The Limits of Functional Reduction

Abstract

It is obvious why the antireductionist picture of mental causation, which considers mental phenomena to be causally efficacious, is so attractive: it preserves the unique nature of the mental (mental realism), while at the same time it tries to secure a place for the mental in our world which is compatible with a physicalist ideology (physical monism).

But Kim’s so called argument from supervenience reminds us of the dilemma that we face while favouring antireductionist solutions of mental causation which might force us to abandon them and look for some other more plausible ones. The trouble is, namely, the following: either mental properties have causal powers or not. If they have them then we violate the causal closure principle which means a denial of physicalism. If not, then we embrace epiphenomenalism, which denies the mental causal powers of any sort. So, either we give up physicalism or accept epiphenomenalism. So, either we give up physicalism or accept epiphenomenalism. Since antireductionism loses both ways neither of these options represents a true alternative for its proponents. For this reason some authors think that we should look at reductionism in order to explain mental causation in a more satisfactory manner. However, it is not the traditional Nagel model of reduction that is in play here, but some rather more sophisticated ones. The first part of the article presents the reasons for dropping classical reduction, the second part describes Kim’s functional model of reduction as one of its possible successors, and the final part of the article discusses the reasons for its failure.

Key words

reductionism, anti-reductionism, Nagel’s classical model of reduction, Kim’s functional model of reduction, Lewis-Armstrong causal theory, mental causation

I.

In the standard sense reduction means the process of contracting the laws of secondary (or reduced) science to the laws of primary science by the use of bridge-laws as a means for relating their predicates. Bridge-laws have forms of equivalences and assure the connection of each property from the secondary theory with an adequate property, i.e. its nomological antipode, from the primary theory. In order to arrive at a successful reduction in the traditional Nagel sense the availability of bridge-laws to always connect the same two properties is crucial: if mental property $M$ (say pain) in the secondary theory is related exclusively to physical property $N_p$ in the primary theory, we can successfully reduce $M$ to $N_p$, under the assumption, of course, that both theories are true. But the disposability of bridge-laws in the sense $M\rightarrow N_p$ is problematic. The multiple realization argument proves that each mental property can be realized by various physical properties. This prevents the connection of the higher-order property $M$ with only one lower-order property (say $N_p$) and makes the bridge-law $M\rightarrow N_p$ impossible. In other words, the multiple
realization argument destroys any hope of a successful reduction in the classical Nagel sense.

One of the more promising answers to the question regarding the availability of bridge-laws, which is supposed to show that standard reduction is still possible despite the multiple realization argument, says that this availability must not be understood in the global sense, i.e. holding for all different biological kinds, but merely in the local sense, i.e. holding within a single biological kind. Neural properties to which pain is related are clearly different from kind to kind, so that in humans pain is realized, for example, by $N_{ph}$, in reptiles by $N_{pr}$ and in Martians by $N_{pM}$. And if $N_{ph}$ is the physical property that realizes pain in one specific biological kind, i.e. humans, a specific bridge-law must hold in it, that is, $M_{ph} \leftrightarrow N_{ph}$, where $M_{ph}$ designates pain in humans. A proper system of such equivalences that includes all mental states in humans would enable us to reduce the mental to the physical within kinds, for example, human psychology to human neurophysiology and so on.

However, this does not mean that Nagel’s proposal has to be understood in such a sense and that all problems simply vanish if classical reduction is limited to kinds. Two problems concerning his suggestion still remain; they question the idea of traditional reduction and bridge-laws in general. And this in turn poses a threat to everything of a local nature. The first problem refers to the character of the reduction that is supposed to be explanatory, which means that the reduction has to demonstrate why or how a phenomenon at a higher level arises from a phenomenon at the lower level. The question is why are all organisms in pain in the exact physical state $N_p$, and not some other, say $N'_u$? Kripke writes:

“It would seem, though, that to make the C-fiber stimulation correspond to pain, or be felt as pain, God must do something in addition to the mere creation of the C-fiber stimulation; He must let the creatures feel the C-fibres as pain, and not as tickle, or as warmth, or as nothing, as apparently would also have been within His powers.”

