L of the form $M' \rightarrow M''$ to be derived from a primary science (base) law L^* of the form $P' \rightarrow P''$; such a derivation will not be possible if we merely rely on one-way bridge laws such as $P' \rightarrow M'$ and $P'' \rightarrow M''$, as Richardson proposes: for we *cannot* deduce L from L^* via these bridge laws alone.[30] So if one insists on the derivability of the *very laws* of the secondary science from the laws of the primary science via bridge laws as an essential constraint on Nagelian reduction, as Kim does, then we cannot hope to have a Nagelian reduction of psychology to physiology if all we have is the one-way bridge laws that Richardson proposes. Those bridge laws would not give us the requisite derivations, let alone the property identifications and ontological simplification that Kim expects from *bona fide* reductions.

However, Richardson's conception of psychophysical reduction (or, in general, "special-science" reduction in Fodor's sense) *can* be reconciled with the Nagelian model if this is construed (or reconstrued) along the lines I suggested in Section 3 – if, that is, we take the essential core of the reduction to be *not* the derivation of the *actual laws* of the target theory from the laws of the base theory, but merely the derivation of the *images* of such laws under appropriate boundary conditions. *If*, for every law L of the theory T to be reduced we are able to derive a law L^* from the base theory T^*, such that L^* is an image of L, then we have effectively reduced T to T^*. A "comparison" of L^* with L, as Nagel pointed out, may lead to the postulation of bridge laws linking certain expressions in L^* with certain expressions in L. In the simplest case the bridge laws will have a biconditional form; in this case they will obviously license the derivation of L from L^* and will inductively support the theoretical identification of the corresponding properties. On the other hand the bridge laws may only be one-way conditionals, in which case the derivation of L from L^* may or may not be possible depending on the form of L and L^* and on the direction of the one-way conditionals: as noted in the previous paragraph, the derivation will not be possible if the conditionals are exclusively from T^* expressions to T-expressions (in which case it may be inappropriate to call such conditionals 'bridge laws' at all – 'supervenience laws' would be a more accurate label), but it may be possible in other cases.[31] But insofar as the bridge laws are not biconditional in form, theoretical property identifications are of course out of the question: T may have alternative reduction bases, or its laws may have alternative images in the same or in different base theories. This would seem to be the case for psychology, if multiple realizability is taken seriously. I am inclined, with Richardson, to count the sort of reduction he envisages for psychology as a case of genuine intertheoretic reduction – even as a particular case of *Nagelian* reduction, if we are entitled to regard as Nagelian the liberalized form of reduction I described earlier. On Richardson's conception, a reduction of psychology requires no more than a *mapping* of types from physiological laws to psychological laws (within specified boundary conditions) – a mapping effected by means of one-way " bridge laws" or supervenience conditionals which, while not licensing the deduction of the psychological laws from the physiological laws, do guarantee that each psychological law has a physiological *image* (leaving it open whether it has other images in the same or different "physiologies" or physical structures). Consistently with the multiple realization thesis, there is no problem in supposing that psychology is "Nagel-reducible" to physiology if, as Richardson suggests, Nagelian reduction, in general, requires no more than the holding of a functional relation from the reducing to the reduced domain. Such functional relation may in some cases (as in the case of the reduction of theormodynamics to statistical mechanics) take the form of biconditional bridge laws, or of bridge laws licensing the deduction of laws; but this, as we have suggested, need not in general be the case, and it is not likely to be the case for psychology. On this conception Nagelian reduction admits of various "grades" of reduction: in its strongest form, when biconditional bridge laws are invoked, it *permits* (*pace* Kim) property identifications and ontological simplification, but it does not in general *mandate* them: a reduction is capable of achieving its goal by merely effecting the derivation, *within the reducing theory* and within specified boundary conditions, of the "base" laws on which the laws of the reduced theory, in virtue of satisfying the appropriate mapping, can be said to supervene.[32] So-called non-reductive materialists (at least those of the "non-anomalous" sort) should have no problem with this weaker form of Nagelian reduction, for it provides no challenge to their cherished multiple realizability thesis. Mental types remain distinct from physical types even while being brought together under a system of asymmetrical nomological dependencies. Surely, ontological *unification* through nomological unification must not be confused with ontological *simplification* through reductive type-identification; for certain domains, a reduction with the former more limited aim is all the reduction we should expect, and that sort of reduction is no threat to the champions of multiple realizability.

