Why Functionalism Is a Form of ‘Token-Dualism’

Meir Hemmo†, Orly Shenker‡

Draft

Abstract

We present a novel reductive theory of type-identity physicalism (called Flat Physicalism), which is inspired by the foundations of statistical mechanics as a general theory of natural kinds. We show that all the claims mounted against type-identity physicalism in the literature don’t apply to Flat Physicalism, and moreover that this reductive theory solves many of the problems faced by the various non-reductive approaches including functionalism. In particular, we show that Flat Physicalism can account for the (alleged) appearance of multiple realizability in the special sciences, and that it gives a novel account of the genuine autonomy of the kinds and laws in the special sciences. We further show that the thesis of genuine multiple realization, which is compatible with all forms of non-reductive approaches including functionalism, implies what we call token-dualism; namely the idea that *in every token* (that partakes in this multiple realization) there are non-physical facts, which may either be non-physical properties or some non-physical substance. In other words, we prove that non-reductive kinds necessarily assume non-reductive tokens, i.e., *token dualism*. Finally, we show that all forms of non-reductive approaches including functionalism imply a literally multi-leveled structure of reality.

Word count: 15292 (total); 14,080 (excluding references and abstract)

Keywords: anomaly; autonomy of special sciences; functionalism; computational functionalism; multiple-realizability; non-reductive physicalism; reductive physicalism; type-identity; token; levels of reality.

1. The state of art concerning reductive type-identity physicalism

> 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.)

These words of Bertrand Russell, written almost a hundred years ago, still describe the state of the art in philosophy, if the term ‘materialism’ is interpreted as *reductive type-identity physicalism* (*“reductive physicalism”* for short). In contemporary philosophy there are two major lines of thinking that reject reductive type-identity physicalism. One rejects physicalism of all sorts, and endorses dualism; for example, arguments around the so-called “hard problem

† Philosophy Department, University of Haifa; meir@research.haifa.ac.il
‡ The Hebrew University of Jerusalem; orly.shenker@mail.huji.ac.il

of consciousness". We don't address these arguments here.[1],[2] The other line of thinking that rejects reductive type-identity physicalism consists of the variety of approaches all of which fall under the general title "non-reductive physicalism". Some central representatives of this line of thinking are our target in this paper. In contemporary literature, this second line of thinking is so dominant, that there is sometimes tendency to disregard the alternative altogether. Here are two very recent examples. Elpidrou (2018), in "Introduction: The Character of Physicalism" to a special issue of the journal *Topoi*, begins by saying that "Not many issues in philosophy can be said to match, let alone rival, physicalism's importance, persistent influence, and divisiveness" (p. 435), and continues to characterize physicalism in such a way that "various identity theories will not be forms of physicalism as understood here." (p. 437). Tiehen (2018), in an overview of contemporary literature in the journal *Analysis*, titled "Physicalism", begins by characterizing "Physicalism" as the thesis that "*that there is nothing over and above the physical*", and continues to characterize the notion of "nothing over and above" in such a way that reductive physicalism is conspicuously not an option taken seriously.[3]

Of course, there are arguments against reductive physicalism (for example, the argument that the mental is computational); but unless this view is presented, taken seriously, and (even!) defended, the game isn't over; and, as we show in this paper, it certainly isn't. Reductive type-identity physicalism should resume a position at the center of the philosophical stage, as an important detailed view that has a lot to be said in its favor, and which is – as we will show – much stronger than the varieties of non-reductive physicalism in solving some of the major problems faced by contemporary philosophy of mind and of science. Reductive type-identity physicalism is at one end of the spectrum of ontological theories (we will not try to characterize the other end of this spectrum); and understanding it is necessary in order to obtain a good understanding of the entire spectrum. Indeed, in this paper we shall point out some results of clarifying this point.

One reason why reductive physicalism was rejected by so many thinkers in the past, and is still rejected (or ignored) by the majority in the present, might be that the theories of reductive physicalism that have been described in the literature are, simply, not good. While people have made a lot of efforts to develop non-reductive theories, and there is a variety of them in the literature, that incorporate a variety of metaphysical assumptions and respond to a variety of

---

[1] We don't find arguments based on the so-called "hard problem" very convincing. For example, conceivability arguments, such as Kripke's (1980) or Chalmers's (1996), presuppose that the non-identity of mental kinds with physical kinds is conceivable. We reject this intuitive assumption. Since according to reductive type-identity physicalism mental kinds *are* physical kinds, the option of non-identity is a *contradiction*, and therefore it is *inconceivable* in any interesting way. But as we said, this is not our topic.

