**Multiple Realization, Reduction and Mental Properties**

in *International Studies in the Philosophy of Science* 13 (1999), No. 2, pp. 135-149.

**Abstract.**

The paper tries to remove some obstacles standing in the way of considering mental properties as both genuine natural kinds and causally efficacious rather than epiphenomena. As the case of temperature shows, it isn't justified to conclude from a property's being multiply realizable to its being irreducible. Yet Kim's (1992) argument to the effect that if a property is multiply realizable with a heterogeneous reduction base then it can't be a natural kind and possesses only derivative "epiphenomenal" causal efficacy isn't conclusive either. The fact that temperature is, but jade is not, a natural kind cannot be established by comparing the heterogeneity of their respective reduction bases, but rather by the fact that the former is and the latter isn't embedded in laws of nature.

**Multiple Realization, Reduction and Mental Properties**

There has been a broad consensus in the philosophy of mind regarding the functionalist conception of the nature of mental states. According to functionalism, a token mental state belongs to a given type because it is causally related, or potentially causally related, to other mental states and to sensory input and behavioral output. When this idea was first proposed by Putnam (1967), the major competing theoretical hypothesis about the mind was that mental state types should turn out to be reducible to brain state types. Functionalism won the assent of the philosophical community mainly by Putnam's now famous argument that such a reduction is impossible because mental states of the same type can be biologically or mechanically realized in many ways. According to Putnam himself, this is an empirical possibility; as a result - what came later to be called - the "argument from Multiple Realizability (MR)[1]" appeared to stand on an empirical premise. Since Lewis (1969), however, this argument has been reinterpreted: it is now considered to be purely conceptual. Our concept of a mental state implies that its physical realization does not condition its identity. Materialism requires each mental state to be physically realized in some way or another, but *a priori* the possible realizers form an open disjunction.

This way of putting things was responsible for some intriguing features of the functionalist consensus. On the one hand, the argument from MR linked functionalism with antireductionism. If the possible realizers of a given mental state type are many, one can't reduce it to any single type of realizer. But on the other hand, the concept of realization deprives the (realized) mental state not only of independent existence (this much, every materialist will concede anyway), but also of independent causal powers. To speak of a mental state M as being *realized*, in a given cognitive agent and in a particular situation, by a token of a physical state P, means that P is where we have to look if we want to know about M's causal powers.

The aim of this paper is to examine whether the fact that mental properties are multiply realizable warrants the conclusion that they are not natural kinds. Such a result would have far-reaching consequences for if mental properties are not natural kinds then they cannot constitute the subject matter of a genuine science. Kim (1992) has argued along these lines against the scientific character of general psychology. Considering his arguments will constitute my starting point for

tackling the issue of the status of multiply realizable properties. It will soon become clear that assessing them presupposes a critical examination of the role of theoretical reduction of higher-level by lower-level theories. Eventually both the traditional argument from MR to the irreducibility of the mental and Kim's argument from MR to the thesis that the mental doesn't constitute the subject matter of a genuine science will be found wanting.

## 1. Functionalism, realization, and causal powers

The conception according to which a mental state is "realized" by a neural state in the brain contains the idea that the "realizer" monopolizes causal power (Cf. Kim 1984; 1993a, p. 366) at the expense of the realized mental state. Now, there is a considerable tension between this implication of the concept of realization and the original functionalist conception of a mental state (Cf. Block 1990). The question of where the causal powers lie is essential for functionalism: it is based on the postulate that the identity of a mental state M is determined by and only by its causal interactions (including its potential causal interactions) with other mental states (and with sensory input and behavioral output), but not by the properties of the brain in which these mental states occur. Functionalism thus crucially depends on the causal efficacy of the mental states themselves as opposed to the brain states that realize them[2]. But it is difficult to see how functionalism could resist the following argument leading to the opposite conclusion: if a mental state has a realizer which is a neural or other biologically or physically specified state, then this realizing state has causal powers of its own. Now attributing causal power both to realizer and realized mental state makes it look as if there were systematically causal overdetermination each time a mental state acts as a cause of the same effect as the realizer causes. Kim proposes to avoid this apparent overdetermination by appealing to a principle of explanatory exclusion (Kim 1988; 1989b): of a given effect there cannot be two complete but independent causal explanations. But where a mental property and the neural property realizing it compete for causal power, the neural property wins. Kim reconciles this with the apparent causal efficacy of mental states by way of the thesis that mental causation - as indeed all other higher-level causation - is "epiphenomenal" or "supervenient" causation (Kim 1984). Kim explicates this relation as follows:

"The general schema for reducing a macrocausal relation between two events, x's having F and y's having G, where F and G are macroproperties, is this: x's having F supervenes on x's having m(F), y's having G supervenes on y's having m(G), where m(F) and m(G) are microproperties relative to F and G, and there is an appropriate causal connection between x's having m(F) and y's having m(G). Any causal relation conforming to the pattern set forth above will be called a 'supervenient causal relation'." (Kim 1984, p. 262).

In (Kim 1992), Kim does not argue for his theory of supervenient causation, but simply reminds us of its conclusion, namely the "Causal Inheritance Principle" which says that "if mental property *M* is realized in a system at *t* in virtue of physical realization base *P*, the causal powers of *this instance* of *M* are identical with the causal powers of *P*." (Kim 1992, p. 18). From this he draws the conclusion that functionalism is wrong in thinking that it is possible to identify mental states on the basis of their causal interactions. In trying to do so, functionalists in fact just apply a generally valid methodological principle for the identification of natural kinds, a principle which Kim calls the Principle of Causal Individuation of Kinds. It says that "kinds in science are individuated on the basis of causal powers; that is, objects and events fall under a kind, or share in a property, insofar as they have similar causal powers." (Kim 1992, p. 17).

