

SPECIAL ISSUE ARTICLE

Flat Physicalism

Meir Hemmo¹  | Orly Shenker² ¹University of Haifa, Haifa, Israel²The Hebrew University of Jerusalem, Jerusalem, Israel**Correspondence**

Orly Shenker, The Hebrew University of Jerusalem, Jerusalem, Israel.
Email: orly.shenker@mail.huji.ac.il

Funding information

Israeli Science Foundation, Grant/Award Numbers: 690/21, 1148/18

Abstract

This paper describes a version of type identity physicalism, which we call *Flat Physicalism*, and shows how it meets several objections often raised against identity theories. This identity theory is informed by recent results in the conceptual foundations of physics, and in particular clarifies the notion of ‘physical kinds’ in light of a conceptual analysis of the paradigmatic case of reducing thermodynamics to statistical mechanics. We show how Flat Physicalism is compatible with the appearance of multiple realisation in the special sciences, and how and in what sense the special sciences laws are autonomous from the laws of physics, despite the full reductive picture of Flat Physicalism. We compare our view with a recent proposal by William Bechtel that accounts for the appearance of levels in mechanistic explanations.

KEYWORDS

identity theory, multiple realisation (realisability), physical kinds, physicalism, reduction, special sciences, statistical mechanics

1 | INTRODUCTION

Materialism as a theory of the nature of the world has had a curious history. Arising almost at the beginning of Greek philosophy, it has persisted down to our own time, in spite of the fact that very few eminent philosophers have advocated it. ... A system of thought which has such persistent vitality must be worth studying, in spite of the professional contempt which is poured on it by most professors of metaphysics. (Russell, 1925 p. V)

Although by and large,

“down to this day, the mind-body problem ... has been that of finding a place for the mind in a world that is fundamentally physical (Kim, 1998, p. 2),

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2022 The Authors. *Theoria* published by John Wiley & Sons Ltd on behalf of Stiftelsen Theoria.

the prevalent physicalist views are non-reductive ones, for example, causal or computational functionalism. *Reductive*-type identity physicalism is *not* itself a mainstream view in contemporary thinking in philosophy as well as in science.¹ We believe that one main reason for this situation is that the theories of reductive-type identity physicalism that are available in the literature are not good and hence are easily dismissed.

Our task in this paper, therefore, is to offer a better theory of reductive-type identity physicalism that will withstand the main criticisms mounted against type identity physicalism in the literature, and that will do better than the varieties of non-reductive physicalism in solving major difficulties. We call our theory *Flat Physicalism* because in it there are no levels of existence or of explanation, and because it is physicalism *tout court*.

Flat Physicalism is an expansion of recent results in the foundations of statistical mechanics (see, e.g., Hemmo & Shenker, 2012, 2016; Shenker, 2017a, 2017b), an expansion that turns the scope of statistical mechanics from a theory originally aimed at reducing the special science of thermodynamics to the fundamental theory of mechanics² – to a general *theory of physical kinds*. Theories of *special sciences kinds* have been developed, some of them as part of the varieties of non-reductive approaches; but a theory of *physical kinds*, which takes into account not only the philosophical considerations concerning the notions of “kinds” but also some details of physics itself, thus bringing about a notion of physical kinds from within physics (as it were), is still called for, and is offered as part of our Flat Physicalism. Once a sound notion of physical kinds is in place, reductive physicalism can explain (i) the appearance (i.e., the seeming) of multiple realisation,³ and (ii) the appearance (i.e., the seeming) of an autonomy of the special sciences.

Our task can be seen as this: to explain how in a *fully* physical world multiple realisation and an autonomy of the special sciences *appear* to obtain. And in both (i) and (ii), our line of thinking consists of the following two ideas.

- (I) Because everything is physical, it follows that fundamentally, ontologically (truth-functionally, if you like), *all* the kinds are physical kinds, including *all* the special sciences kinds – full stop. Hence, ontologically there is *no* multiple realisation and *no* autonomy of the special sciences kinds and laws.
- (II) However, the so-called special sciences kinds are the physical ones that have the following two features: (a) These physical kinds are perceived by *us* (as human beings) as a result of our physical interactions with the environment plus the physical making of our brains (our measuring devices enable us to perceive additional physical kinds; different creatures may perceive different physical kinds); and (b) these physical kinds exhibit regularities. Physics fully accounts for both (a) and (b), and thus fully explains how, out of all the physical kinds, those of the special sciences are selected.

One interesting implications of this result is this. The reductive-physical account of (a) and (b) entails that it may happen – as it in fact does! – that the regularities exhibited by the special sciences kinds may be very different from the regularities exhibited in the fundamental physical theory (they may even appear anomalous in the Davidsonian sense!⁴). This is not a paradox but is rather a consequence; in fact, it is a well-known consequence in the science of explaining the laws of thermodynamics in statistical mechanics (as we explain below). All we do here is extend this known result and generalise it to all the special sciences.

So far, what we propose may seem to be an alternative to non-reductive physicalism, which is the standard account for (a) and (b). But this is misleading. What we also show in this paper is that so-called non-reductive physicalism is an *incoherent* view; therefore, it is not an option at

¹Smart (2017) is an overview of contemporary literature on identity physicalism.

²Thermodynamics has the status and structure of a special science, despite being part of physics.

³In cases where there is such an appearance as a matter of fact; see Polger and Shapiro (2016).

⁴See Shenker (2015, 2017c), Hemmo and Shenker (unpublished manuscript).

all. Showing this is a major task of this paper, and here is the idea very briefly. “*Non-reductive physicalism*” is the name of the idea that that so-called “high-level” kinds or properties *supervene* on *heterogeneous* “low-level” (especially physical) kinds or properties; that is, the “low-level” heterogeneous kinds can *multiply realise* the “high-level” ones. (Our present argument is not aimed at non-reductive approaches that rule out multiple realisability.) Metaphysically, this relation is often called “*realisation*,” but other sorts of metaphysical relations can be used, and we do not distinguish between them in this paper. And when we mention “realisation,” the reader can replace this notion with other metaphysical relations. As long as *supervenience is required and heterogeneous “low-level” kinds are allowed to fall under the same “high-level” kind*. Ever since this idea was suggested by, for example, Putnam (1967), Fodor (1974, 1997), and Davidson (1970), it became extremely influential, and it may not be an exaggeration to call it part of the contemporary mainstream in the philosophy of science and of mind.⁵ In this paper, we argue that *any theory that allows for multiple realisability (regardless, as we said, of the nature of the metaphysical relation involved) is a version of dualism, not of physicalism*; thus, the name “non-reductive physicalism,” to the extent that it refers to a theory that allows for multiple realisability, is a misnomer and even an oxymoron: *either non-reductionism (and allowing for multiple realisability) or (excluding or!) physicalism*. Moreover, it is a form of *token dualism* according to which *each token* contains non-physical elements. This is the claim we set out to argue for in this paper. And this is why we insist on adding the prefix “*so-called*” to the term *non-reductive physicalism*. (As we said, our argument is not aimed at non-reductive approaches that rule out multiple realisability. We think those are dualistic as well, but we do not address this point here.)

In sum, accounting for the appearance of multiple realisation and the autonomy of the special sciences in terms of so-called non-reductive physicalism is incoherent (or is simply dualistic); however, there is a full strong account of the fact that these empirically appear to be the case (to the extent that they do) within Flat Physicalism, which is a theory of type-identity reductive physicalism. This is the point we wish to make in this paper.

The paper is structured as follows. Section 2, after a brief presentation of our theory of physical kinds, Flat Physicalism, explains how the appearance of multiple realisation in the special sciences can be understood within this physicalist picture (assuming that multiple realisation obtains in our world) and what are the mental kinds in this picture. In section 3, we explain how Flat Physicalism accounts for the autonomy of the special sciences (in the above sense). Finally, to clarify our view and put it in the context of other relevant approaches, we compare (section 4) our view with a recent proposal by Bechtel (2017) that explains the appearance of levels in mechanistic explanations. Whereas mechanistic explanations are often understood within a non-reductive framework, we present Bechtel’s considerations to the effect that these explanations can be *better* understood in a reductionist framework, and then address some of his worries and arguments that lead him, nevertheless, to adhere to a non-reductionist view.

2 | PHYSICAL KINDS

Reductive physicalism and (the varieties of) so-called non-reductive physicalism differ (among other things) on whether the special-sciences kinds or properties are identical to physical kinds or properties. What are physical kinds, in this context?

Scientific disciplines frequently divide the particulars they study into kinds and theorize about those kinds. To say that a kind is natural is to say that it corresponds to a grouping that reflects the structure of the natural world rather than the

⁵See Polger and Shapiro (2016); but see Kim (2012) for criticism and references for this debate.

interests and actions of human beings. We tend to assume that science is often successful in revealing these kinds; it is a corollary of scientific realism that when all goes well the classifications and taxonomies employed by science correspond to the real kinds in nature. The existence of these real and independent kinds of things is held to justify our scientific inferences and practices. (Bird & Tobin, 2018)

Because the reduction of the special sciences to physics is of great interest in philosophy of science and philosophy of mind, theories, and notions of kinds in the special sciences have been vastly discussed in the literature. By contrast, the study of the notion of *physical* kinds is unsatisfactory, perhaps because it is (mistakenly!) thought to be trivial. We begin with some general discussion of kinds and proceed to present what physical kinds are and how *physics* takes them to be related to the special-sciences kinds, that is, a theory of special sciences kinds from within physics, as it were.

