A physicalist account of multiple realizability in the special sciences

Meir Hemmo[1] and Orly Shenker[2]

This is the text of a talk given in the PSA2016 meeting in Atlanta. For this reason there are no references in it.

One of the main achievements of analytic philosophy in the 20th century is the formulation and defense of the thesis called non-reductive physicalism (in various versions). This project is central to the work of Putnam, Davidson, Fodor and to some extent Lewis, and many others. Non reductive physicalism is considered to be a major achievement in analytic philosophy precisely because it seems to many that – on the one hand – it says that everything is fundamentally physical, while claiming – on the other hand – that there are high-level facts in the world that are not reducible to the fundamental facts of physics, and consequently there are high-level laws in the special sciences that are genuinely autonomous from the laws of physics. The thesis of multiple realizability is central to these ideas. One of the prevalent arguments in support of multiple realization is *empirical*.

In their recent book, Polger and Shapiro examine in detail various case studies of alleged multiple realization and show very convincingly that – as a matter of fact – the empirical evidence is not conclusive. Their conclusion is that multiple realization is not as wide-spread as some people tend to think. However, they do not reject the idea that multiple realizability is a genuine logical possibility. We join Polger and Shapiro in this conclusion. However, I will argue here on *a priori* grounds that multiple realizability is *incompatible* with physicalism, because it entails what we call *token dualism*. We shall explain below in detail this implication and how it comes about.

The plan of the paper is as follows: I will first put forward a type-type identity theory, I will defend it, and then I will argue for the following two points.

[1] Philosophy Department, University of Haifa, 31905 Haifa, Israel, e-mail: meir@research.haifa.ac.il
[2] Program in the History and Philosophy of Science, The Hebrew University of Jerusalem, 91905 Jerusalem, Israel, e-mail: orly.shenker@mail.huji.ac.il

1. Multiple realizability is *incompatible* with physicalism. This holds even if high-level kinds *supervene* on the lower-level physical kinds.

2. A type-identity physicalist theory can explain away multiple realization in the special sciences in all cases where multiple realization seems to hold.

I need to put forward some conceptual tools taken from the foundations of physics concerning the way in which fundamental physics accounts for natural kinds in some special sciences. I will focus here on the high-level kind 'temperature' in thermodynamics and statistical physics. This will set the stage for understanding in terms of our best theories of physics what is genuine multiple realizability and what it implies.

According to our best physical theory (for example, classical mechanics or quantum mechanics or quantum field theory) the world at any given time has a *complete* state. This state is given (metaphysically speaking) by what is called the '*microscopic state'* (or for short '*microstate'*). The microstate of the world (or of any subsystem of it) is considered to be the *complete* and *exhaustive* state of the world at a given time *without remainder*. In other contexts, the term 'microstate' sometimes means *small* or *part of a whole*. This is not the intended meaning here. Here – I repeat – the term microstate denotes the *complete physical* state of the world at a time which means that given this state there is no further information that one can add in order to describe in more detail or more accurately the physics of the world. We can now understand the physicalist thesis as follows: *every token microstate* of the world is strictly and wholly physical.

Now suppose that we wish to say something about the world, in some high-level special science, say biology, which seems *not* to be captured by the world's physical microstate, but nonetheless we want it to be compatible with the physicalist thesis. The only compatible thing we can say, which will be informative and nonetheless not repetitive, is say much *less*, that is say something about an *aspect* of the complete microstate, which is given by some *partial description* of it and compare it to aspects of *other* possible microstates.

The special science of thermodynamics is understood in physics *precisely* in this way. *Temperature* of an ideal gas in equilibrium, for example, is *identical* to – a certain *aspect* of the microstate of the gas, namely the average kinetic energy of the particles that make up the gas. So temperature is *not* some high-level fact over and above the microstate of an ideal gas in equilibrium, but rather it is just the average kinetic energy of the particles that make up the gas. This is an *aspect* of the gas's microstate which may be common to other possible microstates. That is, temperature is the understanding of temperature in the Boltzmann approach to statistical mechanics. What we have here is a cross-level *identity* theory in terms of partial aspects of microstates. These aspects are called *macrovariables* or *macrostates* in statistical physics, and the regularities they turn out to exhibit just are the laws of thermodynamics.