Not the least of Kim's merits is the fact that he has spelled out more explicitly than others that the idea that mental states are (multiply) "realized" by physical states leads to an irrealist position with respect to mental states. The reason is that causation is the best indicator we have to know whether a property is real. But if the Causal Inheritance Principle is correct, the functionalist's attempt will not lead to the specification of any well-determined natural kind. Kim concludes that "mental kinds cannot satisfy the Causal Individuation Principle, and this effectively rules out mental kinds as scientific kinds." (Kim 1992, p. 18)[3].

## 2. Theoretical reduction and property identity

Kim (1992) attacks antireductionism on two grounds. We have seen his first argument: mental kinds are realized by physical kinds; kinds are individuated causally and the causal powers of the realizing property are ontologically prior to the causal powers of the realized state. From these premisses he concludes that "mental kinds are not causal kinds, and hence are disqualified as proper scientific kinds" (Kim 1992, p. 18)[4].

His second argument begins with a careful spelling out of his opponents' - the antireductionists - line of reasoning according to which MR entails the irreducibility of the higher-order multiply realized state. Why is it that MR is said to lead to irreducibility? If M can be realized by a large number of different physical states $N_1$, $N_2$,..., then the reduction base would have to be a "disjunctive property" $N_1 \vee N_2 \vee N_3$ ... with many or indeed infinitely numerous disjuncts[5]. But why is a mental state type irreducible if the only candidate for such a reduction is a disjunction? According to Kim,

> "for genuine reduction, the bridge laws must be construed as *property identities*, not mere *property correlations* [...] This of course requires that each $T_2$-term [i.e. each term of the reduced theory, M.K.] have a nomic (or otherwise suitably modalized) coextension in the vocabulary of the reduction base. To put it in another way, ontologically significant reduction requires the reduction of higher-level *properties*, and this in turn requires (unless one takes an eliminativist stance) that they be identified with complexes of lower level properties. Identity of properties of course requires, at a minimum, an appropriately modalized coextensivity." (Kim 1992, p. 9; italics are Kim's).

This requirement is the first premiss of the argument from MR to irreducibility. It amounts to a very restrictive interpretation of the concept of theoretical reduction. Kim cites Hellman and Thompson's (1975) formulation according to which reductionism is the thesis "that all scientific terms can be given explicit definitions in physical terms" (Hellman and Thompson 1975, p. 551; Kim 1992, p. 2). But instead of talking in terms of the *definition* of *terms*, Kim himself then goes on to formulate his discussion of reduction in terms of *laws* linking higher-level to lower-level *properties*. In the latter - realist - terminology, reduction requires bridge laws linking reduced to reducing properties. This in turn requires that both the reducing and the reduced properties are natural kinds.

Now, the argument from MR precisely denies that this latter requirement can be fulfilled in the case of multiply realizable properties. The denial depends on a second premiss which Kim sets out to explicate and to defend. The potential realizers form a heterogeneous disjunction, but "a disjunction of heterogeneous kinds is not itself a kind." (Kim 1992, p. 9). Kim's main argument for this claim is that predicates expressing a property with a disjunctive realization basis, like "being jade", are not projectible: they are not confirmed by their instances[6].

With these premises the irreducibility of M follows from its being multiply realizable: If M is "realized" in many different ways, the only plausible candidate for reduction is a "non-kind",

namely the disjunction of all these actual and potential realizers. But Kim pursues the entailments of the argument further than its antireductionist proponents:

"If M is identified with non-kind Q (or M is reduced via a biconditional bridge principle 'M $\leftrightarrow$ Q', where Q is a non-kind), M could no longer figure in special science laws; e.g., the law, 'M$\rightarrow$ R', would in effect reduce to 'Q $\rightarrow$ R', and therefore loses its status as a law on account of containing Q, a non-kind." (Kim 1992, p. 10).

This consideration leads to turning the arguments' conclusion against the antireductionists. If being nomically equivalent to a non-kind means to *be itself* a non-kind, then if the only plausible reduction base for M is disjunctive, we can't have both: hold that M can be reduced to a disjunctive property *and* hold that M continues to be a natural kind, figuring in laws on the level of mental properties. Faced with this alternative, instead of choosing to maintain M as a natural kind and abandoning the perspective of its reduction, Kim proposes to make the opposite choice: M is reducible though only locally, i.e. there is a local reduction of M whose validity is restricted to precisely one type of realizer. On the other hand, Kim is prepared to accept the conclusion that M itself can't be considered as a natural kind any more. He is thus led to "a form of mental eliminativism" (1992, p. 25)[7].

The above argument defended both by the antireductionists and their opponents relies crucially on the *identity* of reducing and reduced property. Indeed, only if M and Q are identical, do we get the crucial conclusion that "'M$\rightarrow$ R' would in effect reduce to 'Q $\rightarrow$ R'" (Kim 1992, p. 10). But it seems to me that it is wrong to impose such a strong requirement on the possibility of reduction. If I can show that this conception of what theoretical reduction comes to is misguided, I shall have shed doubt on the soundness both of the argument from MR to irreducibility and of Kim's argument that local reduction leads to the elimination of the reduced kind.

The relevant condition Nagel (1961) imposes on the reduction of a predicate $P_1$ in a higher level theory $T_1$ by a predicate $P_2$ in a lower-level theory $T_2$ is weaker. Indeed, in order to accomplish a successful reduction of $T_1$, $T_2$ must be supplemented with linking postulates establishing a connection between those predicates of $T_1$ which do not appear in $T_2$, and $T_2$. In general $T_2$ will not contain any predicate *synonymous* with $P_1$, e.g., neither "mean molecular energy" nor any other term of statistical mechanics is synonymous with "temperature". Once this possibility is excluded, the linkage could in principle either be the result of a convention which stipulates a coordinating definition of $P_1$ by $P_2$, or the object of a factual hypothesis[8]. What is

crucial for our present argument is that neither of the remaining two interpretations - convention or factual hypothesis - implies that a successful reduction of $P_1$ by $P_2$ shows that both predicates designate one and the same property.