The theory we shall propose is inspired by statistical mechanics, which – as is well known – is usually confined to accounting for the special science of thermodynamics. Thermodynamics, despite the fact that it is usually seen as part of physics, has the status of a special science, and this is how its ontology is analysed within physics: it is understood as a generalisation from phenomena that is explained by the fundamental theory of mechanics (together with some auxiliary hypotheses, described in Shenker, 2017a, 2017b). The other special sciences, we propose, are to be understood along the same lines. Accordingly, we shall expand statistical mechanics to account for *all* the special sciences, and this will enable us to understand the relation between physical kinds and the special sciences kinds, and more generally the project of reduction of the special sciences to physics. (One difference between thermodynamics and the other special science is that the former’s kinds and laws happen to be much simpler than those of the latter, and therefore the reduction of thermodynamics to physics is manageable in *practice*; we submit that this is the *only* reason why it became part of physics, unlike the other special sciences, as we explain below.)

Instead of discussing “kinds” in general, we hope that the notion that we have in mind will become clear as we proceed.

2.1 | Microstates: Fundamental physics

Statistical mechanics is the name of the project of accounting for the thermodynamic regularities, taken to be generalisation of observation (e.g., as in Einstein’s, 1919, “principle theory”) by fundamental mechanics (e.g., take to be an Einsteinian 1919 “constructive theory”); here we shall focus, as an example, on classical mechanics.⁶ This project consists of two (related) stages. First, express the thermodynamic quantities (e.g., volume and temperature) in terms of mechanical quantities (like mass, position and velocity). Second, prove that due to the laws of mechanics (e.g., $F = ma$, together with some auxiliary hypotheses⁷) these mechanical quantities evolve in a way that mirrors the laws of thermodynamics (e.g., the second law). In this paper, we focus on the first stage (for the second stage, see Hemmo & Shenker, 2012, 2016; Shenker, 2017b, 2017c), and therefore we begin with the ontology of fundamental physics.

According to mechanics, the universe is at every moment in some well-defined state called a *microstate*.⁸ In classical mechanics the microstate consists of the precise positions and velocities⁹

⁶Much of the study of the foundations of statistical mechanics is carried out under the assumption that classical mechanics is the fundamental theory. Although strictly speaking, according to contemporary physics, classical mechanics is false; it arguably preserves some of the salient explanatory and predictive aspects of the true fundamental physics under the appropriate conditions. See Wallace (2001), section 1 and Ladyman et al. (2007) for a critical discussion of the use of outdated physics in philosophical discussions.

⁷These auxiliary hypotheses are described in Shenker (2017b, 2017c).

⁸This microstate can be thought of as an *event*. We do not expand on this point here.

⁹Or momenta, or generalised magnitudes, that take into account constraints. For the notion of instantaneous momentum, see Arntzenius (2000).

of all the particles.¹⁰ Given the parameters and constraints (e.g., masses and total energy) and the microstate at some moment, the equations of motion determine the microstates at all other moments. The following terminological remark is of utmost importance for understanding statistical mechanics. There are various notions of *microscopic* in the literature: the term sometimes means *small* and sometimes *part of a whole*. But in statistical mechanics it is customary to use the term *microstate* to denote the *complete* mechanical state of the system of interest, or of the world (at a moment¹¹). It is complete in the sense that (given the parameters and constraints) the theory cannot say more about the state of affairs in the world, and does not need more, nor settle for less, in order to predict and retrodict other microstates.¹² For instance, in classical mechanics the precise positions and velocities of the particles of the entire universe are a microstate despite the fact that it is neither small nor part of anything.

2.2 | Aspects: Macrovariables and macrostates

At first sight we might conclude ... that, as the number of particles increases, so also must the complexity and intricacy of the properties of the mechanical system, and that no trace of regularity can be found in the behavior of a macroscopic body. This is not so, however, ... when the number of particles is very large, new types of regularity appear. (Landau & Lifshitz, 1980, p. 1)

These regularities pertain to certain *aspects* of the microstates of the complex system given by partial descriptions of these microstates: the partial description (epistemology) pertains (if it is true) to an aspect of the microstate (ontology). For example, if in a sample of diluted gas (idealised as *ideal gas*) the energy is distributed among the particles according to the so-called *Maxwell-Boltzmann energy distribution*, then this aspect – regardless of other details of the microstate! – *determines* that if the gas is confined to an isolated container and one increases the pressure on it, its temperature will increase in direct proportion, in accordance with the ideal gas law. This ontological statement has its epistemological counterpart: if one *only knows* that in a sample of diluted gas (idealised as *ideal gas*) the energy is distributed among the particles according to the so-called *Maxwell-Boltzmann energy distribution*, and knows nothing about the other aspects of this gas, then from this *partial information* one can *infer* that if the gas is confined to an isolated container and one increases the pressure on it, its temperature will increase in direct proportion, in accordance with the ideal gas law. And so, the central idea of statistical mechanics is that, in order to provide a successful and informative account of thermodynamic phenomena, one needs only *partial information* about the microstate of a system.

Because the term *aspect* is central for our discussion, let us clarify its meaning (for us). An aspect of a microstate is an *ontological* notion, it is *part of* or *in* the microstate, but to explain it, it is easier to look at it from the epistemological perspective. A complete knowledge of the world is knowledge of its precise full microstate (possibly plus knowledge of parameters and constraints, which for simplicity we include here in the notion of microstate). Knowledge of an aspect of that microstate is a partial knowledge of that microstate. Every partial knowledge of the microstate is knowledge of an aspect of that microstate. If the microstate is represented mathematically, then an aspect might be thought of as a function of that mathematical representation; every function of the microstate is a mathematical representation of an aspect of that microstate (and the mathematical formulae represent the Maxwell-Boltzmann distribution are examples of this). Of course, the fact that a given microstate of the world satisfies the Maxwell-

¹⁰For the notion of microstate in quantum mechanics, see Shenker (2021).

¹¹Relativistic considerations need to be taken into account; we do not do this here.

¹²Constraints may decrease the number of degrees of freedom that are necessary for prediction.

Boltzmann energy distribution is a fact about the world, regardless of whether anyone knows this to be the case or bothers to represent it in any way. (If in the paper we mention an aspect of a physical system, what we mean is an aspect of the *microstate* of that system.)

The fact that an aspect (which is in the *ontology*) is given (*epistemologically*) by a partial description has the following implication, which is the fundamental idea in the foundations of statistical mechanics.

As an example to illustrate this important point, consider a *particular* sample of an ideal gas in a container, observed during some *particular* time interval from t_0 to t_1 , and seen to be in *thermodynamic equilibrium* throughout this interval, and in particular, the gas stably fills out the entire container with uniform distribution of pressure and temperature. We know from mechanics that during that interval the gas particles evolve through a continuous sequence of *different* mechanical *microstates*, and we know from the work of Maxwell and Boltzmann that in all of these microstates the energy is distributed among the particles according to the Maxwell-Boltzmann energy distribution.¹³ We shall say that all of the microstates of the gas during the observed interval *share an aspect*, namely, the aspect of the Maxwell-Boltzmann energy distribution. Of course, they *differ in other aspects*; for instance, they presumably differ in the precise positions of the particles.¹⁴

This last example illustrates that the notion of *aspect* has two characteristics, and their interplay is the core of statistical mechanics and the notion of a physical kind based on it.

- (i) Let us focus our attention, for example, on *one particular microstate* out of the above sequence that obtains at the point of time t_a between t_0 and t_1 : We know that this particular microstate has the aspect of the Maxwell-Boltzmann energy distribution. When we say that the system is at time t_a in the Maxwell-Boltzmann energy distribution, we are talking about *the aspect of this single particular microstate that actually obtains at t_a* , given by a partial description of this single particular microstate; and when we talk about, say, the volume of the gas *at t_a* , we refer to a function of the particular positions of the gas particles at t_a , as described in full detail in the particular microstate that actually obtains at t_a , and not to any other (counterfactual) microstate. The reason is, simply, that at t_a it is this particular microstate that obtains in the world, and since – according to mechanics – the microstate (together with parameters and constraints) is a complete description of the world, *there simply is nothing else we could refer to*.
- (ii) At the same time, we notice that this particular microstate, that actually obtains, also has (ontology) other aspects of which we (by hypothesis) know nothing (epistemology). Consequently, the partial information that we have concerning our t_a microstate is compatible with the system's being in any one of a continuous infinity of *counterfactual* microstates, that *share* this aspect but differ in other aspects. For all *we* know, it could have been any of those counterfactual microstates, and *we* would be able to tell the difference. In other words, all *we* know is that the actual microstate at t_a is a member of a *set* of microstates, one of which is actual (at t_a) and the rest are counterfactual (at t_a). Some of those that are counterfactual at t_a will obtain during other moments, but many will never obtain for that system. An aspect of a microstate (given by its partial description) is sometimes called *macrovariable*, and the set of counterfactual microstates that share an aspect is sometimes

¹³The Gibbsian picture of statistical mechanics differs from the Maxwellian-Boltzmannian one in important details (see Callender, 1999; Frigg, 2008; Goldstein et al., 2020; Uffink, 2004, 2007; Wallace, 2015), and therefore the predictions of these two approaches agree only approximately (see Jaynes, 1965; Werndl & Frigg, 2017). However, in both traditions thermodynamic magnitudes are associated with macrovariables or functions of macrovariables. See more on the connection between these pictures in Hemmo and Shenker (2012, ch. 11).