To say that this is just the neural state which realizes pain in different (or similar) systems that are capable of feeling does not seem to be sufficient because it tells us absolutely nothing about why precisely pain and $N_p$ are related. Bridge-laws, therefore, occur in reduction as additional premises that are not specifically explained and as such do not contribute anything at all to our understanding of the world. It seems that because bridge-laws are unable to account for the connection between the mental and the physical, traditional reduction simply takes this relation as a brute fact. But from a physicalistic point of view this spells disaster: such a reduction turns out to be incapable of providing substantial physical explanations of non-physical events.

“So [Nagel] reductions, whether global or local, do not give us reductions that explain. Even if all phenomena have been Nagel-reduced to basic physical theory, of which let us suppose we have a completed version, the world would still be full of mysteries, mysteries that defy our completed physics.”

The second problem refers to the function of the reduction that is supposed to simplify our ontological picture, and here bridge-laws again fall short. Since we do not know why in the bridge-law $M \leftrightarrow N_p$, these two properties are connected we can say that they are non-rigid, which makes bridge-laws merely contingent (it is namely not necessary that if $M$ then $N_p$). $M$ and $N_p$, therefore, remain two different properties; they would be identical, i.e. they would be one and the same property, only if they were, as Kripke says, rigid, which would make the bridge-law $M \leftrightarrow N_p$ necessary (it would not be contingent that
that if \( M \) then \( N_p \). Contingency of bridge-laws that results from the non-rigidity of their properties prevents the identification of \( M \) and \( N_p \) and so buries the reductionists’ hopes for a simpler ontological picture of our world.

“In any case the metaphysically significant fact is that Nagel reduction gives us no ontological simplification, and fails to give meaning to the intuitive ‘nothing over and above’ that we rightly associate with the idea of reduction.”

Since bridge-laws do not help us to solve any of the cited problems, the classical model of reduction does not play a significant role in metaphysical discussions. Moreover, a rejection of such reductions cannot count as an outstanding philosophical contribution. However, we should not diminish its importance, since it can surely take at least some credit for the rise of other, maybe more plausible, models of reduction. But the inability of the traditional model of reduction to play a notable role in contemporary metaphysics does not mean that all models of reduction face the same fate, or that there is not some other reduction able to avoid the described imperfections and take centre stage in the contemporary discussion of the mind-body problem.

II.

Kim is one of the philosophers who offer a different model of reduction. He starts by saying that in order to simplify our ontological picture, bridge-laws, such as \( M \leftrightarrow N_p \), have to be substituted by identities \( M = N_p \). By doing so, we kill two birds with one stone, because, at the same time this also explains why specifically \( M \) and \( N_p \) are connected. The question is, of course, can we do it and how?

He thinks that the first step is the formulation of \( M \) as a functional property, i.e. extrinsic,\(^4\) in terms of its causal-nomological relations to other properties, and not intrinsic, i.e. in terms of its internal character or material structure. Take a gene and a DNA molecule.\(^5\) A gene is the biological mechanism that according to physical laws does the causal work in the sense that it transmits internal and external properties from one generation to another, whereas a DNA molecule is the physical compound that actually performs the causal role, i.e. it is what does the causal work of the gene in our world and all others physically similar to it. What about temperature? Temperature is the property of an object which, for example, causes steel to melt or water to freeze, but this causal role is in the actual world and all others that are in all physical aspects similar to it performed by molecular energy. What about pain? Pain is the mechanism that according to natural laws does the causal work in the sense that it detects and dismisses injuries, whereas its physical realizer \( N_p \) is the state that actually performs the causal role, i.e. it is what does the causal work of pain in our world and all others physically similar to it.\(^6\)