In conclusion, I think there are good reasons to resist Kim's contention that Nagelian reduction is "the wrong battlefield on which to contest the issue of reduction" (p. 26); to the contrary, the Nagelian model, broadly conceived, still provides an effective paradigm for explicating various forms or "grades" of reduction in science. Kim's own "functional" model of reduction, I argued, can itself be accommodated within the Nagelian model. To the extent that it insists on kind-restricted reductions involving the theoretical identification of the "functionalized" properties with their kind-specific realizers or "role fillers", Kim's model affords the equivalent of biconditional bridge laws mediating the derivation of the higher-level laws that articulate the "role" of the functionalized properties in the target theory, from the lower-level laws conformity to which enables the role-filler properties in the base to fill their respective roles. As we saw in Section 5, the role-filling relation implies a mapping relation; but the mapping relation need be one-to-one only if the functionalized properties have unique realizers. Kim insists on unique realizers in order to secure (kind restricted) reductive identifications and ontological simplification. But if kind-restricted reductions do not, as previously maintained, provide a generally satisfactory answer to the multiple realizability challenge, then there is no reason to insist that the mapping relation between the two domain be one-to-one: it could well be one-many, so as to allow different potential role fillers to be mapped onto the same higher-level functional property. This would be exactly the situation that Richardson, Fodor, and many others envisage for the reduction of psychology to physiology: such reduction should require no more than the specification of sufficient conditions at the physiological level capable of accounting for the phenomena at the higher level. Richardson and I are happy to subsume such reduction under the Nagelian model; Fodor and Kim, apparently, are not. But they are not, I submit, altogether for the wrong reasons: Fodor because he thinks that

Nagelian reduction requires property identifications while psychophysical reduction should not; Kim because he thinks that psychophysical reduction should yield property identifications but Nagelian reduction, even with biconditional bridge laws, can't provide any. I hope enough has been said to see why both views are mistaken.

ACKNOWLEDGEMENTS

I am grateful to Jaegwon Kim for stimulating conversation which sparked reflection on the central issues discussed in this paper. I am also grateful to Cliff Hooker for careful comments on an earlier draft of this paper, and for further subsequent discussion. I also benefitted from the helpful comments of an anonynous referee for this journal. Support from the Social Sciences and Humanities Research Council of Canada is also hereby acknowledged.

NOTES

1 See, e.g., various papers in Kim 1993a (esp. 1984, 1989, 1990, 1992), 1993b, and 1998.

2 Kim (1998), Nagel (1961).

3 See Bickle (undated) for a concise survey of critical work on Nagelian reduction.

4 All further page references will be to Kim (1998) unless otherwise indicated.

5 See, e.g., Hooker (1981, p. 49), for an explicit endorsement of the deductive structure of theory reduction. (As we shall see, it's not deduction as such that is often at issue in theory reduction, but *what*, exactly, can or needs to be deduced. See Sklar (1967) for an early discussion of this issue.)

6 I have been speaking of *properties* here, as Kim does. Nagel, however, warns us against the supposition that reduction (*qua* derivation of statements from sets of statements) is concerned with the "derivation of *properties*" (Nagel, 1961, pp. 364ff). It would thus be more appropriate (as an anonymous referee has pointed out) to speak of *predicates* rather than properties. Since the property or "nature" associated with a predicate is determined by the theory to which the predicate belongs, we need not suppose that the properties in the scope of the reduced theory are thereby less than real or somehow "illusory": we need merely suppose that the predicates of the reduced theory have been provided with new conditions of application, and the corresponding phenomena with an explanation, in terms of the primitives of the reducing theory (cf. Nagel, 1961, p. 366). Although the significance of the predicate/property distinction will later become an issue (see Section 4), nothing substantive in my present dicussion of Kim's views on the failings of Nagelian reduction turns on this distinction.

7 See, e.g., Kim (1984, 1989).

8 Still, one might ask, if reduction is explanation, and if the explanation (derivation) of T requires *both* T^* and *BL* as explanans, does it not follow that T is reduced to (T^* & *BL*)? One could speak that way, I suppose; even so, it would not follow that T^* has changed in the way Kim claims it has. But I think it's more appropriate to say that T has been reduced to T^* with the *mediation* of *BL* as auxiliary hypotheses.

[9] I call attention to this point in Marras (1993, p. 287): "The best case for property identification is, of course, one based on *theory reduction*: then the 'bridge' laws linking 'P' and 'Q' acquire the status of intertheoretic definitional equivalences, and on that basis we can make outright property identifications". (The "definitional equivalences" are, of course, empirically based and thus a posteriori). See also Causey (1972) for an argument for the claim that bridge laws must be "contingent attribute-identities". Bridge laws guarantee that the co-instantiated properties play isomorphic theoretical roles – a condition often taken as sufficient for "contingent" property identity (in the sense of Putnam, 1970). See also Hooker (1981, Part II).