¹⁴The notion of aspect is in some sense perspectival. But, because this perspective is derived from the complete microstate by partial description, there is no danger of inconsistency between the perspectives, nor of not knowing "perspective of what," especially where theoretical entities are concerned (see Arabatzis, 2018).

called *macrostate* (see Hemmo & Shenker, 2016): this is how we shall use these terms here.¹⁵

We draw attention to the subtle notion of “description” in this context. As already Ben Menahem (2001) has shown, different descriptions, that in our terms here pertain to different aspects of the microstate, and that give rise to different partitions of the state space into the sets called macrostates, can give rise to qualitatively different behaviours, all of which exist together within the same microstate undergoing the same micro dynamics. This is the way to explain the appearance of regularities such as the time-directed second law of thermodynamics (see Hemmo & Shenker, 2012, 2019a). And to the extent that “a description” is understood epistemologically (and not as referring to the aspects being described), it assumes (implicitly) a descriptor: different descriptors will see different things in the same microstate under the same micro dynamics. Of course, if the picture is to be fully physical, the descriptor has to be fully physical itself; we return to this point later (see on Davidsonian descriptions Shenker, 2015; Hemmo & Shenker, unpublished manuscript).

These two characteristics of aspects are the idea that gives rise to statistical mechanics; all the ideas about macroscopic probabilistic regularities are built on them, as we now turn to see.

2.3 | Special sciences: A reductive account

To complete our account of the special sciences we now need to ask, how are certain mechanical aspects (macrovariables) associated with measured thermodynamic magnitudes? Consider for example the association of the thermodynamic magnitude of the *volume* of a gas, as measured by a volume measuring device, with the mechanical macrovariable *distribution of the positions* of the gas particles. The association can be analysed into four steps.

- a. By making certain reasonable *theoretical statistical and dynamical assumptions* concerning the gas, we infer that *if* the microstate of the gas has a certain aspect (e.g., a certain distribution of the positions of all its particles), then its particles closest to the walls will interact with the walls in a certain way.
- b. Hypothetically, suppose that we built a (new) device that (based on our understanding of micro mechanics) interacts with the gas particles close to the walls, and reflects the interaction predicted in (a).
- c. The new device is built in such a way that this interaction triggers a chain of reactions, by the end of which the device’s pointer reads “the volume is v .”
- d. If our new device reads “volume v ” in precisely the same cases that we (human beings) would say that the sample has volume v in some pre-theoretic sense of the term, then we would say that our new device is a volume measuring device. And then, if we think that the world is fundamentally described by mechanics – that is, in a reductive framework – then by the *name volume*, we shall then call the distribution of the positions of all of its particles, mentioned in (a).

The actual historical story is different, of course. We do not need this new device; all we need is to understand what old volume measurement is all about. But we hope that this little story helps to make our conceptual point. Notice that on this account the identification of “volume” with “positions of particles” is not a matter of conceptual analysis concerning the meaning of *volume*

¹⁵In the literature these terms are sometimes used interchangeably, or other terms are used. See various examples in Ehrenfest and Ehrenfest (1912), Sklar (1993), Lebowitz (1993), Goldstein and Lebowitz (2004), Frigg (2008), and Uffink (2007). In Hemmo and Shenker (2016), the notion of *macrovariable* refers to any aspect of a microstate, and *macrostate* refers specifically to the sets of those that appear in our experience.

and of *distribution of positions* because, for example, one can use the pre-theoretic term *volume* even if one believes that matter is continuous or has no theory of matter at all (as is indeed the case with thermodynamics). The identification is based on a conjecture as to the matter of fact in the mechanical universe that gives rise to our observations or measurements.

Other identity statements, concerning more complex aspects, such as “heat is the motion of molecules” (Kripke, 1980¹⁶), are to be analysed along the same lines. And here is a complexity that often arises in the special sciences (for which the idea of multiple realisation is needlessly invoked). *Motion of molecules* is a very general term; for example, the aspect (macrovariable) associated with temperature in an ideal gas (in which there are no interactions between the molecules) is different from the one in a van der Waals gas (in which there are the simplest interactions between the molecules), and the two cases agree quantitatively only in certain limiting conditions.¹⁷ The question arises: why are these two distinct aspects associated with the *same* thermodynamic quantity of “temperature”? This question is made even more difficult when we consider the notion of “spin temperature,” which is an aspect of the microstate as described in quantum mechanics. *What is the fact in virtue of which* these three cases are all subsumed under the term *temperature*? Is this fact a shared aspect of their microstates? It has been suggested that these cases are physically heterogeneous, and therefore they are subsumed under the kind of “temperature” in virtue of some other sort of matter of fact. What is this fact? How can heterogeneous cases form a kind? In particular, what makes special sciences kinds, in cases where the physical kinds that fall under them are heterogeneous? (See discussion of the case of temperature in Frigg & Hoefer, 2015, addressed in Hemmo & Shenker, 2019a).

To generalise this question, suppose that we are given some particular microstate J and are asked whether it is of the kind Q. There are two sorts of ways for us to answer the question.

- (1) There may be a list of the microstates fall under the kind Q, and J’s Q-ness is a fact *about this list* or about a relation between J and the list.
- (2) J’s Q-ness is a fact *about J itself*, knowable (ideally) by examining J alone.
(By the way, the same question and the same two possible answers arise with respect to the formation of kinds more generally, but here we focus on the case of microstates because we are within the conceptual framework of Flat Physicalism.)

Is the case of “spin temperature” one of case (1) or of case (2)? We now turn to explain the difference between these cases in a physicalist picture.

Should it turn out that the case is (2), that is, that the microstates of the ideal gas, the Van der Waals gas, and the spins system share some aspect that we associate with “temperature” (as in the (a)–(d) stages above), then the magnitude of “temperature” would be understood in terms of this shared aspect. It would be a physical kind.

But it has been suggested (e.g., by Frigg & Hoefer, 2015) that this is not the case; the microstates of the ideal gas, the Van der Waals gas, and the spins system *do not share any aspect* that we could associate with “temperature.” (Remark: Perhaps there is always *some* aspect in common; but if so, then multiple realisation is impossible by definition. Because we want to discuss the meaning of multiple realisation, we focus on *relevant* aspects that might explain why they are all cases of “temperature.”) How can we explain the subsumption of these heterogeneous microstates under the same kind? To understand this, recall that in the case of volume above we emphasised that we give the name “volume” to the aspect of distribution of positions *because* it is this aspect that brings about a certain pointer state of the measuring device. (Note: the pointer state of the measuring device is an aspect of the microstate of that measuring device!)

¹⁶See a recent defence of the Kripke-Putnam approach in Hoefer and Marti (2019).

¹⁷Heat and temperature are different concepts; for example, one is intensive and the other is extensive. But they converge *quantitatively* in an ideal gas.

Let us apply the same idea to the case at hand, *mutatis mutandis*: Here we shall have to say that the *same* pointer state of the measuring device (which – recall – is an aspect of the microstate of the device!) is brought about by interacting with the *different* aspects (macrovariables) in the microstates of the measured systems (the idea gas, the Van der Waals gas, and the spins system); thus – by hypothesis – the only fact that is shared in these cases is the pointer state, and it is due to this shared pointer state – which is a *shared aspect (macrovariable) of the microstate of the measuring device (but not of the microstates of the measured system)* – that all the cases belong to the same kind.

Consider now the notion of *extended system* and its *extended microstate*, which we present by way of an example. In our case, the extended systems are: the idea gas plus the measuring device (and their joint microstate in their combined state space), the Van der Waals gas plus the (same) measuring device (*ditto*), and the spins system plus the (same) measuring device (*ditto*, assuming that it is possible to come out with a suitable description in terms of quantum mechanics). The extended microstate, that is, the microstate of the joint systems in their combined state space, shares an aspect, namely, the pointer position. And it is due to the “pointer position aspect” that they fall under the same notion of “temperature.”

Disregarding the measuring device and focusing on the measured systems only, one gets the impression that this is a case of “multiple realisation” of the kind “temperature” by distinct macrovariables of the measured system, and this indeed seems *mysterious*: why should heterogeneous matters of fact fall under the same kind? But the mystery is removed once we realise that this is not a case of genuine multiple realisation but rather a case of a *shared physical aspect of the extended system* of measured and measuring systems.

The notion of an *extended system* or *extended state*, although simple, needs emphasis, and therefore we pause to elucidate it a bit more. Consider three physically *heterogeneous* microstates – A, B, and C – all subsumed under the kind H. (To illustrate: H is temperature; A is a microstate that has the aspect of average kinetic energy associated with temperature in an ideal gas; B is a microstate that has the aspect of a van der Waals gas associated with temperature; and C is a microstate that has the aspect in quantum mechanics associated with spin temperature.) Why are A, B, and C all cases of H? One option is that microstates A, B, and C share an aspect, which is H; however, this cannot be the case because by assumption A, B, and C are heterogeneous (and because they are microstates, they are complete with respect to their systems). Another option is that there is a device such that whenever it interacts with a system in microstates A or B or C, it ends up in a microstate that has the aspect O in which its pointer says “H.” And so, the extended microstates A + O, B + O, and C + O do share an aspect, namely, the pointer’s position. A, B, and C, in their own right, do not form a physical kind; however, the extended microstates do. If we ignore the device (as we often do), we have the mistaken impression that a “molto mysteriosos” fact is responsible for the subsumption of the heterogeneous cases A, B, and C under the same kind H (Fodor, 1974, 1997). Noticing the notion of extended system and its extended state does the trick, non-mysteriously. But then, when the mystery is removed, we are left with a physical kind, the kind of systems in which the pointer state is H; there is no multiple realisation.