4 C, for example, is an extrinsic property if its existence depends on the existence of other properties.
6 Let us ignore multiple realization to get a greater simplicity.
We described genes, temperature and pain in terms of their causal functions, that is, in terms of their typical causes and effects. And then we discovered that DNA molecules, molecular energy and \( N_p \) were those properties which fulfilled their causal specifications. What does this mean? This means, Kim says, that genes, temperature and pain are properties which have certain other properties, i.e. DNA molecules, molecular energy and \( N_p \), and because, in general, properties which have other properties are one and the same property, genes are nothing more than DNA molecules, temperature nothing more than molecular energy and pain nothing more than \( N_p \). Functionalization of the second-order property \( M \) enables us to identify it with the first-order property \( N_p \), and provides us with a simpler ontological picture of the world. Moreover, it also provides us with an explanation of why exactly \( M \) and \( N_p \) are related: when \( N_p \) occurs \( M \) occurs also because \( N_p \) and \( M \) are one and the same property. Functional reduction is clear: the latter was neither able to solve the explanatory problem, i.e. why \( M \) and \( N_p \) are connected, nor was it able to solve the ontological one, i.e. to simplify our ontological picture. None of these represents a problem for the former.

However, there are three questions regarding functional reduction that must be answered. The first one refers to the status of psycho-physical identities. Functionalization of second-order properties, or their formulation in extrinsic terms, is the core of functional reduction. But since \( M \) is defined in terms of its causal-nomological relations based on laws that hold in the actual world there is still a danger that the property which will fulfil \( M \)'s causal specification will be again merely coincidental. The property that realizes \( M \) changes from world to world which makes it non-rigid (this is in contrast to Kripkes’ requirement that theoretical identities are supposed to be metaphysically necessary) and its identity with \( N_p \) is therefore metaphysically contingent. Kim’s answer to the question is as follows:

“Since whether or not \( N_p \) is a realizer of the functional property \( M \) is determined by the prevailing laws of nature, the realization relation remains invariant across all worlds with the same basic laws. Thus \( M=N_p \) holds in all nomologically possible worlds (in relation to the reference world). Accordingly we may say ‘\( M \) is nomologically rigid or semi-rigid.’”

This makes the identity of \( M \) and \( N_p \) nomologically necessary and this is enough. The second question is even more transparent: if \( M \) is a second-order property and \( N_p \) is a first-order property is it then not contradictory to say that they are identical? It is not possible that \( M \) is realized and is at the same time the property that realizes it. Kim answers the question by saying that in order to identify pain with \( N_p \) we first need a concept of pain, that is, we must first know what the word ‘pain’ means at all; just as we must have a concept of a gene first in order to identify it with a DNA molecule. And what is a concept of a gene? It is a concept of the internal mechanism responsible for the transmission of hereditary material. By this, gene is defined in terms of its causal function. And after it was discovered that DNA molecules were what performed the function it was not difficult to identify genes with them. Evidently we first had a conceptual identity and then went over to a metaphysical one after seeing the results of empirical investigations. Kim thinks that the same holds for pain. A concept of pain is a concept of the internal state caused by tissue damage that itself causes wincing and groaning. By this, pain is defined
in terms of its causal role. And after it was discovered that $N_p$ was what performed the role (ignore multiple realization) it was not difficult to identify pain with it. This can be illustrated by the following scheme:

2nd level  
concept of pain: causal specification

1st level

\[ \text{pain} \]
\[ N_p \]

(The solid line denotes the relation of identity.)

Since pain is identical to $N_p$, it is in fact a first-order property; it is the concept of pain that is a second-order property. Kim therefore says that to have the second-order property $M$ is to have the first-order property $N_p$, which makes our discussion of $M$ as an independent or autonomous property redundant. Again, it is the concept of pain that belongs to the higher level, while pain itself, because it is identical to $N_p$, belongs to the lower level. He thinks that “… it is less misleading to speak of second-order descriptions or designators of properties, or second-order concepts, than second-order properties” and believes that

“… we may want to, or perhaps should, give up the talk of second-order properties altogether in favor of second-order designators of properties, or second-order concepts”.\(^{11}\)

Second-order designators are used when we do not want to describe the given property in terms of its realizers, for example, the fact that I shook hands with George Bush yesterday is sometimes, or always, expressed instead as ‘yesterday I shook hands with the president of the USA’. In this way, listeners obtain information which is important in a given context and which otherwise they would not get, e.g. that we speak of a leader of the greatest military force in the world and not of someone who did not serve in the army, which might be one of the implications resulting from the use of the words George Bush, but irrelevant in the given context. Kim concludes that

“… from the ordinary epistemic and practical point of view, the use of second-order property designators probably is unavoidable, and we should recognize that these designators introduce a set of useful and practically indispensable concepts that group first-order properties in ways that are essential for descriptive and communicative purposes.”\(^{12}\)

---

\(^{7}\) Kim says: “/…/ by definition, having $M$ is having a property with causal specification $D$, and in systems like $s$, $P$ is the property (or one of the properties) meeting specification $D$.” – J. Kim, Mind in a Physical World, p. 24.