[10] The "abundant" conception of properties, however, turns out to be one that Kim later rejects: "...I am advocating here a "sparse" conception of properties as distinguished from the "latitudinarian" or "abundant" conception. An extreme form of the latter would regard every predicate as denoting, or representing, a property, with synonymy or logical equivalence taken as the condition under which two predicates denote the same property" (p. 105). Why then insist that the predicates biconditionally related in logically contingent bridge laws must denote distinct properties?

[11] It's comforting to know that this isn't just my problem. Here is Donald Davidson (1987, p. 451): "It is possible to insist that 'water' doesn't apply just to stuff with the same molecular structure as water but also to stuff enough like water to be odorless, potable, to support swimming and sailing, etc. I realize that this remark ... may show that I don't know a rigid designator when I see one. I don't".

[12] It is now commonplace to distinguish between "retentive" reductions and "eliminative" reductions, as two extremes along a spectrum. Although Nagel likely intended the reduction of thermodynamics to mechanics to illustrate a case of reduction pretty close to the retentive end of the spectrum, he may have been mistaken in doing so. As Hooker (1981, p. 49) claims, "[i]n a fairly strong sense, thermodynamics is simply conceptually and empirically wrong and must be replaced. The evidence for this is the fact that thermodynamical concepts and laws can only be exactly reconstructed in statistical mechanics under very peculiar, empirically unrealized, limiting conditions, and that otherwise at least some thermodynamical concepts have so far proved not even roughly applicable (e.g., phase transitions)". However, as I will presently indicate, Nagel was fully aware of the idealizing assumptions and boundary conditions under which thermodynamics can be shown to be reducible (1961, pp. 353ff, 360ff.). Anyhow, if the reduction of thermodynamics is an inappropriate example of the kind of "retentive" reduction that Nagel's model was intended to capture, there is no reason why Nagel could not have chosen a more appropriate one – e.g., the reduction of physical optics to electromagnetic theory, which Hooker places closer to the retentive end of the spectrum (as does Sklar, 1967, pp. 117–118).

[13] Cf. Schaffner (1967), Hooker, (1981), Bickle (undated).

[14] In Nagel's simplified exposition, p (as well as V) occur in both L and L^*. A more precise account would have the mechanical counterpart of the Boyle-Charles' law be $MV = 2E/3$, where M (the mechanical counterpart of p) is the average momenta of the gas molecules per unit area. See note 26 below.

[15] Indeed, as Ager and Aronson (1974) pointed out long ago, *BL* is formally deducible from L and L^*. This, however, as an anonymous referee has pointed out, may only be possible when the two theories share a relevant part of the vocabulary, as in the case under consideration. In other cases a *BL* is posited as a contingent hypothesis which, as Cliff Hooker has pointed out to me in conversation, can be as nontrivial and explanatory as any other contingent hypothesis. (See also note 16.)

[16] This is not to deny that once a bridge law has been postulated at a given stage in the reduction – e.g., in the derivation of the Boyle-Charles' law – it may serve a heuristic purpose in suggesting further intertheoretic connections and testable hypotheses about the derivability of additional laws, thus indirectly contributing to the augmentation, and even correction, of the body of currently accepted laws. (Cf. Nagel, 1961, p. 360; Schaffner, 1967.)

[17] It is interesting to note that Bickle (undated) characterizes Schaffner's and Hooker's accounts of reduction as fundamentally Nagelian (the section in which their positions are discussed is entitled: "Nagel's Insights Revised and Modified"); yet Hooker's account makes no use of bridge laws (except for such *reflexive* "bridge-laws" as those connecting, e.g., V in the Boyle-Charles' law with V in the corresponding mechanichal law), and on Schaeffner's account it's not the T-laws themselves that are deduced from T^*, but only *approximations* to such laws. Anyhow, if such "revised and modified" accounts of Nagel's classic model should be deemed not to deserve the label 'Nagelian', on account of their not requiring the derivation of the *actual* T-laws, let them be re-labelled '*Nagelian'. I would be happy to so re-label my own account, insofar as it does not regard such derivations as *central* to the reduction.

[18] Kim's account of functional reduction is clearly in the spirit of the Armstrong-Lewis version of "analytical functionalism" and psychophysical reduction. Kim acknowledges this link (p. 132, note 10).