One subtle point needs emphasis here: the device in the above scenario is *not strictly speaking a measuring device* because A, B, and C are physically heterogeneous; the fact that the device ends up in microstate O with its pointer saying “H” when interacting with all three of them cannot be due to its measuring some “H property” in them – there is not any such shared property, by hypothesis! What, then, explains this shared pointer state O? Within physicalism, because one is committed to there being only physical fact, there is no physical account for this coincidence apart from mere coincidence or the details of the initial condition of the universe (both options are naturalist versions of good-old accounts of psychophysical parallelism such as occasionalism or pre-established harmony). This solution, although consistent with the

principles of physics, is not ideal and sounds strange within scientific methodology. But we emphasise that there are *only two options at this stage*:

In virtue of what are tokens A, B, and C subsumed under the same kind H? There are two options: either they share a physical aspect of an *extended* token that comes about by sheer coincidence (and not by measurement), in which case the kind H is a *physical kind (no multiple realisation)*, or tokens A, B, and C share no physical aspect such that there is *genuine multiple realisation*. This case is discussed in the next section, where we show that this is a case of *token dualism*.

Let us connect the result thus far to three influential ideas in the literature, those of Putnam, Kripke, and Lewis.

We propose here that Putnam's famous *Twin Earth* thought experiment, in which two substances, XYZ and H₂O, causally bring about in the observer the same experience, should be analysed along these lines. *Pace* Putnam, the conclusion in this case is that, according to Flat Physicalism, "meanings are in the head"; the meaning is water; and the kind is water, not XYZ or H₂O, which are the references (so meaning does not fix a unique reference). Of course, one significant difference is that in our example above we make use of measuring devices, whereas Putnam discusses (presumably) human observers and the meaning of terms in language. We expand our analysis to the case of mental kinds below.

Another famous example of what Kripke (1980) calls *theoretical identifications* is the statement "water is H₂O" (see Hoefer & Marti, 2019). This complex example includes the statement that the stuff that we call water, which has a "characteristic feel, appearance and perhaps taste" (Kripke, 1980, p. 128), consists of H₂O molecules that interact with each other in a certain way such that each such molecule consists of particles that interact in a certain way. Importantly, water is not a mere "Lego"-like composition of particles (see Chang, 2012, 2017): the interactions responsible for the existence of each molecule and for their combination to form water are *fully* described by the theories of fundamental physics, and they form *part and parcel of the aspects of the fundamental physical state of affairs* that we call *water* (mereological accounts are mistaken in this context). We emphasise that the physical account here is all that one needs to explain H₂O molecules and the way they combine in water and affect our senses: there is no residue that calls for explanation in terms of other metaphysical relations such as grounding, determination, and realisation.

Lewis (1994) distinguishes between "natural" and "unnatural" properties, the latter illustrated by the examples of being a pub, a club, or a tub. One important implication of Flat Physicalism concerns the distinction commonly used between properties that appear in the sciences (e.g., temperature) and properties that seem to depend on human interests or culture: it is sometimes said that the latter cannot be accounted for in terms of physical aspects (macrovariables). But in a physicalist picture, there is no room for this distinction. All the properties that appear in our experience, including what Lewis takes to be "unnatural" ones, are to be explained by full reduction to physical macrovariables. The cases of Lewisian "unnatural" properties are likely to be accounted for in terms of *extended system* that include the human observer, as explained above, and in this sense *all of them equally carve nature at its joints*. In a reductive theory, there are no properties other than physical macrovariables. The kinds that appear in the special sciences are macrovariables that may appear to behave in ways that are not easily described in physical terms, but that is a pragmatic difficulty that does not indicate fundamental anomaly (we expand on this in section 3 below, see also Hemmo & Shenker, 2015a, 2015b). This in no way implies that these macrovariables are "unnatural" or "less natural" than, say, thermodynamic macrovariables that no one would say are unnatural.

2.4 | Genuine multiple realisability entails *token dualism*

Suppose that our three token microstates above, A, B, and C, all of which are subsumed under the so-called "high-level" kind H, turn out to be fully completely and genuinely heterogeneous

in the sense that there is not even an extended system that could explain why they are all cases of H. We call such a case *genuine multiple realisation*. Presumably, this is the case envisioned by Putnam (1967), Davidson (1970), Fodor (1997), and more. What then? Assuming that such a case is indeed found empirically, how are we to explain it? In order to describe our account of such a case, it is helpful to begin by asking two questions concerning this case.

Question (i) What facts make it the case that a given microstate A (or B or C) is a case of high-level property H rather than another, say L?

Question (ii) How can the high-level kind H (should that be the case) be identified, for instance, by us? It is a fact, after all (presumably!) that we are able to say of a given particular case that it is a case of the high-level kind. How do we come to know, when A (say) obtains, that it is a case of H rather than L?

Recall that above we said that there are two sorts of ways for us to answer the question of whether the token J has the property Q:

- (1) There may be a list of the things that are Q, and J's Q-ness is a fact *about this list* or about a relation between J and the list.
- (2) J's Q-ness is a fact *about J itself*, knowable (ideally) by examining J alone.

In the present discussion, if multiple realisation obtains, then we must rule out option (2) and are therefore left with option (1). Because, *by assumption* there is no physical fact in the microstate A, no aspect (macrovariable) of A that can provide an answer to **Question (i) and Question (ii)** above, a remaining option might be that there is some *list* of all the physical microstates that realises the high-level kind H, and then the answer is:

Answer to Question (i): An item that belongs to the list is of kind H; and

Answer to Question (ii): Knowledge of the list is the way to know whether an item is in the list.

And here is a dilemma. Either the list is part of physics, or it is not. If the list is physical, then we should add it to the microstate, thus obtaining an *extended microstate* – and then the multiple realisation would be only *apparent*, not *genuine*. Flat Physicalism would hold. If, on the other hand, the list is not physical, then we have a case of *token dualism*: the non-physical fact of the existence of the list must hold in each and every token case in which we know whether A is H (and it is a *fact* about us and about our experience that, upon encountering a particular cases, we *are* able to know whether they are of the high-level kinds of interest). And so multiple realisation, if genuine, is a case of *token dualism*.

Another way to see this is by invoking Laplace's Demon, who by assumption has access to all the possible and actual fundamental physical facts. Laplace's Demon is able to know whether a given microstate A is of kind H only in case the fact that makes A be of the kind H is the sort of fact accessible to the Demon, namely, an aspect of the (extended, if you will) microstate. If, on the second horn of the dilemma, the fact that makes the microstate A case of the "high-level kind" H is not in the (extended) microstate A; it is not accessible even to Laplace's Demon, and this is a case of *token dualism*.

Notice that nothing in the above argument depends on the fact that microstates are *instantaneous* states (Arntzenius, 2000). The same analysis and argument (*mutatis mutandis*) apply for *temporal sequences of microstates*, which are taken to multiply realise high-level kinds in functional theories, such as the computational and causal theories of mind.

This sort of token dualism holds even if one takes it that the list is learned only a posteriori through experience. Multiple realisability implies token dualism regardless of whether the list is learned a posteriori; in order to build the list on the basis of experience, upon encountering a

particular case one must be able to experience whether this case is of the kind H, in which case one should add it to the list. And as we argued above, if the list is not physical or is not fixed entirely by the physics of its tokens, including the physics of the (token) experience of a token, learning the list by experience presupposes that one has access to a *non-physical* fact. The alternative is that upon encountering a particular case, one decides for no reason at all whether to add it to the list. In that case, different agents will in general have different lists, and the lists will become similar only by mutual learning. This, however, is (too) a case of an extended microstate, albeit a more complex one.

Fodor (1974) thinks that the items in the list do *not* share anything physical, and because on the one hand he is aware that the way the list is built is inexplicable and in particular has no physical explanation, and on the other hand he is reluctant to admit dualism explicitly, he calls the way the list is built a *mystery*. He says that the multiple realisability of special sciences kinds by physical kinds is just a *brute fact* (“molto misterioso” Fodor, 1997, p. 1961) that has no further explanation. From the perspective of Flat Physicalism, this view is a straightforward metaphysical dualism (see also Hemmo & Shenker, 2015a, 2015b, 2019a).

In sum, in virtue of what are tokens subsumed under the same kind? There are two options: Either they share a physical aspect (either in them or in an *extended* token), in which case the kind is a *physical kind (no multiple realisation)*, or there is *genuine multiple realisation*, in which case these tokens share a non-physical fact that is somehow accessible to us when we “know” that a given token belongs to a given kind, and this is a case of *token dualism* because the non-physical fact must obtain within each of the tokens.

Whether there are cases in the actual world is a matter of methodological and empirical controversy (see, e.g., Polger & Shapiro, 2016). We take it, however, that if as a matter of fact there are cases of *genuine* multiple realisation, then physicalism is empirically refuted and metaphysical dualism prevails.

2.5 | Mental kinds

Recall for a moment the notion of an extended token: Suppose, for example, that the token microstates A, B, and C are physically heterogeneous and nevertheless are subsumed under the same “high-level” kind H – and this is explained by observing that there is another system with microstate O with the same aspect (macrovariable) whenever any of A, B, and C obtain such that the extended microstates A + O, B + O, and C + O share the macrovariable of O in all of these cases. We mentioned that in this case A, B, and C may *seem* to be (mysteriously) of the same “high-level” kind despite being heterogeneous, but this is a mistake made by ignoring O. We now add that this mistake is made frequently, and understandably, when O is an aspect of the microstate *of us*, human observers; the pointer state is our mental state in which we perceive that we have a case of the kind H. If A, B, and C are in our environment and we observe them, it is natural to search for the explanation in the environment, ignoring the fact that it is us who make them into a kind. This is most clearly seen where the kind to which the physically heterogeneous tokens A, B, and C belong is culturally dependent (e.g., three physical realisations of a stop sign¹⁸): nothing connects them, except us. The next step in our analysis, then, needs to explore the physics is “us” in this context.