Despite this identity theory, the high-level laws of thermodynamics are radically different from the laws of the low-level mechanics (laws such as F=ma or the Schroedinger equation in QM). For example, the laws of thermodynamics are *not* time-symmetric (i.e. they are not invariant under the reversal of the direction of time), despite the fact that the fundamental laws of mechanics are time-symmetric, and that all there is in the world is the time-symmetric evolution of microstates! The explicit target of statistical mechanics is to *reduce by means of cross level identities* the entirety of thermodynamics to mechanics and nonetheless obtain the thermodynamic time-asymmetric laws from the time-symmetric laws of fundamental mechanics. We set aside here the question of whether this target has indeed been already achieved. What I wish to stress here is that thermodynamics is understood by this reductive approach as a case of a special science. And what I wish to do here is advance this reductive approach to all the special sciences. If one wishes to construct a theory of any special science that is compatible with the universal imperialism of fundamental physics -- as just explained -- there is simply no other way.

Let us stress the nature of the *identity theory* that comes out from this discussion. We are talking about a cross level *identity* theory, in which a high-level kind is identified with a macrovariable or an aspect of the token low-level microstate where this macrovariable is a fact pertaining to each and every *token-microstate*. For example, the high-level thermodynamic kind, that is the macrovariable of

temperature, *is a fact pertaining to each and every individual token*-microstate of the world. This fact is not picked out by the entire microstate, but only by an aspect of it. This aspect then explains the partition of the tokens into the kinds by the fact that this aspect is *shared* by all the tokens of this kind. This type-identity theory is suggested by statistical physics (according to our understanding of it) and we propose to generalize it to all the special sciences. I am not sure whether this is the standard way of thinking about the identity theory in the literature. If not, this identity theory is new.

This identity theory has the following consequence. Often in the literature it is argued that since there are many *different* microstates that may realize the *same* temperature, we have here a genuine case in physics of multiple realization. But this is a *mistake*! It presupposes that temperature is realized by the *entire* microstate of the gas, and this assumption is not entailed by physics! Temperature is identified with an aspect of the microstate of the gas and not with the full microstate. And since all the microstates that realize the same temperature share the same aspect, we say that there is no multiple realization in statistical mechanics.

Let us now turn to multiple realizability. A set of microstates is said to (genuinely) multiply realize a special science kind just in case the microstates in the set that realize this kind are *heterogeneous*. By heterogeneous I mean that the microstates *don't share any aspect or macrovariable that can be identified with the special science kind.* The crucial point here is the heterogeneity: Multiple realization holds if and only if there is no shared aspect according to our identity theory.

I will now argue that if multiple realization holds the high-level facts must be *new* primitive facts which are *not* physical. To see why consider the following argument. Suppose that God created the world by creating a collection of token-microstates A, B and C, each of which is wholly physical, without remainder. Suppose also that together with the microstates God created the laws of fundamental physics. Is this enough in order to yield the world as we know it, in particular is this enough to give the special sciences high-level kinds? Suppose also that the partition to the special sciences kinds is such that the high-level kind P *supervenes* on the low-level physical tokens A and B. And finally suppose that the token C is not P. If God's creation of the

tokens and the laws is enough to fix this partition, then it *must* be the case that the partition is based on shared physical macrovariables, essentially along the lines of statistical physics. This is the type identity physicalism we have in mind.

We stress that if the partition to the high-level kinds is brought about by the laws of physics so that the laws of physics make it the case that A and B are P and C is not P, this means that the partition to the high-level kinds is based on shared macrovariables. If this is God's world, it isn't a world in which multiple realization holds.