It order not to beg any questions, it may be useful to compare reduction with a non-reductive systematic link between two families of properties where those properties are clearly not identical with each other. Thus, basic colors are systematically correlated to light of definite wavelengths. This link presumably has the status of an empirical psychophysical hypothesis, but it could also be interpreted as a set of coordinating definitions. Yet neither interpretation would imply that, e.g., the secondary property of being yellow is *identical* to the property of possessing a wavelength of 589 nm. Other than electromagnetic waves may have that wavelength without thereby being yellow.

Linking postulates are necessary but not sufficient for successful reduction. Rather, the decisive criterion is the theoretical fruitfulness of the connection achieved by those postulates. One is justified in considering a reduction as successful if it suggests new and interesting generalizations, both in the reducing and in the reduced theory. There are indeed good reasons for not requiring the linking postulates either to be "bridge laws" or to express property identities. On looking closer to paradigmatic examples of successful property reductions in physics, it turns out that they don't have the properties philosophers often think they have. Why should the reduction of the mental to the neural fulfil requirements so rigourous that even a paradigmatic case of successful reduction in physics, such as that of temperature to mean molecular energy, fails to satisfy them?

First, temperature is multiply realizable, but its being multiply realizable is not an obstacle to its reduction. Second, even if one takes into account only the main reduction base of temperature, mean molecular energy, it turns out that there are objects which possess the property in the reduction base - mean molecular energy - without possessing the reduced property - temperature. Coextensivity[9], and *a fortiori* identity of reduced and reducing property, fail[10].

In order not to bring in any controversial issues which are inessential for the point presently at stake, let us focus on gases: these are the paradigmatic substances for which the reduction is valid. Consider an object constituted by only *two* molecules of some gas. This object definitely has mean molecular energy - the mean being taken both over time and the two molecules - but it doesn't have any temperature. The concept of temperature can only be applied to macroscopic objects[11]. Or

consider the artificial object constituted by the sum of a great number of gas molecules, but where these molecules are isolated from each other. For example, take the mereological sum of one molecule out of each of a series of boxes filled with some gas. Again, the resulting object will possess mean molecular energy, but no temperature. This is because the number of elements is not the only restriction placed on the application of thermodynamic concepts like temperature to a given object. The elements constituting the object in question must furthermore form an interacting whole. But in virtue of the constitution of our mereological sum, its elements do not interact with each other; thus the object doesn't have a temperature. Nevertheless, it has mean molecular energy, for the application of *this* concept requires neither a minimal number of elements nor interaction between these elements. This should be sufficient to show that in the paradigmatic case of the reduction of temperature to mean molecular energy, this reduction does not result in the *identity* of the reduced with the reducing property. In fact, these properties aren't even coextensive[12], let alone nomically or necessarily coextensive.

Secondly, the reduction base of temperature is multiple[13]. The concept of temperature has recently been applied both to describe the state of the matter forming the atomic nucleus and to empty space. In the first case, there are no molecules which could possess mean molecular energy; in the second case there aren't any particles at all. In view of the case of nuclear temperature, we can drop the mention of molecules which now appears to reflect the narrow application of the temperature concept before the discovery of the inner structure of the nucleus. If we redefine the reduction base to be simply mean kinetic energy, this property can reduce both nuclear temperature and the temperature of solids, liquids and gases. However, the reduction base so redefined still doesn't cover the application of the concept of temperature to empty space. Indeed, this application isn't based on kinetic energy but on the fact that black bodies[14] have a spectrum of radiation which is specific for each temperature. This perfect mapping allows one to extend the concept of temperature to objects deprived of matter but containing radiation - which is the case for empty space. This means that temperature is, after all, neither reducible to mean molecular energy nor to mean kinetic energy but to mean molecular energy *or* to the black body radiation spectrum.

How does this help us with our problem about knowing whether a mental property with a multiple reduction base can nevertheless be considered as a natural kind, and even an irreducible

natural kind - as antireductionists argue - or whether "as a property or kind [it] must go", for as Kim argues, "to be reduced is to be eliminated as an independent entity" (Kim 1992, p. 24)? I propose to infer from the analogy with temperature that neither of these conclusions is correct.

1. The argument from MR to irreducibility is wrong: Temperature has been reduced[15], pain and other mental states are potentially reducible, all with a multiple reduction base. This means that a disjunctive predicate is on one end of the translation hypothesis which is the core of theoretical reduction. The argument from MR to irreducibility is wrong, because on the basis of the multiplicity of their (actual or potential) reduction base, temperature and mental states are in the same boat. Parity of reasoning prevents us from taking MR as a ground for denying reducibility to mental states in so far as the application of the same reasoning to the case of temperature leads to the manifestly wrong result that temperature is irreducible.

2. Kim is wrong, too: Reduction with a multiple reduction base doesn't entail elimination "as an independent entity". Kim's thesis can be interpreted as a thesis about elimination *as a causally efficacious property*. The extension of the concept of temperature from matter to empty space shows that reduction is not the basis for considering a property as real. As Kim himself points out, to be considered as real - where "real" always means "(potentially) causally efficacious" - a property has to figure in genuine laws[16]. Kim (1992, p. 12) proposes to distinguish the latter by the fact that they are confirmed by their instances, whereas other generalizations are not[17]. This is not the place to enter into the difficult subject of criteria for lawfulness. But it is certainly the case that the success of macroscopic thermodynamics combined with the fact that if that theory were true there would be many objective laws of nature involving temperature, are responsible for our considering temperature as an independent property. "Independent" of the issue of its reducibility, and *a fortiori* of whether it is multiply or simply reducible. We should consider both the reality of a property and its reducibility as empirical issues: to find out about the former, we have to look at scientific inquiry and see if there are any lawful connections between the property in question and other properties on the same ontological level. This criterion of when some property is "real", i.e. constitutive of a natural kind, and when on the contrary it "must go" as a natural kind property, also proves successful in more controversial cases. Kim elaborates the case of jade - which is a kind for common sense and for gemology but which has a dual reduction base in mineralogy, consisting of

jadeite and nephrite. Presumably, neither common sense nor gemology are sciences. They don't identify laws of nature. Being jade is not a natural kind property because no law contains this property, whereas being jadeite and being nephrite are kinds because they are lawfully connected to other properties identified by mineralogy.