Kripke (1980) argues that, in the identity statement “heat is molecular motion,” *heat* refers to the state of affairs in the world that (causally) brings about in us a sensation that we call (for lack of a better name) *sensation of heat*. The state of affairs called *heat*, he says, is – as discovered scientifically – molecular motion. And, to complete the picture, what is the state of affairs called *sensation of heat*? In a physicalist framework, this element should be accounted for in

¹⁸An example we came to think about while discussing this problem with Yemima Ben Menahem.

physical terms. Kripke (1980) famously thinks that this is impossible and therefore that the identity statement “the sensation of heat is physical state (or process) x ” is false (for any x). Let us address this point in the terms of Flat Physicalism.

The statement “heat is molecular motion” is understood by Kripke by reverting to the “sensation of heat” which is brought about (in the observer) by molecular motion (in the observer’s environment). If we apply the same schema of analysis to the statement “the sensation of heat is physical state (or process) x ,” we end up in a regress because we shall conclude that the term *sensation of heat* should refer to the causal trigger of a “*sensation of the sensation of heat*,” and so on ad infinitum. To avoid this result, one may decide to end the causal chain that starts at the molecular motion and continues in some brain (or other physical) state at the stage in which the “sensation of heat” occurs, thus *departing from the above schema* in one of the following ways:

Either: The fact which is a “sensation of heat” lies *beyond* the physical chain of causal interactions.

Or: The fact which is a “sensation of heat” is *part* of the physical chain of causal interactions.

(We emphasise that, at this point of the argument, other options often mentioned in the literature, such as that the “sensation of heat” merely somehow supervenes on the physical, are not available to us because that would lead us to an infinite regress, which we are now trying to stop (or to prevent). The argument for why, at this stage, opting for mere supervenience leads to infinite regress is provided in the previous section.)

In the first of the above two options, the mental state of “sensation of heat” may be causally or otherwise *correlated* with brain states, but it *itself* needs to be accounted for in some non-physical terms, together with their “non-physical dynamical” laws concerning how the correlation with physical brain states comes about. This is a well-known problem for dualism, and in this paper we focus on physicalism and therefore we do not expand on this line of thinking here.

We focus on the second option, in which the term *sensation of heat* refers to an *aspect* (or *macrovariable*) of the physical microstate of the brain (Here, we use the term *brain* to denote whatever in the body or elsewhere in the physical world is relevant to the mental states and processes). This is a type-identity reductionist theory of the mind: mental states are (identical) with aspects of microstates (given to us by partial descriptions of these microstates).

With regard to what is “type-identity reductive physicalism,” it is important to note that “type identity” often is understood (e.g., ever since Putnam, 1967) as requiring that the *entire microstates* be identical. This is simply a *bad theory* of physicalism, which is (moreover) not in line with the scientific research of the brain, nor is it in line with well-known ideas in physics in the foundations of statistical mechanics. By contrast, according to *Flat Physicalism* – a type-identity theory that is in line with actual science – *type-identity means shared aspect*. Two microstates are type-identical if (and only if!) they share the same aspect (given by their partial description and including the aspect of extended systems), and this aspect is identical with the special science kind in question (see more in the next section on special sciences laws). This is the essence of Flat Physicalism, and it leads to solving all the problems faced by bad physicalist theories.

Kripke (1980) argues that identity statements like “the sensation of heat is brain macrovariable H ” are false (because they hold between rigid designators, and a “possible world” in which the sensation of heat occurs, whereas the macrovariable H does not, or vice versa, as emphasised in Chalmers’s 1996 *Zombie* thought experiment, is conceivable, and hence – on Kripke’s view – possible.). In our view, Kripke and Chalmers beg the question on this point. In the first (dualist) option above, “brain macrovariable H without a sensation of heat” (and vice versa) is conceivable; in the second option, it is *inconceivable* (i.e., it is conceivable only in the sense of mistakes, e.g., in the same sense that one can make arithmetic miscalculations and

believe they are correct).¹⁹ Identities, when discovered and if true, do not require explanation, and their negations are *inconceivable* (hence the question of whether conceivability entails possibility does not even arise). We agree with Papineau on this point:

Consider this parable. Suppose that there are two groups of historians, one of which studies the famous American writer Mark Twain, while the other studies his less well-known contemporary, Samuel Clemens. The two groups have heard of each other, but their paths have tended not to cross. Then one year they both hold symposia at the American Historical Association, and late one night in the bar of the Chicago Sheraton the penny drops, and they realize they have both been studying the same person. At this stage there are plenty of questions they might ask. Why did this person go under two names? Why did it take us so long to realize Mark Twain and Samuel Clemens were the same person? But it doesn't make sense for them to ask: why were Mark Twain and Samuel Clemens the same person? If they were, they were, and there's an end on it. (Papineau, 1993, pp. 180–181)²⁰

This holds, in particular, with respect to the hypothetical discovery that “the sensation of heat is brain macrovariable H,” no less than the discovery that “water is H₂O,” despite the abovementioned difference between the schemas that give rise to them.

An important implication is that genuine multiple realisation of mental types by physical types are *impossible* in a physicalist picture (recall, in this sense, so-called non-reductive physicalist views are forms of token dualism). The option of “omitting the observer” to account for an alleged *appearance* of multiple realisation is of course unavailable here because we are already talking about the observer. (An observer-of-observer is not only the beginning of an infinite useless regress but also does not make sense because it implies that you, the reader, have a certain mental state in virtue of the fact that someone else, or a homunculus, is observing you.)

Should it turn out that as a matter of empirical fact certain mental types are genuinely multiply realised by certain physical types, then Flat Physicalism with respect to the mental will end up being *false*: one would have to opt for ending the causal chain beyond physics, as in the first option above, and then metaphysical dualism (substance, properties, or what have you) would turn out to be true of the world.

3 | REDUCTION AND THE LAWS OF THE SPECIAL SCIENCES

As we have shown in section 2 above, to ensure reduction to physics one must assume something like Flat Physicalism, a *type–type identity* theory, in which the special sciences kinds and properties are associated with aspects (macrovariables) of physical microstates (or with sets of microstates that share the same aspect/macrovariable, sets called *macrostates*). The question we wish to examine now concerns the status of the laws of the special sciences in this fully reductive theory. Fodor (1974) famously emphasised that one argument in favour of a non-reductive picture is the fact that the special sciences progress *independently* of physics, and (moreover) the forms of their laws appear to be very distinct from the forms of the laws of physics; the special sciences are *autonomous* in these two senses. Our task is to show that these observations are completely compatible and can be *better* explained by Flat Physicalism. In particular, we will now show how special sciences laws that have forms that are very different from those of the

¹⁹An interesting case here are probabilistic algorithms for finding big prime numbers.

²⁰For various notions of identity, see Noonan and Curtis (2017).

laws of physics can come about. This will explain the sense in which the special sciences are *autonomous* despite the fact that they are completely *reducible* to the laws of physics.

3.1 | The autonomy of the special sciences in Flat Physicalism

Because special sciences kinds are (identical to) aspects of microstates, and the special science property that is observed at a given time t_a is (identical to) that aspect of the actual microstate that obtains at t_a (because this microstate is everything that there is at t_a !), the following question arises. Every microstate (e.g., the one that obtain at t_a) has infinitely many aspects (given by infinitely many possible partial descriptions), but only some of them appear in our actual experience and in our theories (e.g., the thermodynamic macrovariables). What is special about these aspects? What makes them preferred over the other macrovariables of the same microstate that obtains at t_a , and that do not appear in our experience and theories? Here, we need to pay attention to two special features of some of the aspects of the microstates:

- (1) We (human beings) are physical systems that interact physically with our environment, for example, via our sense organs. Due to our particular physiology, we interact (physically) with our environment in particular ways, and because of this our sense organs are sensitive to certain aspects of our environment and not to others. The proverbial Laplacian Demon, who knows the complete microstate and micro-trajectory of the universe, if it is told which system is “us” and which system is “our environment,” can calculate which are the aspects of our environment to which we are sensitive. This is one feature of the aspects of the microstates of the environment which makes them special *relative to us*. Other creatures can be – and are! – sensitive to other aspects of the environment (e.g., some animals see a different spectrum of light than we do). We can build measuring devices that are sensitive to additional aspects of the environment and “translate” them to aspects that we can perceive, thus making those aspects indirectly accessible for us, and adding them to the aspects of the environment to which we can empirically relate and that therefore feature in our experience and theories.
- (2) The microstates of the universe evolve according to the laws of physics. As they evolve, their aspects evolve with them (trivially). Presumably, only some of the infinitely many aspects of any given microstate satisfy regularities: those include the thermodynamic aspects, but maybe there are additional aspects of the world’s microstates that exhibit regularities as they evolve along the world’s micro-trajectory, and because we are not sensitive to those aspects – in the sense of (1) – we do not observe their regularities. Among the aspects of the world’s microstates that we can perceive as explained in (1), we are naturally especially interested in those that exhibit regularities. The task of the special sciences is to pick out those aspects and describe their regularities.

In this sense the special sciences are *not* independent of physics: they are part of it; they are *in principle derivable* from it. Nevertheless, as we said, the special sciences can and do progress in a de facto autonomous way: you do not need to follow the details of physics in order to study biology or economics; in fact, the study of physics will not help you at all in studying those special sciences and will even hinder their progress. Why is that so? What is the notion of autonomy that is relevant to understanding the special sciences, given this reductive picture?