[19] There is, however, an important difference between the two kinds of identification. On the Nagelian model, identifications are posited as best explanations of certain nomic regularities (e.g., of the fact that an ideal gas obeys the Boyle-Charles' law iff it obeys the corresponding mechanical law). On Kim's model, identifications are entailed by the very definition of what it is to have a functionalizable property M: having $M =_{\text{df}}$ having some property or other, P, that occupies the causal role definitive of M. This difference, however, concerns the *grounds* for the identification, not the *consequences* of the identification for the explanatory and ontological-simplificatory aims of the reduction. Compare, on this point, Lewis 1972, p. 207)

[20] This paragraph and the next two extend and revise some points of criticisms that I presented in my critical notice of Kim's book (Marras, 2000).

[21] Like Lewis (1980), Kim takes functional terms to be nonrigid designators: "To functionalize M is to make M nonrigid. ... [For] M is defined in terms of its causal/nomic relations to other properties, and since these relations are contingent, it is a contingent fact whether a given property satisfies the causal/nomic specification that is definitive of M" (Kim, p. 99).

[22] The threat of eliminativism with respect to such functional states as "pain-as-such" is acknowledged by Kim (1992, p. 334). See also Horgan (1996).

[23] Kim's way of dealing with multiple realization is as kind-restricted property identity is similar to Lewis' (1980): e.g., to say that pain is realized by C-fiber firings in humans and by Φ in Martians is to say that human pain = C-fibre firing and Martian pain = Φ. This is a consequence of treating 'pain' as a nonrigid designator. See Block (1990) for objections to this account.

[24] D. Lewis's primary allegiance to a similar, kind-restricted form of reductive materialism about mind leads him to question the ostensibly functional aspect of his position: "I am a realist and a reductive materialist about mind. I hold that mental states are contingently identical to physical – in particular, neural – states .... In view of how the term is con-

tested, I don't know whether I am a 'functionalist' (Lewis, 1994, p. 412). Kim may well acknowledge as much about his position.

25 More precisely, there must be a L^* law corresponding to each L-law in which M_i occurs.

26 As remarked, sometimes L and L^* share some of the properties they interrelate (as in "under conditions C tissue damage causes pain etc". and "under conditions C tissue damage causes C-fiber firings"; analogously for $pV = kT$ and $pV = 2E/3$). Furthermore, since P_i in L^* fills the role of M_i in L only if each property P in L^* fills the role of some property M in L, and since for Kim role fillers get identified with the properties whose role they fill, we end up identifying each property in L^* with some property in L, and conversly. (Thus, for example, the mechanical law corresponding to the Boyle-Charles law could be represented as $MV = 2E/3$, where M is the average momenta of the molecules per unit area – the mechanical equivalent of p (pressure) in the Boyle-Charles' law.) $2E/3$ fills the role of kT only if M fills the role of p; and just as $2E/3$ gets identified with kT, so too M gets identified with p. This role isomorphism corresponds to the "dynamical pattern isomorphism" developed by Hooker (1981, Part III) as a criterion for theoretical identification.

27 Terence Horgan (1996) takes this more radical form of multiple realizability to be irreconcilable with kind-restricted reductive identifications.

28 As I pointed out in Marras (1993, pp. 290–291), the kind-restricting strategy, if unconstrained, would have the consequence that any arbitrary domain of individuals at t would constitue a reductive domain for a given psychological property M as long as the domain is specified by some property P which happens to realize M for those individuals at t: just call that domain S, and, unsurprisingly, relative to S, $M \leftrightarrow P$ holds. Surely reductions can't come that cheap.

29 This points to a conception of reduction that corresponds essentially to the one that Fodor (1974) endorses for the reduction of psychology and which he opposes to the "classic" conception commonly associated with Nagel: "The point of reduction is *not* primarily to find some natural kind predicate of physics coextensive with each kind predicate of a special science. It is, rather, to explicate the physical mechanisms whereby events conform to the laws of the special sciences" (Fodor, 1974, p. 127). The physical mechanisms are detailed by the lower level physiological laws on which the special science laws supervene. In accordance with multiple realizability, there may be a plurality of such implementing mechanisms within each species or structure-type.

30 Richardson acknowledges this point in a later paper (1982, p. 126). It should be noted that *other* one-way bridge laws may, of course, enable the deduction: e.g., $M' \to P'$ taken together with $P'' \to M''$ But while enabling the deduction, these one-way bridge laws do not licence reductive, type-identifications.

31 See preceding note.

32 These base laws, as previously noted, are the ones that, as Fodor put it, "explicate the physical mechanisms whereby events conform to the laws of the special sciences" (Fodor, 1974, p. 127).