Questions about reducibility must be postponed to a time when both the property itself is embedded in a scientific theory and there are candidate theories describing reality on a lower ontological level. Pain (and other mental states) in general - as opposed to pain in a given species - will turn out to be a natural kind if there turns out to be a general psychology providing evidence that there exist laws linking pain (and other mental states) to other mental or non-mental properties which are instantiated by more than one species. This is an empirical issue, and it is independent of the question about the reducibility of these kinds - multiple or not.

### 3. Homogeneous and heterogeneous reduction bases

One might object to this reasoning that it doesn't do justice to the important question as to *how heterogeneous* the set of properties in the reduction base is. In this sense, Kim says that "there is nothing wrong with disjunctive predicates as such; the trouble arises when the kinds denoted by the disjoined predicates are heterogeneous, 'wildly disjunctive', so that instances falling under them do not show the kind of 'similarity', or unity, that we expect of instances falling under a single kind." (Kim 1992, p. 13). There is, it might be argued, a crucial difference between temperature on the one hand and jade and pain (and other mental states) on the other. Both the functionalists' argument from MR to irreducibility and Kim's argument from MR to the elimination of the multiply realized property as a natural kind are motivated by, if not explicitly dependent upon, the supposed heterogeneity of the lower level properties[18]. Thus, the objection might go, temperature is reducible, although reducible by a multiple reduction base, because the properties contained in the reduction base are relatively homogeneous. This is why the functionalist needn't consider temperature as irreducible and why Kim isn't forced to conclude that temperature is not, after all, a natural kind. But, according to the objection, the situation is crucially different with pain and jade: *Their* reduction bases are heterogeneous to a degree higher than some treshold.

In order for this objection to be successful and thus to be able to rescue either the functionalists' or Kim's argument, it would have to be supplemented by a clear-cut criterion for distinguishing homogeneity from heterogeneity in the relevant sense. A plausible criterion seems to be that two properties are homogeneous if and only if they have (important) properties in common, and (totally) heterogeneous, if and only if they have few (or no), or only unimportant, properties in common. It should be uncontroversial that such a distinction must allow for degrees. But how can it be made precise? - sufficiently precise to justify the judgment that the pair jadeite/nephrite is significantly more heterogeneous than the pair mean kinetic energy/specific distribution of radiation (obeying the law of the black-body), and - what is certainly even more difficult - to justify the further judgment that pain is more similar to jade than to temperature in this respect. It is clear that neither pair is *totally* heterogeneous: Jadeite and nephrite have sufficiently many properties in common to make for the very similar macroscopic properties of the crystals of those types. Similarly, mean kinetic energy and the specific distribution of radiation obeying the black body law have something in common which enables objects possessing either one or the other of these properties to reach a common state of thermodynamic equilibrium[19].

Plausibly, though, neither of the common properties shared by the pair in question could be used decisively in the issue at hand, namely that of distinguishing homegeneous from heterogeneous in the context of different types of MR. For both the properties common to jadeite and nephrite underlying the similar superficial properties of the respective crystals, and the property common to mean kinetic energy and the specific distribution of radiation obeying the black body law, enabling the objects possessing them to reach a state of thermodynamic equilibrium, are properties which are characterised by reference to a property which belongs to the *level of the reduced property*. But the fact that they have something in common on *that* level is trivial for we precisely undertook to consider those pairs of properties because they are reducing the same property! It means nothing else than that they share the property of belonging to the same reduction base.

Kim proposes a criterion of confirmability: A general statement with an antecedent denoting a "wildly disjunctive" property is not in general confirmed by its instances. Thus, according to Kim, we would not be justified in considering

(1) "Jade is green"

as confirmed by its positive instances if all samples ever considered turned out to have been jadeite (Cf. Kim 1992, p. 12). In that case, he argues, all those jadeite samples do not licence any conclusion about the properties of nephrite and thus don't in the end confirm (1) either.

I think one should resist that conclusion. The samples of green jadeite *do* confirm the statement (1), for the green appearance is precisely one of the properties which are responsible for the grouping together of jadeite and nephrite under the higher-level property of being jade. To make his point, Kim would have to find a property belonging exclusively to jadeite and not to nephrite, say containing atoms of aluminium (Al). The jadeite samples still confirm the general statement

(2) "Jade contains Al",

but this time, the confirmation is deceptive. (2) is refuted by the first observed sample of nephrite which does not contain Al. However, its being deceptive is perfectly compatible with its being a confirmation. We know that induction is logically invalid; confirmation is inductive and can therefore never provide a proof of the confirmed general statement. Many of the best confirmed statements of laws of nature have turned out to be false. Nevertheless they had been confirmed, until cases refuting them were discovered. Nothing else happens when the (hypothetical) generalization (2) is refuted by a sample of nephrite, i.e. a sample of jade not containing Al. The reasonable move is then to modify statement (2) and choose as a substitute a more modest conjecture:

(3) "Jadeite contains Al",

as long as *it* is not discovered to be false in turn. The case of temperature is analogous. Before discovering that empty space has temperature as well, one might have conjectured that

(4) "Everything which has temperature has moving parts".