The idea is, simply, this. Concerning (1), in order to find to which aspects of the environment we are sensitive, we do not need to make physical calculations or to ask Laplace’s Demon; we know these aspects directly from our very experience. To see this, it is important to recall the nature of the identity statement, and we repeat it by way of an example. We know from experience what temperature is, even if only approximately and not quantitatively; and from

statistical mechanics we know that temperature (of an ideal gas in equilibrium) *is identical with* average kinetic energy. This means that we know from direct experience that we are sensitive to average kinetic energy, even if we cannot spell this mechanical quantity explicitly and we never heard about the science of mechanics. Concerning (2), among these accessible aspects, only some satisfy regularities; discovering those is a highly non-trivial task to be carried out empirically. And this is the job of the special sciences. Again, we need not go through physics to make these discoveries, and a famous example is the special science of thermodynamics that describes regularities in terms of quantities such as temperature, pressure, volume (and adds theoretical terms like *entropy*; we do not go into this point here).

In this sense, the special sciences are *epistemologically* autonomous despite being *ontologically* reductive to physics. This epistemic autonomy is very strong for two reasons.

First, to illustrate, a case in which there is no such autonomy is the special science of thermodynamics, and the reason why it is not autonomous is that the connection between its terms and laws and those of the fundamental physics are very simple. In other special sciences, the complexity is enormous and prevents such a reduction in practice.

To the immense complexity of the world, we need to add here the difficulty coming from the fact that we ourselves are part of the world, and the very act of making predictions is itself a physical process in our brain, that is, in the world that is to be predicted.

Thus, it is the job of the special sciences to teach us which are the aspects of the world we are sensitive to, what regularities do they satisfy, and which other aspects – accessible only via complex measuring devices – satisfy interesting and useful regularities. *Physics alone cannot do this, despite the fact that everything is physical.* This fact underlies the autonomy of the special sciences, which is epistemic and nevertheless extremely strong (see also Hemmo and Shenker, 2015).

Thus, in Flat Physicalism the autonomy of the special sciences is fully explained and is not mysterious. By contrast, for thinkers like Fodor (1974, 1997) who insist on making the autonomy ontological, the autonomy of the special sciences is “molto misterioso.”

3.2 | Special sciences laws that differ in form from those of physics

Here are some examples of forms of special sciences laws that seem, on the face of it, to be so different in the form of the regularities they describe from the laws of physics that one may be tempted to see them as supporting a version of non-reductive ontology. However, all three have full accounts within the reductive picture of Flat Physicalism.

- (i) The Second Law of thermodynamics is *time-directed*, despite the time-symmetry of all the laws of fundamental physics.²¹ This directedness can be accounted for in terms of non-temporal local asymmetries (see Hemmo & Shenker, 2019b).
- (ii) The laws of statistical mechanics are probabilistic, and despite the deterministic nature of the laws of fundamental physics, the statistical mechanical probabilities describe objective features of the world (see e.g., Callender, 2007; Dürr et al., 1992; Frigg & Hoefer, 2015; Hemmo & Shenker, 2012, 2016, 2012, 2016; List & Pivato, 2015; Loewer, 2001, 2012; Maudlin, 2007).²² In Flat Physicalism, this is accounted for in terms of the harmony

²¹The charge-parity-time (CPT) theorem and the violations of time-symmetry and charge-parity symmetries in the quantum electroweak theory are considered irrelevant for the workings of the brain because of the high level of energy at which these violations occur. But, even if they are relevant, it is conjectured (see Atkinson, 2006) that the origin of these violations (from which the CPT theorem follows) is in the low-entropy past hypothesis introduced in (quantum and classical) statistical mechanics. If this conjecture is true, the fact that the CPT theorem originates in the low-entropy past is another example for what we call the *autonomy of the special sciences!*

²²On some versions of quantum mechanics, the fundamental laws are probabilistic. But the quantum probabilities are different from the probabilities that appear in statistical mechanics; see Hemmo and Shenker (2012, appendix).

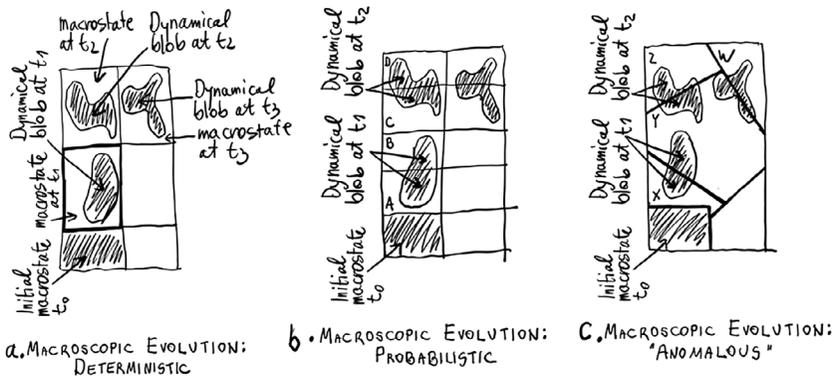


FIGURE 1 Dependence of *macroscopic* evolution on the interplay between dynamics and partition of the state space

between the micro-trajectories and the partition of the state space into sets according to the aspects (macrovariables) to which the measuring devices are sensitive in the sense explained above; see Hemmo and Shenker (2012, 2019a). This is illustrated in Figure 1a,b. The grey areas depict the bundle of micro-trajectories that start out, at time t_0 , from the bottom left region in the state space, as it evolves in times t_1 and t_3 (satisfying Liouville's theorem).²³ The state space is further partitioned to sets of microstates that share an aspect to which a given measuring device is sensitive. In Figure 1a, the aspects to which the device is sensitive happen to be such that the device will detect a deterministic behaviour, whereas in the case of Figure 1b the measuring device is more sensitive; that is, it can distinguish between finer values of the aspects of the microstates, and the result is that this device will detect a probabilistic behaviour: at time t_1 the system will have either property A or property B, with some probabilities, and at time t_2 it will have either property C or property D, with some probabilities.

- (iii) As is well known, Davidson (1970, 1993) conjectured that the mental is anomalous, and offered a model in which that might be possible (for a critical analysis of Davidson's approach, see Hemmo & Shenker, unpublished manuscript; Shenker, 2015). Whether the mental is anomalous is a question of fact, and we do not purport to answer it. Suppose, for the sake of the argument, that the mental is anomalous. Flat Physicalism is *committed* to providing an outline of how to go about providing a reductive fully physical explanation of anomalous facts in a world governed by the laws of physics. How can Flat Physicalism do this?

Consider Figure 1. Each of the three parts of the figure depicts the state space of some system (the same system in all three cases), in which every point represents a microstate of the system. In each of the three parts of the figure, the state space is carved up into sets of microstates in a different way. Each such set is a macrostate, that is, it is a set of microstates that shares some aspect (or macrovariable). In each of the three parts of the figure the carving up is different, reflecting the fact that each of them represents *different* aspects of the *same* microstates. For simplicity as well as for emphasising our point, the starting macrostate and the dynamics in all three cases 1a, 1b, and 1c are the same, and only the partitions of the microstates into sets are different in the three cases. At the initial time t_0 , the system is *prepared* in such a way that the *actual microstate* at t_0 has a given aspect, shared by all the *counterfactual* microstates in the set in the bottom left rectangle; we say that the system is prepared at t_0 such that its *actual*

²³The sequence of grey rectangles in Figure 1 are the Poincaré sections of the bundle of trajectories of the time evolution of that initial set of microstates, at several points of time.

microstate belongs to the corresponding *macrostate* (which is, recall, a set of microstates that share an aspect). We now want to predict the evolution of the system. We do this in three steps.

- Step 1: Dynamics.** Ideally, for every microstate in the initial macrostate we calculate its future microstates at t_1 , t_2 , and t_3 . The sets of point to which the initial microstates are mapped are depicted by the coloured blobs in all three parts of Figure 1.
- Step 2: Macrostates.** We partition the state space in sets corresponding to the kind of aspects that we would like to measure. For example, we might be interested in the temperature of the brain in one case, in its neuronal state in another, and in the distribution of a given hormone in the third, and so one.
- Step 3: Dynamics-macrostate harmony and probabilities.** In the final state, we measure the size²⁴ of the overlap between the dynamical blobs and each of the macrostates at times t_1 , t_2 , and t_3 . This size is the probability of finding the system, in those times, in each of the macrostates, which means, in microstates having each of the corresponding aspects. (For more details on this kind of probabilistic prediction, see Hemmo & Shenker, 2012, 2019a).

With this schematic procedure in mind, consider the difference between Figure 1a–c. In Figure 1a, there is a nice harmony between the dynamics and the macrostates: at each moment we are certain to find the system in a given macrostate. In the case described by Figure 1b, the macrostates are more finely grained. Therefore, we can only make probabilistic predictions. For example, at t_1 the system will have either aspect A or aspect B, with probabilities given by the overlap of the dynamical blob with the sets A and B. Figure 1b illustrates the kind of cases discussed in the statistical mechanical account of thermodynamics.