REFERENCES

Ager T. A. and Aronson, J. L.: 1974, 'Are Bridge Laws Really Necessary?', *Noûs* **8**, 119–134.

Beckermann, Ansgar: 1992, ‘Supervenience, Emergence, and Reduction’, in A. Beckermann (ed.), *Emergence or Reduction?*, De Gruyter, Berlin and New York.
Bickle, John (undated), ‘Concepts of Intertheoretic Reduction in Contemporary Philosophy of Mind’, in M. Nani and M. Marraffa (eds.), *A Field Guide to the Philosophy of Mind*, published electronically (http://www.uniroma3it/Kant/field/cir.htm).
Block, Ned: 1990, ‘Can The Mind Change the World?’, in G. Boolos (ed.), *Meaning and Method*, Cambridge University Press, Cambridge, MA.
Brook, D. H. M.: 1994, ‘How to Perform a Reduction’, *Philosophy and Phenomenological Research* **4**, 803–814.
Causey, R. L.: 1972, ‘Attribute-Identities in Microreductions’, *Journal of Philosophy* **69**, 407–422.
Cummins, Robert: 1983, *The Nature of Psychological Explanation*, MIT Press, Cambridge, MA.
Davidson, Donald: 1987, ‘Knowing One’s Own Mind’, *Proceedings and Addresses of the American Philosophical Association* **61**, 441–458.
Fodor, Jerry: 1974, ‘Special Sciences, or the Disunity of Science as a Working Hypothesis’, reprinted in N. Block (ed.): 1980, *Readings in Philosophy of Psychology*, Vol. 1, Harvard University Press, Cambridge, MA.
Horgan, Terence: 1996, ‘Kim on the Mind-Body Problem’, *British Journal for the Philosophy of Science* **47**, 579–607.
Hooker, C. A.: 1981, ‘Towards a General Theory of Reduction’, *Dialogue* **20**, 38–60, 201–236, 496–629.
Kim, Jaegwon: 1984, ‘Concepts of Supervenience’, in Kim (1993a).
Kim, Jaegwon: 1989, ‘The Myth of Nonreductive Materialism’, in Kim (1993a).
Kim, Jaegwon: 1990, ‘The Philosophical Concept of Supervenience’, in Kim (1993a).
Kim, Jaegwon: 1992, ‘Multiple Realization and the Metaphysics of Reduction’, in Kim, J. (1993a).
Kim, Jaegwon: 1993a, *Supervenience and Mind*, Cambridge University Press, Cambridge, MA.
Kim, Jaegwon: 1993b, ‘Can Supervenience and “Non-Strict Laws” Save Anomalous Monism’, in Kim (1993a).
Kim, Jaegwon: 1998, *Mind in a Physical World*, MIT Press, Cambridge, MA
Lewis, David: 1972, ‘Psychophysical and Theoretical Identifications’, *Australasia Journal of Philosoph* **50**, 249–258.
Lewis, David: 1980, ‘Mad Pain and Martian Pain’, in N. Block (ed.), *Readings in Philosophy of Psychology*, Vol. 1, Harvard University Press, Cambridge, MA.
Lewis, David: 1994, ‘Lewis, David: Reduction of Mind’, in S. Guttenplan (ed.), *A Companion to the Philosophy of Mind*, Blackwell, Oxford and Cambridge, MA.
Marras, Ausonio: 1993, ‘Psychophysical Supervenience and Nonreductive Materialism’, *Synthese* **95**, 275–304.
Marras, Ausonio: 2000, ‘Critical Notice of Kim (1998)’, *Canadian Journal of Philosophy* **20**, 137–160.
Nagel, Ernest: 1961, *The Structure of Science*, Harcourt, Brace & World, New York.
Putnam, Hilary: 1970, ‘On Properties’, in *Philosophical Papers*, Vol. 1, Cambridge University Press, Cambridge, MA.
Richardson, Robert: 1979, ‘Functionalism and Reductionism’, *Philosophy of Science* **46**, 533–558.
Richardson, Robert: 1982, ‘How *Not* to Reduce a Functional Psychology’, *Philosophy of Science* **49**, 125–137.

Schaffner, K.: 1967, 'Approaches to Reduction', *Philosophy of Science* **34**, 137–147
Sklar, Lawrence: 1967, 'Types of Inter-Theoretic Reduction', *British Journal for the Philosophy of Science* **18**, 109–124.

Department of Philosophy
University of Western Ontario
London, Ontario
Canada N6A 3K7