Such a putative law[20] would have been regularly confirmed by all material samples[21]. However, it was refuted by the discovery that regions of space devoid of matter also had temperature. As with jade, the correct move is to modify the statement by reducing the width of its scope:

(5) "Everything material which has temperature has moving parts".

The important point to note is that the refutation of a general statement like (2) or (4), i.e. the discovery that the preceding confirmation was deceptive, does not automatically entail that the property in question was not after all a natural kind. Temperature is and remains a natural kind even though there are generalizations involving temperature - like (4) - which have been refuted. By parity of reasoning, Kim would have to say that it then turned out that (4) was "not confirmed by its instances", for the conclusions one can draw about the properties of empty space on the basis of the observation of samples of matter with temperature are just as limited as the conclusions one can draw about samples of nephrite on the basis of the observation of samples of jadeite. But unlike jade, temperature's having a disjunctive reduction base doesn't threaten its being a natural kind. The reason is that there are *other* laws involving temperature which we are well justified to believe, such as the 0th law of thermodynamics, or the Boyle-Mariotte law for ideal gases: $pV = nRT$. It suffices that there exists one law implying a property in order for this property to be a natural kind. If, on the other hand, we don't take jade to be a natural kind, this is not because, as Kim suggests, there are general statements - like (2) - involving the predicate "jade" and whose confirmation by its instances was deceptive, but because there is no law at all which we have good reasons to believe and which involves the property of being jade.

I conclude that the capacity of a general statement to be confirmed by its instances does not provide a criterion for whether the predicates linked by that statement denote natural kind properties or not. In the end, the criterion Kim proposes doesn't allow one to make precise the intuitive distinction between predicates denoting "wildly disjunctive" properties and predicates denoting genuine natural kinds. In the form Kim gives it, parity of reasoning forces him once more to rule out temperature together with jade from the realm of genuine natural kinds. This consequence suffices to undermine the credibility of Kim's criterion.

It seems to me that the burden of proof lies with the advocate of the homogenous/heterogeneous distinction: In order to block my argument by invalidating the analogy with temperature, both the functionalist and Kim must either spell out a clear-cut criterion of what counts as being sufficiently heterogeneous for membership in the same category as jade and what counts as being sufficiently homogeneous for membership in the same category as temperature, or at least show that such a criterion could be found in principle. In the absence of such a criterion, the

analogy stands and with it my argument against the functionalists' irreducibility thesis and Kim's elimination thesis.

### 4. Local reduction and eliminativism

Before finishing I would like to point to an unwelcome consequence of Kim's argument for the "conclusion that pain as a property or kind must go" (Kim, 1992, p. 24). If it were correct, it would lead to radical eliminativism, Kim's explicit denial nonwithstanding. According to Kim, "the present approach is not, in its ontological implications, a form of the standard mental eliminativism currently on the scene" (Kim 1992, p. 25). I shall briefly try to show that Kim's argument, if it were correct, couldn't both show that general pain *is not* a natural kind *and* that its species-specific substituents *are* natural kinds. By parity of reasoning, all the local species-specific restrictions of our usual general psychological properties would fall prey to analogous arguments from reduction by a disjunctive reduction base.

According to Kim, whenever a property has a disjunctive reduction base, neither the reducing disjunctive property nor the reduced property are genuine kinds. "Local reduction" permits one to show where the genuine kinds are. In the case of mental properties like pain, the genuine properties are species-specific. "The present view allows, and in fact encourages, 'species-specific psychologies', but the standard eliminativism would do away with all things psychological - species-specific psychologies as well as global psychology." (Kim 1992, p. 26). In an earlier text, Kim (1989a) himselfs hints at a general difficulty for the intended substitution of species-specific psychologies for general psychology. He considers the possibility that science may discover species-specific bridge laws reducing mental properties separately for each biological species. These laws would have the form "$Si \rightarrow (M \leftrightarrow Pi)$" (Kim 1989a, p. 273). For each particular species Si, there is some physical property Pi whose instantiation by an individual of this species is both necessary and sufficient for its instantiation of the mental property M. Kim pauses to note that

> "in order to generate laws of this kind, biological species may turn out to be too wide; individual differences in the localization of psychological functions in the brain are well known. Moreover, given the phenomena of learning and maturation, injuries to the brain etc., the neural structure that subserves a psychological state or function may change for an individual over its lifetime" (Kim 1989a, p. 273)[22].

Kim doesn't seem to appreciate the fact that there is an enormous difference between a linking proposition - I hesitate to call it a law - establishing a nomic link between a mental state type and the neural state of *one* individual at *one* moment of his lifetime on the one hand, and on the other hand a genuine bridge law between that mental state type and a general neural state type, i.e. a neural state which could be instantiated in all (or at least many) individuals of the species, and more than only once in their lifetime. On the contrary, Kim immediately continues by playing down this difference. "What is important, he says, is that these laws are relative to physical-biological structure-types, although for simplicity I will continue to put the matter in terms of species" (Kim, 1989a, p. 273). But the thesis that for a given species, there exists a non-disjunctive reduction base for mental properties, is far from being a negligible detail in Kim's argument which could be set aside for simplicity's sake. On the contrary, it is the centerpiece of his response to Putnam's (1967) argument according to which MR blocks all hopes for reduction of mental state types. If it really turned out to be the case that the neural reduction base for a given mental state M is different not only for each species, but for each individual and even for each moment in the lifetime of this individual, this would according to the notion of reduction used by Kim indeed make even "local", i.e. species-specific, reduction impossible.