Suppose now that multiple realization holds, so that A, B and C are physically heterogeneous. If this is true, it means that the laws of physics don't make it the case in terms of macrovariables and supervenience that A and B are P and C isn't. And so in order to create a world in which multiple realization holds, God had to make *another* creation *over and above the creation of the tokens and the laws* which will make it the case that A and B are P and C isn't. That is, God must in this case *add the partition of the tokens into the high-level kinds as another fact*, which is *not dictated by the physics of the tokens*! And this must be a new unphysical fact built into *each and every token*, namely the fact that *fixes or determines to which kind the token belongs*. I stress that this holds even though we have assumed supervenience! This is *token dualism*!

Of course we don't take the God fairy-tale seriously. But the serious point in this story is this: According to physics all there is in the world is the world's token-microstate. If multiple realization holds, the physics of the token-microstate is not enough to determine the partition to sets! And so regardless of supervenience it cannot be the case that each token is wholly and only physical. But this means that there is a fact about a token, which is *not physical*. And it is *this non-physica*l fact about the token, which determines whether or not the token belongs to any special science kind. Finally: This means that supervenience does not secure the minimal physicalist idea -- that the physics of a token determines -- or grounds or fixes -- the token's high level behavior.

There are in the literature ways of talking about the type=token relation other than realization, such as grounding and constitution. Our conclusion holds with respect to

all of these ways, as long as the partition of the high-level kinds is not given in terms of shared macrovariables and the laws of physics.

In the literature the high-level kinds are sometimes thought of in terms of functional kinds (computational or causal functions). If the function and the functional role of tokens are determined by the laws of physics, functionalism is a special case of type-identity. But if the laws of physics don't determine the functions and the functional role of tokens, then there is an *extra non-physic*al fact at play here, and exactly as before this implies token dualism.

Some non-reductive physicalists say that "God knows" what it is that partitions the tokens into the sets (as Fodor says). By this they mean presumably that this partition is a *brute fact*. The point here seems to be that if the partition is a brute fact, then it is *not* dictated *by* the tokens, and therefore the tokens can be wholly physical. But this is wrong, since, as we saw, the brute fact is an *additional* fact over and above the tokens and the laws of physics, and this means that even God cannot base the partition to the high-level kinds wholly and only on the physics of the tokens.

We conclude that if the world is completely physical, the fact *that determines* to which high-level kinds (or sets) the token belongs must be a *physical* fact *encoded in the token*. Since non-reductive physicalism rejects this conclusion, by accepting multiple realization, it collapses into token dualism! And by now it is immaterial whether the dualism in question is property or substance. By this we have proved the first point of the paper.

Let's now turn to the *second claim* that the *appearance* of multiple realization (if there is such an appearance) in the special sciences can be explained by our type-*identity* theory. Our treatment here equally applies to both classical and quantum mechanics and is compatible with low-level physical laws that are deterministic or stochastic.

The set up is such that we have a straightforward physical interaction between a system (to which we ascribe some high-level behavior) and a device (which could

ultimately be the observer). The interactions bring about the following time evolutions:

(1) READY_d * A_s → X_P_d * A_s
(2) READY_d * B_s → X_P_d * B_s

Here the interactions are written in terms of *microstates*. In both time evolutions we assume that the device begins in the *same* READY microstate and ends up either in the same *microstate* P, or more generally, in two different final *microstates* that have the shared macrovariable P. The microstates A and B of the system in the two evolutions are *heterogeneous*. Since the device ends up in both evolutions in the *same* macrovariable P, we mistakenly interpret the interactions *as if* they are measurements of the macrovariable P indicated by the device. And since A and B are heterogeneous here, we mistakenly suppose that the high-level kind P pertains to the system. And this is why we think that P is multiply realized by A and B. But nothing of this sort happens. What does happen is that the device doesn't really measure anything in the system, but arrives to its final state P because of the dynamical laws of the evolution. This is consistent with everything we know in fundamental physics. So we have the *appearance* of multiple realization, accounted for by physics, and explained here by *the shared property P of the device*.