In the third case (Figure 1c), there is even less harmony between the dynamics and the partition, and there may not be any regularity – not even a probabilistic one – in the evolution of the aspects that fix this partition. This case may appear to be anomalous with respect to the evolution of the aspects to which the measuring device is sensitive. Another case of even stronger anomaly (not depicted here) is one in which the partition can remain as, for example, in Figure 1a, but the dynamics is different (call it the *alternative dynamics*) so that the result is as in case 1c: the alternative dynamics in this case may lack *any* macro-regularity of the aspects described by the 1a partition. Interestingly, given the alternative dynamics, which appears to result, under the 1a partition, in no regularity, it may be possible (a fact to be determined empirically) that there is some other alternative partition to aspects (e.g., the 1c one, or even a more complex-looking partitioning) under which the alternative dynamics will exhibit an interesting regularity.

Discovering the right partitions and making the calculations is a highly non-trivial task, and it may be that due to the sheer complexity involve we shall never find them out. Nevertheless, this schema is important because it explains how seemingly anomalous phenomena can come about in a world governed by the laws of mechanics; it explains how the world as we know it can be fully physical.

4 | ARE MECHANISTIC EXPLANATIONS COMPATIBLE WITH REDUCTION TO PHYSICS?

The picture we propose in Flat Physicalism is flat in the sense that according to this theory there are *no levels of reality*: So-called high-level kinds (e.g., temperature and water) are understood here as aspects, or macrovariables, given by partial descriptions of the *microstructure* of a single level of entities and their interactions, as described by the fundamental theories of physics.

²⁴Using the right measure; we do not address here the problem of the choice of measure; see Hemmo and Shenker (2012, 2015b).

An influential approach to explaining biological and other processes and system in contemporary literature is the mechanistic approach in which the notion of levels is central. To discuss the way Flat Physicalism sees the mechanistic approach, it is useful to consider a recent account of this approach by one of its proponents, Bechtel (2017). According to Bechtel (2017), the appearance of mechanistic levels in biological and physical explanations is a feature of using a causal network picture in which the entities that are basic (in the context of a given mechanism) are represented by *nodes* connected by causal links, and clusters of nodes connected by causal links that satisfy certain criteria form *modules*. Due to the criteria for the formation of modules, one may get the impression that they are high-level entities in the sense that the causal connections between them exhibit regulations that seem independent from those holding in the lower levels. But, as Bechtel shows, a precise analysis of the network reveals that this is not the case: fundamentally, it is all a matter of basic nodes, their causal links, and their clustering according to the criteria. Therefore, Bechtel says,

The network representation seems to favor a highly reductionist account that represents all activity at one lowest level. (Bechtel, 2017)

Indeed, he uses the phrase *flattening levels*. But he hastens to emphasise that “This [reductionist] interpretation, however, is mistaken” for three reasons:

First, the nodes in a network need not belong to a common level in any of the standard senses. ... Second, although in any graph representation there will be a set of nodes that correspond to what are taken as the basic entities, they should not be treated as representing entities at some base level. ... Third, in developing a graph representation, one might deliberately represent a set of entities as a single node. (Bechtel, 2017, pp. 269–270)

For Bechtel, the main interest in reduction or in non-reduction is epistemological or even methodological or pragmatic rather than metaphysical, and therefore he is reluctant to draw any ontological conclusions from his analysis.²⁵ In this sense, we may go beyond his intentions when analysing these ideas metaphysically. We nevertheless find that such an analysis is instructive for understanding the idea of mechanistic explanation within Flat Physicalism.

It seems to us that if Bechtel’s proposal is interpreted along the lines of (or combined with) our Flat Physicalism approach, then these three problems do not arise (and then whether one endorses reductionism will depend on other considerations). We suggest understanding Bechtel’s picture as follows. The grouping of elements (of nodes, and then of modules) reflects the fact that they share a macrovariable (an aspect of their physical state), and it is this macrovariable that interacts with other macrovariables, that is, in Bechtel’s picture, with other modules. It is these macrovariables (aspects of the physical states) – rather than the full entities and their full microstates – that participate in the causal network and satisfy regularities.²⁶ The causal relations between the modules are the physical interactions between macrovariables. Importantly, the macrovariables here include the physics of the interactions (see Chang’s (2017) “no Lego” argument), and indeed Bechtel’s modules contain the entities as well as their causal relations. The picture is that of relations that hold between macrovariables, either qualitatively different or with different degrees of coarse graining, but all of them are mere aspects of the single ontological level given by different ways of partially describing it.

Bechtel’s first objection (above) to reductionism is met by noticing that macrovariable in this picture can differ in the degree of coarse graining and other features. The second objection

²⁵As he explains in discussions of this paper.

²⁶See Frisch (2015) and Ben-Menahem (2018) for recent analyses of notions of causation and their relation to physics.

is met by realising that in every context different macrovariables are the most relevantly basic aspect of the microstate. The third objection is met by realising that one can focus on different macrovariables of the same microstate in different contexts.

Consider two physical examples that Bechtel discusses. One is that of a marble sliding on an inclined plane, which constraints its motion. In our proposal, this case should be thought of as an interaction between certain macrovariables of the microstate of the marble (e.g., the centre of mass of its particles) and some macrovariables of the microstate of the plane (e.g., the positions of its particles, or the quantum effects that make the plane impenetrable). Because these are partial descriptions of the microstates of the marble and the plane, the picture is flat: there are no levels of reality; the constraint and the constrained are on a par; and ipso facto the (causal) interactions are not “top-down.” A similar analysis applies to the paradigmatic case of a gas in a container in statistical mechanics, where the distinction between the gas and the constraints dissolves in small systems, for example, in Feynman’s picture of Maxwell’s Demon (Feynman et al., 1975, chap. 46).

Another example that Bechtel brings is the patterns of convection in Rayleigh–Bénard cells, in which every molecule is constrained to move only within its vortex. This example can be analysed in physics by applying a method analogous to the Hartree-Fock approximation used for heavy atoms: treat each molecule in the liquid as if it is constrained by the other molecules, but bear in mind that this holds for *each* molecule, and therefore it is essential for this treatment that the system and constraint are treated on a par: the interaction holds between certain macrovariables of the molecules. The picture is flat. There are no levels of reality.

5 | CONCLUSION

In this paper, we have outlined a reductive *type–type identity* theory for all the special sciences based on ideas recently developed in the foundations of statistical physics (classical or quantum). This theory is called *Flat Physicalism*. We take it to be a theory of physical kinds as well as a theory of all the natural kinds within a physicalist framework. We have argued that if the world is entirely physical, that is, if there are no non-physical facts, then all the natural kinds that appear in the special science, including mental kinds in psychology, just are physical kinds. We have further argued that all these special-cum-physical kinds are physical macrovariables, which are functions (or aspects) of the *complete physical microstate of the world and nothing else*. Contrary to the received wisdom usually called *non-reductive physicalism* (including functionalism of all sorts), we have shown that reality is flat in the sense that there are no levels of reality because physical microstate and evolution over time are all there is to the world. It is no surprise that no questions of top-down or bottom-up relations (e.g., causation, multiple realisability, or what have you) come up in this picture because the relation between the special sciences kinds and physical kinds is that of strict identity. This means that all the special sciences (e.g., biology, psychology) are ultimately branches of physics describing physical kinds which are entirely and exhaustively fixed by the microstate of the world, being nothing more than partial descriptions of it. Nevertheless, the special sciences are autonomous in an important sense, and their role cannot be taken by physics because it is the task of the special sciences to provide the complete inventory of the physical kinds that we encounter in our experience and that are significance for describing our experience in the world, and the ways they behave over time that are relevant to their domains of applicability; most of these physical kinds are yet to be discovered. We have shown (see also Hemmo & Shenker, 2019a) how – *despite this reductive picture* – the laws of the special sciences can, and should be expected to be, radically different from the fundamental physical laws (e.g., they can be chancy, even if the fundamental physical microstate satisfies a deterministic equation of motion). We have further argued that every theory in which multiple realisability holds entails token dualism, and therefore that the prevalent view called *non-reductive physicalism* is a form of *dualism* (substance or property) in disguise.

ACKNOWLEDGMENTS

We are grateful to the helpful discussions with the participants of the workshop on The Multi-Level Structure of Reality at the Israeli Institute for Advanced Studies, May 2018. The research towards this paper was supported by grants 1148/18 and 690/21 of the Israeli Science Foundation.