We have good reasons to expect that the properties which are at the basis of a species-specific reduction turn out to be themselves multiply reducible. Let M be reducible, in humans, to the neural property P. Then due to individual differences and to the evolution of each individual, P is likely to be itself reducible to a disjunction of complex chemical and physical properties of the brain. If Kim's argument concerning M were sound, he couldn't help applying it to P as well. The neural state P would be eliminated as a natural kind for exactly the same reason for which M has been eliminated as a natural kind. His argument according to which either both the reduced property and the reduction base are natural kinds or neither of them is, naturally drives him toward the eliminativist position according to which only fundamental physics individuates objective natural kinds. Kim succeeds in avoiding eliminativism only by drawing the opposite conclusion in two arguments whose logical structure is strictly the same. Pain in general *is not* a respectable natural kind *because* it is multiply realizable on the neural level, but human pain *is* a respectable natural kind *although* it is multiply realizable on the chemical and physical level. Kim's conclusion that

general psychological properties are eliminated by local reduction to species-specific psychological properties, where these latter conserve their status as natural kinds, is based on an *ad hoc* distinction.

## 5. Conclusion

The aim of this paper has been to remove some important obstacles standing in the way of considering mental (and other higher-level) properties as constitutive of genuine natural kinds. I argued for the thesis that the causal efficacy of macroscopic properties (which goes hand in hand with their being natural kinds and their being embedded in laws of nature) is compatible with, though logically independent of, their reducibility to underlying properties. In order to show this, I analyzed an argument of Kim's (1992) intending to show that multiply realizable macroproperties are not genuine natural kinds - and thus possess only derivative "epiphenomenal" causal efficacy. My strategy for refuting this claim was to show that temperature is in relevant respects similar to more controversial macroproperties. In particular, it is multiply realizable and it is neither identical nor coextensive with any one of the properties in its supervenience base, e.g. mean kinetic energy. Nevertheless temperature is reducible if anything is.

Against antireductionism, this shows that - by parity of reasoning - the fact that a macroproperty is multiply realizable and neither identical nor even coextensive with any of the properties in its supervenience base is not an obstacle to its reduction. But against Kim, the example of temperature shows that its being in that sense a typical macroproperty isn't an obstacle to its being a genuine natural kind either, and therefore to its being genuinely causally efficacious. A property's being involved in laws of nature is a better criterion for both naturalness and causal efficacy. The fact that temperature is part of the Boyle-Mariotte law suffices to establish its being a natural property and causally efficacious.

There is an important objection against the legitimacy of the analogy between temperature and mental states which is crucial to my argument. It might be said that this analogy overlooks the distinction between multiply realizable properties whose supervenience bases are respectively homogeneous and heterogeneous. This is not a clear-cut distinction, but if the objection is to go

through, it should at least be plausible that a mental state like pain comes closer with respect to the homogeneity of its supervenience base to a property like jade which is *not* a natural kind than to a property like temperature which is one. I tried to show, however, that neither the homogeneity or heterogeneity of the supervenience base of a property P nor the confirmability of general statements containing a predicate denoting that property can serve as a criterion for deciding whether P is a natural kind or not. My argument was based on a comparison of jade with temperature with respect to these criteria. Contrary to what the objection says, these properties turn out to be similar with respect to the heterogeneity of their respective supervenience bases, and this similarity is also brought out by applying the criterion of confirmability. Therefore, these are not satisfactory criteria for establishing that temperature is a natural kind property whereas jade is not. But if these criteria fail for jade and temperature, we cannot expect them to tell us whether *pain* is a natural kind property or not.

I ended by pointing out an additional problem for Kim's position. Let us assume that there are, or will be, species-specific reductions of mental properties to properties of the brain. As Kim recognizes, we can also assume that the neurophysiological reducing property will in general itself be instantiated in different ways, depending on the individual and the moment in his lifetime. If one accepts his thesis that a property whose reduction base is heterogeneous can't be a natural kind, and if this criterion is sufficient for deciding about naturalness, then one isn't justified either in considering the neurophysiological property as a natural kind in its turn. Kim's line of reasoning leads to eliminativism, recognizing only the properties of fundamental physics as genuinely natural and causally efficacious[23].

**References**

BLACKBURN, S. (1993) Losing Your Mind: Physics, Identity, and Folk Burglar Prevention, in: *Essays in Quasi-Realism* (Oxford, Oxford University Press).

BLOCK N. (ed.) (1980), *Readings in the Philosophy of Psychology, Vol. 1* (Cambridge, MA, Harvard University Press).

BLOCK, N. (1990) Can the Mind Change the World?, in: G. Boolos (ed.), *Meaning and Method, Essays in Honour of Hilary Putnam* (Cambridge, Cambridge University Press).

BLOCK, N. (1997) Anti-Reductionism Slaps Back, in: James E. Tomberlin (ed.), *Philosophical Perspectives, 11: Mind, Causation, and World*, pp. 107-132.

CHURCHLAND, P.M. (1988) *Matter and Consciousness: A Contemporary Introduction to the Philosophy of Mind*. Revised Edition (Cambridge, MA, MIT Press).

CHURCHLAND, P.S. (1982) Review of: Austen Clark, Psychological Models and Neural Mechanisms, *Journal of Philosophy,* 79, pp. 98-111.

CHURCHLAND, P.S. (1986) *Neurophilosophy* (Cambridge, MA, MIT Press).

CRANE, T. (1992) Mental Causation and Mental Reality. *Proceedings of the Aristotelian Society*, 92, pp. 185-202.

HELLMAN, G.P. AND THOMPSON, F.W. (1975) Physicalism: Ontology, Determination, and Reduction. *Journal of Philosophy,* 72, pp. 551-564.

JACKSON, F. AND PETTIT, P. (1988) Functionalism and Broad Content. *Mind,* 97, pp. 381-400.

JACKSON, F. AND PETTIT, P. (1990) Causation in the Philosophy of Mind. *Philosophy and Phenomenological Research,* 50, Supplement, pp. 195-214.