So in physicalism there are only two options: High-level kinds are either shared macrovariables of the systems or else they are shared macrovariabes of devices. It seems to us that multiple realization in all the special sciences can be explained in exactly the same way. The analysis we gave here is the only route that is compatible with physics.

We can better understand now the notion of temperature in statistical mechanics. We said before that temperature of an ideal gas in equilibrium pertains just is the *macrovariable* of average kinetic energy. This is the way in which temperature is understood in the so-called Boltzmann approach to statistical mechanics. In the Gibbsian approach the account of thermodynamic macrovariables is different. It is given in terms of functions over the entire phase space of the system. In the standard understanding of the Gibbsian approach these functions are *interpreted as*

macrovariables of measuring devices, such as a thermometer accounted for by the dynamics of the interactions. The thermodynamic kinds on this approach are treated along the lines of the analysis of interactions I just gave. But in both approaches to statistical mechanics there is *no* multiple realization.

So in what sense are the thermodynamic laws and kinds autonomous after all? The strict short answer is that they are *not*. From God's eye point of view, if physicalism is true the fundamental low-level token and its law-like time evolution dictate *everything*. But the special sciences kinds are *new physical kinds* (just like temperature). However, from our point of view the physical macrovariables that the special sciences laws talk about are highly complex and inaccessible and will presumably remain so. It is therefore unreasonable to think that we will ever do without the special sciences. So, despite the fact all there is in the world are only physical tokens, our best attitude should be that the special sciences are autonomous from fundamental physics.

Conclusions:

1. Whether or not *genuine* multiple realization holds is a question of fact, but if it holds, physicalism is false.

2. *Apparent* multiple realization in the special sciences can be explained *away* by type-identity physicalism.

3. The special sciences are autonomous just because they are *new* branches of physics. Since the macrovariables identified with the high-level kinds need not be the kinds familiar from known lower-levels laws, such as the thermodynamic laws for example, the behavior of these *new* physical kinds need not have anything to do with the thermodynamic laws. That is, the special sciences laws may be radically different from the thermodynamic laws and may exhibit a behavior over time of the new higher-level macrovariables that need not resemble neither the fundamental microphysical laws nor the laws of lower-level kinds that are already known.

4. A straightforward consequence of this type-type identity theory is that the distinction between high- and low-level kinds (and laws) has no ontological implications in the following sense. On this view, the so-called levels don't correspond to levels of *reality*: there is only *one* level of reality. Instead of levels, the theory appeals to different degrees of abstraction and different degrees of description, coarse-graining etc. In this way various mysteries that arise due to the talk about levels, such as top-down causation, brute facts that somehow give rise to sets ("high-level kinds") etc., simply evaporate. By contrast, note that if one assumes multiple realization, then it follows (from our argument for the first point) that there are at least two distinct levels of reality, since multiple realization as we agued implies token dualism.

5. Another important consequence of our analysis is that there are essentially only two kinds of fundamental (metaphysical) relations in nature: identity and causation (regardless of how one makes sense of causation). Identity replaces all sorts of *non-causal* relations that have been supposed to hold *across* or between the so-called different levels, and once the high-level kinds are identified with the low-level macrovariables, what remains to be done is provide the causal structure of the macrovariables that holds at the single level of reality (and their temporal and functional behavior at this level). In the literature other metaphysical relations such as realization, grounding, constitution, etc., have been proposed in order to explain the way in which high-level facts arise from the fundamental matters of fact. Our type-type identity theory provides a physical explanation of these relations. Realization, grounding, constitution and the like should be thought of in our picture as ways of describing various combinations of the two fundamental relations, namely identity and causation. Note that the type-type identities here between macrovariables and the so-called higher-level kinds are self-explanatory precisely because they are strict identities that need no further explanations.

6. The resulting picture is parsimonious, appeals to no miracles brute facts and relations other than the most simple and self-explanatory ones of identity and causation and has quite a strong explanatory power. The picture provides a unified account of all the special sciences together with fundamental physics according to which everything is indeed physical.