[2] In this paper we set aside Hempel's dilemma, because it applies to all forms of physicalism and this paper focuses on the distinction between reductive and non-reductive physicalism. For Hempel's dilemma, see e.g., (Ney 2008).

[3] In (Brown and Ladyman 2019), which provides a historical and philosophical overview of materialism and physicalism, the section on "Mind-Brain identity theory" (in Ch. 6) consists of (a) a non-critical summary of the introductory essay of a collection of essays from 1969, and (b) a non-critical presentation of a quotation from an essay by Crane in the *Time Literary Supplement* (2017) that expresses support of the "irreducible reality of the mental" (an essay that ends with the sentence "We will make no progress at all until we move beyond the simplistic brain's eye view.").

objections, the reductive theories that are presented aren't very well developed. Here are two examples that, while not very recent, are still influential.

One very influential description of a reductive type-identity physicalist theory is by Smart (1959),[4] who writes: "It seems to me that science is increasingly giving us a viewpoint whereby organisms are able to be seen as physico-chemical mechanisms. ... That everything should be explicable in terms of physics except the occurrence of sensations seems to me to be frankly unbelievable." (p. 142) Importantly, Smart says explicitly (on p. 143) that "the above is largely a confession of faith", not an argument. To see how Smart characterizes the physicalist view, consider his characterization of dualism: "There does seem to be, so far as science is concerned, nothing in the world but increasingly complex arrangements of physical constituents. All except for one place: in consciousness." (p. 142) Reductive physicalism simply omits the second sentence. This is a starting point for constructing a physicalist theory, but certainly calls for more details and for more arguments; without them it is quite vague.

Putnam (1975) adds some details to this very general characterization of reductive physicalism. In the paper in which he proposed the computational functionalist theory of mind[5], he criticized the "brain-state hypothesis" or the "physical-chemical hypothesis." (section III). What exactly is that hypothesis, according to Putnam? Concerning the computational functionalist hypothesis, that is proposed in that paper, Putnam writes: "This hypothesis is admittedly vague, though surely no vaguer than the brain-state hypothesis in its present form" (p. 434), and a bit later, "I contend, in passing, that this hypothesis, in spite of its admitted vagueness, is far less vague than the 'physical-chemical state' hypothesis is today, and far more susceptible to investigation of both a mathematical and an empirical kind." (1975, p. 435) Of course, both computational functionalism and the physical-chemical state theories were in their scientific infancy at the time, and to a large extent they still are. What, in particular, is the theory that he criticizes as too vague? The characteristics of this theory that Putnam provides are the following. (i) It is about the physical-chemical state of the brain. (ii) It is incompatible with psychophysical dualism (unlike functionalism: Putnam writes that "the functional-state hypothesis is not incompatible with dualism!" (1975, p. 436) (iii) Since the third characteristic has been immensely influential on arguments against reductive physicalism, let us bring it in Putnam's words (despite its length).

> Consider what the brain-state theorist has to do to make good his claims. He has to specify a physical-chemical state such that any organism (not just a mammal) is in pain if and only if (a) it possesses a brain of a suitable physical-chemical structure; and (b) its brain is in that physical-chemical state. This means that the physical-chemical state in question must be a possible state of a mammalian brain, a reptilian brain, a mollusc's brain (octopuses are mollusca, and certainly feel pain), etc. At the same time, it must not be a possible (physically possible) state of the brain of any physically possible creature that cannot feel pain. Even if such a state can be found, it must be nomologically certain that it will also be a state of the brain of any extra- terrestrial life that may be found that will be capable of feeling pain before we can even entertain the supposition that it may *be* pain. It is not altogether impossible that such a state will be found. Even though octopus and mammal are examples of parallel (rather than sequential) evolution, for example, virtually identical structures (physically speaking) have evolved in the eye of the octopus and in the eye of the mammal, notwithstanding the fact that this organ has evolved from different kinds of cells in the two cases. Thus it is at least possible that parallel evolution,

[4] It has been suggested that Smart should be read as a non-reductionist. Be that as it may, our interest here is in his characterization of the reductionist view; see also (Smart 2017).
[5] See (Shagrir 2005) for a retrospective exposition of the computational approach to the mind.