\(^{8}\) Gene = DNA molecule, temperature = molecular energy, and pain = $N_p$.

\(^{9}\) J. Kim, Mind in a Physical World, p. 98.

\(^{10}\) Ibid., p. 99–100. $P$ in the quotation is replaced by $N_p$.

\(^{11}\) Ibid., p. 104, 106.

\(^{12}\) Ibid., p. 104–105.
The final question regarding functional reduction refers to multiple realization. It is obvious that any model of reduction must be compatible with the idea about different physical realizers of the same mental kind, and functional reduction is no exception. Let us see how we are supposed to carry out such a reduction when we deal with the mental property $M$ that has different realizers in different biological species in different worlds all governed by the same natural laws as ours. Suppose that functionalization of $M$ is done and that its description in terms of causal-nomological relations says that it is the internal state caused by tissue damage that itself causes wincing and groaning. So, to have $M$ is to have a property with such a causal specification, which we will call $D$. Functional reduction consists in identifying $M$ with the property that fulfils its $D$, and given multiple realization everything makes sense only if within different biological kinds different physical properties fulfil it, e.g. $N_{ph}$ in humans, $N_{pr}$ in reptiles and $N_{pM}$ in Martians. Therefore every $M$-token is identical to its neural realizer (note that they are both entities of the same level), which is in contrast to classical Nagel reduction via bridge-laws (even if understood in the local sense): by supposing merely a nomological connection between $M$ and its realizer they always remain two different things. Kim himself sees his functional reduction as local reduction:

“In this way multiply realized properties are sundered into their diverse realizers in different species and structures, and in different possible worlds.”

What about functional reduction and mental causation? How does Kim’s model of reduction solve the causal exclusion problem? The latter results from so called causal overdetermination: here two different and independent causes are supposed to bring about the same physical effect. The following scheme represents this situation ($M$ stands for a mental property, $N$ for a physical property and wincing for physical behaviour):

\[
\begin{align*}
S_1 & \quad \quad M \\
& \quad \quad \quad \quad \quad N \\
& \quad \quad \quad \quad \quad \quad \text{wincing}
\end{align*}
\]

(The broken arrow denotes a relation of supervenience; the solid arrows represent a relation of causation.)

The same physical effect has two different causes, mental and physical, each being enough for its occurrence, and is as such causally overdetermined. This leads any physicalist to the problem of causal exclusion because “two or more complete and independent explanations of the same phenomenon cannot coexist”.

The physical cause, since it is alone enough for the occurrence of a physical effect, threatens to exclude the mental one, and this makes the mental cause that brings about the same physical effect redundant. This is in fact the greatest problem concerning mental causation that a theory must solve in order to be satisfactory.

The question, therefore, is what is Kim’s solution to the situation $S_1$? Since in functional reduction every $M$ token is identical to its realizer, every $M$ token also has the same causal powers as its realizer, no matter that all causal work is actually done by the latter. Kim calls this the causal inheritance principle. Its definition says that if some first-order property $N$ at time $t$ realizes some
second-order property $M$ ($M$ occurred because one of its neural realizers occurred) the causal powers of this $M$-token are identical to the causal powers of this $N$-token. The scheme that represents such a local reductionist solution of the causal exclusion problem and mental causation in general, in which a mental property and its physical realizer are always limited to a single biological species, is the following:

$$S_2$$

(M$_{ph}$ stands for a pain-token in humans; the solid line denotes a relation of supervenience; the solid arrow represents a relation of causation.)