KIM, J. (1984) Epiphenomenal and Supervenient Causation. *Midwest Studies in Philosophy,* IX, pp. 257-270; repr. in: Kim (1993b), chap. 6.

KIM, J. (1988) Explanatory Realism, Causal Realism, and Explanatory Exclusion. *Midwest Studies in Philosophy,* XII, pp. 225-239.

KIM, J. (1989a) The myth of nonreductive materialism, in: J. Kim (1993b), chap. 14.

KIM, J. (1989b) Mechanism, Purpose and Explanatory Exclusion, in: James E. Tomberlin (ed.), *Philosophical Perspectives, 3: Philosophy of Mind and Action Theory*, pp. 78-108; repr. in: Kim (1993b), chap. 13.

KIM, J. (1992) Multiple Realization and the Metaphysics of Reduction. *Philosophy and Phenomenological Research,* 52, pp. 1-26; repr. in: Kim (1993b), chap. 16.

KIM, J. (1993a) Postscripts on mental causation, in: Kim J. (1993b), chap. 18.

KIM, J. (1993b) *Supervenience and mind* (Cambridge, Cambridge University Press).

KIM, J. (1997) The Mind-Body Problem: Taking Stock after Forty Years, in: James E. Tomberlin (ed.), *Philosophical Perspectives, 11: Mind, Causation, and World*, pp. 185-207.

KISTLER, M. (forthcoming), Causes as events and facts, *Dialectica*.

LEWIS, D. (1969) Review of Putnam, in: Block (1980).

LEWIS, D. (1972) Psychophysical and Theoretical Identifications, repr. in: Rosenthal D. (ed.), *The Nature of Mind* (Oxford, Oxford University Press, 1991); also repr. in: Block (1980).

NAGEL, E. (1961) *The Structure of Science* (London, Routledge and Kegan Paul).

PACHERIE, E. (1993) *Naturaliser l'intentionnalité*, (Paris, Presses Universitaires de France).

PEREBOOM, D. AND KORNBLITH, H. (1991) The Metaphysics of Irreducibility, *Philosophical Studies,* 63, pp. 125-145.

PROUST, J. (1995) Functionalism and Multirealizability. On Interaction Between Structure and Function, in: K. Gavroglu et al. (eds.), *Science, Mind and Art* (Dordrecht, Kluwer), pp. 169-185.

PROUST, J. (1997) *Comment l'esprit vient aux bêtes* (Paris, Gallimard).

PUTNAM, H. (1967) The Nature of Mental States, in: Rosenthal D. (ed.), *Materialism and the Mind-Body-Problem* (Englewood Cliffs, NJ, Prentice Hall, 1971); also in: *Philosophical Papers, Vol. 2* (Cambridge, Cambridge University Press, 1975).

## Note on contributor

Max Kistler is Lecturer in the Department of Philosophy, Université Blaise Pascal of Clermont-Ferrand, France. He earned his Ph.D. in philosophy from the Ecole des Hautes Etudes en Sciences Sociales (E.H.E.S.S.) in Paris, with a thesis on "Causality, law, and representation" in 1995.

*Correspondence*: Département de philosophie, Université Blaise Pascal, 29 Boulevard Gergovia, 63037 Clermont-Ferrand, France

---

[1]I shall use "MR" to refer the *phenomenon* of multiple realizability, not - as for example Kim (1992) does - to the *thesis* that mental states have the property of being multiply realizable.

[2]It should be mentioned that the opposition between a mental state and its realizer is not inevitable for functionalism as such. David Lewis' (1972), e.g., formulates functionalism as a version of the identity theory which identifies brain states and mental states rather than opposing them. The opposition between a mental state and its realizer exists only for those who take it that MR forbids their identification insofar as the different realizers are not identical to each other.

[3]Kim's way in (1984a) of analyzing mental causation as "supervenient causation" or "epiphenomenal causation" is misleading in this respect : while attributing causal efficacy exclusively to the physical realizers of mental states, and thus denying independent causal efficacy to the mental states themselves, he continues to *call* them causes and effects. More recently, Kim seems to have admitted this fact : he now wonders whether "'supervenient causation' [...] might [...] not be 'causation' in name only" (Kim 1993a, p. 359).

[4]Since 1992, Kim has modified his judgment about the scientific character of properties which are, like mental properties, multiply realizable. In (Kim 1997), he argues for the compatibility of functionalism with the reduction of the mental to the physical, and even for the thesis "that the functionalist conception of mental properties is *required* for mind-body reduction" (Kim 1997, p. 203). But this modification doesn't touch the crucial point of the present paper which concerns the reality of the multiply realized property. In the sense intended here, a property is "real", or is constitutive of a natural kind, if it is capable of being causally efficacious. According to nomological conceptions of causation, this capacity in turn depends on the property's being embedded in laws of nature. But Kim still holds - in line with the position expressed in (Kim 1992) - that if a (mental or other) property M is reducible with a multiple reduction base, then "there is no need [...] to think of M itself as a property in its own right" (Kim 1997, p. 201). The reason is that the scientific respectability of properties which are multiply reducible goes hand in hand with their functional definition as "second order-properties" (ibid., p. 194). But Kim makes it clear that such talk of "second-order properties" is misleading, for contrary to what it suggests, "by existential quantification over a given set of properties, we do not literally bring into being a new set of properties" (ibid., p. 200). What remains in the place of properties which have been reduced (and thus, as Kim might still say, "eliminated as independent entities"), are "second-order *descriptions* or *designators* of properties, or second-order *concepts*, [rather] than second-order properties" (ibid., p. 201).

[5]The former is the case if multiple realization is an empirical fact : the reduction base is the disjunction of the actually existing realizers. The latter is the case if MR is conceptual : there is an infinite number of possible realizers for any given mental state, thus the potential reduction base is an infinite disjunction.