ORCID

Meir Hemmo  <https://orcid.org/0000-0002-0752-1269>

Orly Shenker  <https://orcid.org/0000-0001-5716-6894>

REFERENCES

- Arabatzis, T. (2018) Perspectival realism about what? In: Massimi, M. & McCoy, C.D. (Eds.) *Understanding Perspectivism*. London, England: Routledge, pp. 181–201.
- Arntzenius, F. (2000) Are there really instantaneous velocities? *The Monist*, 83(2), 187–208.
- Atkinson, D. (2006) Does quantum electrodynamics have an arrow of time? *Studies in History and Philosophy of Modern Physics*, 37(2006), 528–541.
- Bechtel, W. (2017) Explicating top-down causation using networks and dynamics. *Philosophy of Science*, 84, 253–274.
- Ben Menahem, Y. (2001) Direction and description. *Studies in History and Philosophy of Modern Physics*, 32, 621–635.
- Ben-Menahem, Y. (2018) *Causation in science*. Princeton, NJ: Princeton University Press.
- Bird, A. & Tobin, E. (2018) Natural kinds. In: Zalta, E.N. (Ed.) *The Stanford encyclopedia of philosophy (Spring 2018 Edition)*, Available at: <https://plato.stanford.edu/archives/spr2018/entries/natural-kinds/> [Accessed 3rd March 2022]
- Callender, C. (1999) Reducing thermodynamics to statistical mechanics: the case of entropy. *Journal of Philosophy*, XCVI, 348–373.
- Callender, C. (2007) The emergence and interpretation of probability in Bohmian mechanics. *Studies in History and Philosophy of Modern Physics*, 38, 351–370.
- Chalmers, D. (1996) *The conscious mind*. Oxford: Oxford University Press.
- Chang, H. (2012) *Is water H2O?*. Dordrecht: Springer.
- Chang, H. (2017) Prospects for an integrated history and philosophy of composition. In: Leitgeb, H., Niiniluoto, I., Seppälä, P. & Sober, E. (Eds.) *Logic, methodology and philosophy of science – Proceedings of the 15th International Congress*. London, England: College Publications, pp. 215–231.
- Davidson, D. (1970) Mental events. In: Davidson, D. (Ed.) *Essays on actions and events*. Berkeley, CA: University of California Press, pp. 207–227.
- Davidson, D. (1993) Thinking causes. In: Heil, J. & Mele, A.R. (Eds.) *Mental causation*. Oxford, England: Oxford University Press, pp. 3–18.
- Dürr, D., Goldstein, S. & Zanghi, N. (1992) Quantum equilibrium and the origin of absolute uncertainty. *Journal of Statistical Physics*, 67(5/6), 843–907.
- Ehrenfest, P. & Ehrenfest, T. (1912) *The conceptual foundations of the statistical approach in mechanics*. New York: Dover.
- Einstein, A. (1919) What is the theory of relativity? *London Times*, 28, 1919.
- Feynman, R.P., Leighton, R.P. & Sands, M. (1975) *The Feynman lectures on physics*. Reading, MA: Addison-Wesley Publishing Company for California Institute of Technology 1975–76.
- Fodor, J. (1974) Special sciences: or the disunity of science as a working hypothesis. *Synthese*, 28, 97–115.
- Fodor, J. (1997) Special sciences: still autonomous after all these years. *Nous*, 31, 149–163.
- Frigg, R. (2008) A field guide to recent work on the foundations of statistical mechanics. In: Rickles, D. (Ed.) *The Ashgate companion to contemporary philosophy of physics*. London: Ashgate, pp. 99–196.
- Frigg, R. & Hoefer, C. (2015) The best Humean system for statistical mechanics. *Erkenntnis*, 80, 551–574.
- Frisch, M. (2015) *Causal reasoning in physics*. Cambridge: Cambridge University Press.
- Goldstein, S. & Lebowitz, J. (2004) On the (Boltzmann) entropy of nonequilibrium systems. *Physica D*, 193, 53–66.
- Goldstein, S., Lebowitz, J., Tumulka, R. & Zanghi, N. (2020) Gibbs and Boltzmann entropy in classical and quantum mechanics. In: Allori, V. (Ed.) *Statistical mechanics and scientific explanation: determinism, indeterminism and Laws of nature*, arXiv:1903.11870.
- Hemmo, M. & Shenker, O. (2012) *The road to Maxwell's Demon*. Cambridge: Cambridge University Press.
- Hemmo, M. & Shenker, O. (2015a) The emergence of macroscopic regularity. *Mind and Society*, 14, 221–244.
- Hemmo, M. & Shenker, O. (2015b) Probability and typicality in deterministic physics. *Erkenntnis*, 80, 575–586.
- Hemmo, M. & Shenker, O. (2016) *Maxwell's Demon*. Oxford University Press Handbooks Online. <https://doi.org/10.1093/oxfordhb/9780199935314.013.63>
- Hemmo, M. & Shenker, O. (2019a) Two kinds of high-level probability. *The Monist*, 102(4), 458–477.
- Hemmo, M. & Shenker, O. (2019b) The second law of thermodynamics and the psychological arrow of time. *British Journal for the Philosophy of Science*. <https://doi.org/10.1093/bjps/axz038>

- Hemmo, M. & Shenker, O. (unpublished manuscript) A dilemma for Davidson's anomalous monism. In: Menahem, Y. B. (Ed.) *Rethinking laws of nature*. Springer.
- Hofer, C. & Marti, G. (2019) Water has a microstructural essence after all. *European Journal for Philosophy of Science*, 9, 12.
- Jaynes, E.T. (1965) Gibbs vs. Boltzmann Entropies. *American Journal of Physics*, 33, 391 Reprinted in E. Jaynes, *Papers on Probability, Statistics and Statistical Physics*, Dordrecht: Reidel, 1983, 79–88.
- Kim, J. (1998) *Mind in a physical world*. Cambridge, MA: MIT Press.
- Kim, J. (2012) The very idea of token physicalism. In: Simone, G. & Christopher, H. (Eds.) *New perspectives on type-identity*. Cambridge, England: Cambridge University Press, pp. 167–185.
- Kripke, A.S. (1980) *Naming and necessity*. Boston, MA: Harvard University Press.
- Ladyman, J., Ross, D., Spurrett, D. & Collier, J. (2007) *Every thing must go*. Oxford: Oxford University Press.
- Landau, L.D. & Lifshitz, E.M. (1980) In: Sykes, J.B. & Kearsley, M.J. (Eds.) *Statistical physics part 1, course in theoretical physics*, Vol. 5, 3rd edition. Oxford: Butterworth-Heinemann.
- Lebowitz, J. (1993) Boltzmann's entropy and time's arrow. *Physics Today*, 46(9), 32.
- Lewis, D. (1994) Reduction of mind. In: Guttenplan, S. (Ed.) *A companion to the philosophy of mind*. Oxford and Cambridge, MA: Blackwell, pp. 412–431.
- List, C. & Pivato, M. (2015) Emergent chance. *Philosophical Review*, 124(1), 119–152.
- Loewer, B. (2001) Determinism and chance. *Studies in History and Philosophy of Modern Physics*, 32, 609–620.
- Loewer, B. (2012) Two accounts of laws and time. *Philosophical Studies*, 160(1), 115–137.
- Maudlin, T. (2007) What could be objective about probabilities? *Studies in History and Philosophy of Modern Physics*, 38, 275–291.
- Noonan, H. & Curtis, B. (2017) Identity. In: Zalta, E.N. (Ed.) *The Stanford encyclopedia of philosophy*, Available at: <https://plato.stanford.edu/archives/spr2017/entries/identity/> [Accessed 3rd March 2022]
- Papineau, D. (1993) Physicalism, consciousness and the antipathetic fallacy. *Australasian Journal of Philosophy*, 71, 169–183.
- Polger, T. & Shapiro, L. (2016) *The multiple-realization book*. Oxford: Oxford University Press.
- Putnam, H. (1967) The nature of mental states. In: Putnam, H. (Ed.) *Mind, language and reality*. Cambridge: Cambridge University Press, pp. 429–440 (1975); Originally published as “Psychological Predicates,” in: William H. Capitan and Daniel D. Merrill (eds.), *Art, Mind and Religion*, pp. 37–48, Pittsburgh, PA: University of Pittsburgh Press (1967).
- Russell, B. (1925) Preface. In: Lange, F. (Ed.) *The history of materialism and criticism of its present importance*, Translated by E.C. Thomas. London: Routledge. V–XIX.
- Shenker, O. (2015) Davidsonian descriptions as a principle theory. *Iyyun: The Jerusalem Philosophical Quarterly*, 64, 171–190 In Hebrew. Available at: <https://www.jstor.org/stable/24307974> [Accessed 3rd March 2022]
- Shenker, O. (2017a) Foundations of statistical mechanics: mechanics by itself. *Philosophy Compass*. <https://doi.org/10.1111/phc3.12465>
- Shenker, O. (2017b) Foundations of statistical mechanics: the auxiliary hypotheses. *Philosophy Compass*. <https://doi.org/10.1111/phc3.12464>
- Shenker, O. (2017c) Flat physicalism: some consequences. *Iyyun: The Jerusalem Philosophical Quarterly*, 66, 211–225.
- Shenker, O. (2021) Foundations of quantum statistical mechanics. In: Knox, E. & Wilson, A. (Eds.) *Routledge companion to the philosophy of physics*. Oxford: Routledge.
- Sklar, L. (1993) *Physics and chance*. Cambridge: Cambridge University Press.
- Smart, J.J.C. (2017) The mind/brain identity theory. In: Zalta, E.N. (Ed.) *The Stanford encyclopedia of philosophy*, Available at: <https://plato.stanford.edu/archives/spr2017/entries/mind-identity/> [Accessed 3rd March 2022]
- Uffink, J. (2004) Boltzmann's work in statistical physics. In: Zalta, E.N. (Ed.) *The Stanford encyclopedia of philosophy*, Available at: <http://plato.stanford.edu/archives/win2008/entries/statphys-Boltzmann/> [Accessed 3rd March 2022]
- Uffink, J. (2007) Compendium to the foundations of classical statistical physics. In: Butterfield, J. & Earman, J. (Eds.) *Handbook for the philosophy of physics, Part B*, Amsterdam, Netherlands: Elsevier, pp. 923–1074.
- Wallace, David (2001) Implications of quantum theory in the foundations of statistical mechanics. unpublished. Available at: <http://philsci-archive.pitt.edu/410/> [Accessed 3rd March 2022].
- Wallace, D. (2015) The quantitative content of statistical mechanics. *Studies in History and Philosophy of Modern Physics*, 52, 285–293.
- Wernrdl, C. & Frigg, R. (2017) Mind the gap: Boltzmannian versus Gibbsian equilibrium. *Philosophy of Science*, 84(5), 1289–1302.

How to cite this article: Hemmo, M. & Shenker, O. (2022) Flat Physicalism. *Theoria*, 1–22. Available from: <https://doi.org/10.1111/theo.12396>