Since on this model of mental causation a mental cause is identical to a physical cause, we do not speak of two different causes anymore but merely of one. Nevertheless, functional reduction ascribes pain causal powers: it has them insomuch as its neural realizer has them, so that pain’s causal role is preserved. However, even in functional reduction causal powers of mental properties are not new and irreducible (genuine), which makes them real only insomuch as physical properties are real. This means that they cannot play any role in causal psychological explanations of behaviour; the latter must refer to entirely real phenomena if psychology wants to be an independent science. It seems that in Kim’s opinion the price we have to pay in order to have at least some sort of mental causation is local reductive physicalism which tolerates not genuine but merely parasitic causal powers of mental phenomena.

### III.

Are we ready to accept functional reduction? What keeps us away from it? It is compatible with multiple realization, granted, only at a local level, but nevertheless. It might also not bother us that it holds merely for intentional states, such as beliefs and desires, and not for all mental states: phenomenal states, such as sensations and emotions, are left out. But the following question remains: is Kim’s functional reduction really completely different from all other reductions that are or were on the market? Some authors think that it is not. Since it has to rely on second-order concepts instead of second-order properties, otherwise it is not clear how a property can be both realized and at the same time its realizer, it seems that it is only a slightly modified Lewis-Armstrong version of reduction, known as the causal theory. Let us take a closer look at it and try to establish if this is true.

---

13 Ibid., p. 111.


According to their causal theory, an identification of mental phenomena with neural ones includes two utterly different steps: (a) a conceptual analysis of different types of mental concepts together with the way in which they are connected; this gives us a thematically neutral model for defining mental processes in terms of their causal roles, and (b) a coincidental identification of these processes and states with those processes and states in the brain, i.e. in the central nervous systems, that are, according to scientific discoveries, relevant for this.\textsuperscript{18} In short, first, pain has to be defined in terms of its causal function and then, after science finds out to which internal states it corresponds, pain can be identified with it. In other words, we first must have a conceptual identity and then go over to a metaphysical one after we see the results of empirical investigations.

“Our view is that the concept of pain, or indeed of any other experience or mental state, is the concept of a state /…/ with certain typical causes and effects. /…/ Or, better, of a state apt for being caused in certain ways by stimuli plus other mental states and apt for combining with certain other mental states to jointly cause certain behaviour. /…/ If the concept of pain is the concept of a state that occupies a certain causal role, then whatever state does occupy that role is pain. If the state of having neurons hooked up in a certain way and firing in a certain pattern is the state properly apt for causing and being caused, as we materialists think, then that neural state is pain.”\textsuperscript{19}

Such a position understands the concept of pain as non-rigid since it is accidental to which internal state it refers:

“The concept of pain… would have applied to some different if the relevant causal relations had been different.”\textsuperscript{20}

So, the concept of pain would be by no chance also the concept of this neural state, or, in other words, the concept of pain is a second-order property whereas pain itself, being identical to some contingent physical state, is a first-order property.

It looks as though Kim’s functional reduction is, indeed, merely a modified example of the Lewis-Armstrong causal theory: it likewise requires for a successful reduction functionalization of mental properties and then their identification with those physical properties that fulfill their causal functions. This is clearly seen in Kim’s answer to the second question referring to his proposal which criticizes him to identify what is realized (pain) with what realizes it ($N_p$). Here, by appealing to the concept of pain that refers to some physical state with which pain is identified, he unmistakably relies on Lewis-Armstrong model of reduction.

Some might object that functional reduction is an example of local reduction whereas the causal theory is an example of global reduction. At first sight this is true, however, a closer look reveals that their reduction also involves elements of local reduction, although they do not explicitly speak of it.

“If the concept of pain is the concept of a state that occupies that role, then we may say that a state is pain for a population. Then we may say that a certain pattern of firing of neurons is pain for the population of actual Earthlings and some but not all of our otherworldly counterparts.”\textsuperscript{21}

It seems that Lewis allows for the possibility that after pain’s causal specification is determined, or after it is functionalized, pain is identified with different physical states in different biological kinds regarding the different physical structure of systems. So, their causal theory has a local nature.