[6]I postpone the examination of this argument until section 3 below. In this section I shall rather concentrate on the more general question whether reduction between properties implies (or presupposes) their identification.

[7]I shall come back to this conclusion later, in section 4.

[8]In the present context, we can leave aside the question which of the remaining interpretations of the character of the nature of the linking postulates is preferable. As Nagel (1961, pp. 356f.) points out, they are not, appearences to the contrary nonwithstanding, incompatible alternatives. It may depend on the context of exposition of the theoretical reduction as a whole whether the link between $P_1$ and $P_2$ appears as a coordinating definition or as a factual hypothesis.

[9]I follow a common contemporary usage - which is also Kim's own - in talking not only of a *predicate's* extension which is the class of objects satisfying that predicate, but also of a *property's* extension which is the class of objects possessing (or instantiating) that property.

[10]If this is true, the reduction can be accomplished neither by nomic equivalence nor by a law - at least on accounts which consider laws to imply necessary, or at least universal coinstantiation of the lawfully linked properties.

[11]"Macroscopic" is a term whose conditions of application are context-dependent. The inferior limit below which temperature and other thermodynamic concepts can't be applied is theoretically fixed by an object's having at least *three* constituent parts capable of independent movement. Objects with only two parts and "atomic" objects

without any parts at all can't be objects of thermodynamics. In particular, it is impossible to ascribe temperature to them.

[12]It is interesting to note that Kim himself notes in passing that Nagel's model of intertheoretic reduction "does not in general require hat each $T_2$-term be correlated with a *coextensive* $T_1$-term" (Kim 1992, p. 9; his emphasis). However, as I argue in the text, his subsequent argument is vitiated by his failure of taking this important fact about reduction seriously. If he did, as I think one should, the resulting picture of the relation between reduced and reducing properties would turn out very different.

[13]This has been noted by a number of authors. Cf. P.S. Churchland (1982, pp. 101f.; 1986, pp. 356ff.), P.M. Churchland (1988, pp. 41f.), Pereboom and Kornblith (1991, p. 138), Crane (1992, p. 195), Pacherie (1993, pp. 31f.).

[14]An object is a black body if it absorbs all incident radiation of all wavelenghts, without reflecting any of it.

[15]Given what I have said about the multiple realizability of temperature, one might rather prefer to say that is has not. But that would amount to interpreting the meaning of the word "reduction" as exclusively defined by its role in the Nagelian theory of intertheoretical reduction. But it is one of the aims of this paper to show that in some cases at least, a property may be reduced without there being a unique bridge law in Nagel's sense. In a way it comes to a terminological decision to keep the name "reduction" for the relation between a macroproperty like temperature and its microphysical basis even when one questions the universal applicability of Nagel's theory to such relations. This decision seems to me to be justified by the fact that the application of the word "reduction" to whatever relation exists between temperature and its microphysical basis, as well as between other paradigmatical cases of reduction, seems to be at least partially constitutive of the meaning of that word.

[16]On the link between the causal efficacy of a property and the laws of nature in which it is embedded, cf. Kistler (forthcoming).

[17]Cf. section 3 below.

[18]It is precisely upon examination of the issue of the heterogeneity which the potential reducers of a given mental property M can be expected to have across the large variety of (natural or artificial) cognitive systems able to possess it, that Block (1997) sheds doubt on the justification of Kim's radical conclusion. Indeed, as Block argues, there are "forces of design or selection" (Block 1997, p. 122) putting heavy constraints on which physical structure could possibly underlie a given mental property. (On this point, see also Proust 1995; 1997). These contraints impose similarities on those underlying properties. If Block is right, both the functionalists' argument from MR to irreducibility and Kim's argument from MR to elimination as a natural kind are *weakened* insofar as they both rely on the assumption of (strong) heterogeneity.

[19]Blackburn (1993) uses the fact that the different realizations of temperature share the capacity to bring bodies possessing them in a common state of thermodynamic equilibrium to argue for temperature's being a perfectly respectable natural kind property and against P.S. Churchland's (1986) contention that its multiple realizability eliminates temperature as a candidate for being a natural kind property sanctioned by science. I would rather argue that temperature is a natural kind not because all of its realizations sustain thermodynamic equilibrium, but because there are laws linking *it* - not the lower-level properties of its particular instances, like mean kinetic energy - to other natural kind properties. The 0th law of thermodynamics (It says that if A and B are in thermal equilibrium, and B and C are as well, then so are A and C.) is such a law, and equilibrium is such a property, but it belongs to the same (macroscopic) ontological level as temperature itself.

[20]I do not claim to have a criterion of lawhood, enabling me to justify that I treat general statements involving "jadeite", "nephrite" and "temperature" as hypothetical laws, but not general statements bearing on "jade". I simply rely on the generally shared conviction that the first three terms designate natural kinds of mineralogy or thermodynamics, whereas the latter designates a kind of gemology which is not a science and whose kinds are not *natural* kinds.

[21]I abstract here from the further difference between *moving* molecules/atoms of gases and *vibrating* atoms/molecules of solids. For the purposes of the present argument, vibration can be considered as a kind of motion.

[22]In (Kim, 1992, pp. 5 and 7/8), Kim repeats the observation that MR might "occur 'locally' as well." He goes on to note that "we may want to avail ourselves of the flexibility of allowing a psychological state, or function, to be instantiated by alternative mechanisms within a single system. This means that Pi can be a disjunction of physical properties" (Kim, 1992, p. 7). He does not, however, take into account the implications of this possibility that I point out in the text.

[23]Parts of an earlier version of this paper have been read at the 5th Annual Meeting of the European Society of Philosophy and Psychology, Barcelona, July 1996. I would like to thank my auditors in Barcelona and Ned Block, Joan Cullen, Marcel Lieberman, Joëlle Proust and an anonymous referee for this journal for their helpful criticisms.