If we agree that both reductions are practically identical then functional reduction faces all the same problems that the causal theory does. The biggest
problem for such reductions is that we can conceive a case in which organisms would be in pain although their outputs, that is, the combination of behaviour and other mental states, would be either different from pain’s causal specification described in the concept of pain or they would not occur at all. Take your brain which was taken from your body in order to be cleaned and renewed. A connection between the brain and body, which went shopping in the meantime, was maintained via radio communication. After the cleaning and renewing are done, the brain is put back in your head. It sometimes happens that while the brain is undergoing this beauty treatment and is only remotely connected with the body the latter suffers an accident that partly destroys it, e.g., since cancer has spread to your vocal chords they must be removed. But the brain connected to the sense organs of such a body will not be able to bring about the effects determined in the pain’s causal specification. With the vocal chords removed they will not groan and ask for aspirin. It seems also sensible, however, to assume that you will nevertheless be in pain. Pain will occur together with its physical realizer after the brain was stimulated in the appropriate way. It is only that your behaviour is different from what is required by the concept of pain. It follows that it is questionable whether or not the behaviour and states described by causal specifications and contained in mental concepts are really something that is essential for them: pain can occur also when our behaviour is completely different from what is prescribed, when we, for example, laugh or sing. If this holds then neither in the Lewis-Armstrong causal theory nor in Kim’s functional reduction can the first required step for a successful reduction, that is, functionalization of mental states, be made. It seems that the reductions would work only if the causal descriptions of mental phenomena included every possible behaviour and state, which seems absurd.

Unfortunately, functional reduction, as tempting it may look, fails to solve the mind-body problem. It does not solve mental causation either since the identification of two causes, mental and physical, is not possible. So, even if we accept that it is limited to biological species and intentional states, we are still left empty handed.

David M. Rosenthal (ed.), The Nature of Mind, Oxford University Press, New York 1991. – Their views are somewhat different but this is ignored because it is of no importance for what we want to prove.

17 They wanted to prove that functionalism was, in fact, a reductive position.


20 Ibid., p. 218.


J. Bregant, The Limits of Functional Reduction

References


Die Grenzen der funktionalen Reduktion

Zusammenfassung


Schlüsselwörter

Reduktionismus, Antireduktionismus, Nagels klassisches Reduktionismusmodell, Kims funktionales Reduktionismusmodell, Lewis-Armstrongs kausale Theorie, mentale Verursachung
Les Limites de la Réduction fonctionnelle

Résumé
Il est évident pourquoi l’image antiréductionniste de la causalité mentale, selon laquelle les phénomènes mentaux sont considérés comme des causes efficientes est si attrayante : elle préserve la nature unique du mental (réalisme mental), tout en essayant d’assurer la place pour le mental dans notre monde, ce qui est compatible à l’idéologie physicaliste (monisme physique). Mais le soi-disant argument de Kim provenant de la survenance nous rappelle le dilemme auquel nous devons faire face lorsque nous favorisons les solutions antiréductionnistes de la causalité mentale, qui pourraient nous forcer à les abandonner et en chercher d’autres plus plausibles. La difficulté est la suivante : les propriétés mentales ont des pouvoirs causals ou non. Si elles en ont, nous violons alors le principe de la conclusion causale, ce qui est la négation du physicalisme. Si ce n’est pas le cas, nous acceptons l’épiphénoménalisme qui conteste les pouvoirs mentaux causals de toutes sortes. Ainsi, soit on renonce au physicalisme, soit on accepte l’épiphénoménalisme. Puisque, dans les deux cas, l’antiréductionnisme est perdant, aucune des deux options ne présente une vraie alternative pour ses tenants. Pour cette raison-là certains auteurs pensent que nous devons nous adresser au réductionnisme afin d’expliquer la causalité mentale d’une manière plus satisfaisante. Pourtant, ce n’est pas le modèle traditionnel de réduction de Nagel qui est ici en jeu, mais il s’agit plutôt d’autres modèles plus sophistiqués. La première partie de cet article présente les raisons pour que soit abandonnée la réduction classique, la deuxième décrit le modèle fonctionnel de réduction de Kim, qui en est un successeur potentiel, alors que la fin traite les raisons de son échec.

Mots clés
Réductionnisme, anti-réductionnisme, le modèle traditionnel de réduction de Nagel, le modèle fonctionnel de réduction de Kim, théorie causale de Lewis-Armstrong, causalité mentale