> all over the universe, might *always* lead to *one and the same* physical "correlate" of pain. But this is certainly an ambitious hypothesis. Finally, the hypothesis becomes still more ambitious when we realize that the brain state theorist is not just saying that *pain* is a brain state; he is, of course, concerned to maintain that *every* psychological state is a brain state. Thus if we can find even one psychological predicate which can clearly be applied to both a mammal and an octopus (say "hungry"), but whose physical-chemical “correlate” is different in the two cases, the brain-state theory has collapsed. It seems to me overwhelmingly probable that we can do this. Granted, in such a case the brain-state theorist can save himself by ad hoc assumptions (e.g., defining the disjunction of two states to be a single "physical-chemical state"), but this does not have to be taken seriously. (Putnam 1975 pp. 436)

Putnam’s account is certainly more detailed than Smart’s, in that he puts forward conditions that a reductive physicalist view must satisfy. However, it seems that Putnam has in mind a very vague idea of a “brain state” theory, and it may be due to the poverty of this theory that he takes it to be unreasonable in the way that he does. Indeed, we present in this paper a much more detailed theory of reductive type-identity physicalism, and the reader may return to the quotation from Putnam later on and see that what Putnam sees as an unreasonable ambition is, in fact, an extremely successful line of *actual* scientific research.[6]

2. The tasks of Flat Physicalism

Task I: Constructing the Flat Physicalism theory of reductive type-identity physicalism.

In view of this state of art, our first task is to propose a *new* theory of reductive type-identity physicalism, that we call “*Flat Physicalism*.” This theory is based on recent results in the philosophy of physics, specifically the philosophy of statistical mechanics, and it takes statistical mechanics as a paradigmatic example on the basis of which the notion of “physical kinds” is to be understood.

Our next tasks are to explain how this theory responds to some major criticisms mounted against reductive type-identity physicalism (e.g., the worries expressed by Putnam in the above quotation), and show that it solves problems faced by varieties of so-called non-reductive physicalism in ways that are much better than those of the latter.

For example, we will show how Flat Physicalism accounts for the appearance of multiple-realizability in the special sciences; explain why the (so-called “high-level”) regularities described by special sciences (for example probabilistic regularities, or even irregularities), seem to be independent of the regularities described by the (so-called “lower-level”) laws of physics, and consequently, in what sense the laws of the special sciences may be genuinely autonomous.

Task II: Provide a Flat Physicalist account of the (alleged) appearance of multiple realization.

As is well-known many forms of so-called non-reductive physicalist approaches allow for multiple-realization of the kinds that appear in the special sciences by physical kinds. Whether or not such multiple realization is observed in experience is an empirical question, that we

[6] It seems to us that this conclusion is also supported by Polger and Shapiro (2016), even though we don’t endorse their characterization of realization and of the special sciences kinds that comes with it.

don't address here (see e.g. Polger and Shapiro 2016, Bickle 2010). Our task is to prove that genuine multiple realizability entails psychophysical dualism (either property or substance); and so *if* genuine multiple realizability obtains in the world, *then* psychophysical dualism (either property or substance) obtains in the world.

The idea of multiple-realization has been introduced (in different terms) as part of a non-reductive approach by Putnam (1967) in his proposal for a computational-functionalist theory of the mind, and developed by (e.g.) Fodor (1974), who took it to be one of the salient motivations for developing non-reductive approaches to the special science. However, as we will show, genuine multiple-realization of special science kinds by physical kinds entails what we call *token-dualism*, that is: it entails that *in every token* (that partakes in this multiple realization) there are non-physical facts, which may either be non-physical properties or some non-physical substance. It is sometimes said that in (so-called) non-reductive physicalism each token is physical and yet the special sciences kinds aren't reducible to physics; this idea is sometimes called "*token physicalism*". We shall prove that non-reductive kinds necessarily assume non-reductive tokens, i.e., *token dualism*. We show that this is the case even if the special sciences kinds supervene on physical kinds: in this sense, *supervenience isn't sufficient* to ensure that the world is physical. The only theory in which the tokens are physical is that of a reductive type-identity physicalism that doesn't allow for multiple-realization.[7]

Task III: Show that if Flat Physicalism obtains, then there are no levels of reality.

The third task of this paper is to explain in what sense the ontology in our proposed type-identity theory is flat, that is, that there *are no* levels of reality; hence the name "Flat Physicalism". Whether or not one chooses to use the term "levels" to describe certain features of reality, so as to obtain so-called "levels of explanation" or "levels of descriptions" etc., is immaterial to this ontological claim. As a corollary we will show that any approach which is compatible with multiple-realizability, regardless of whether it also assumes supervenience of high-level kinds on low-level kinds, entails multiple-levels of reality (not only of explanations, descriptions, etc.).[8]

Task IV: Show that Flat Physicalism allows for an autonomy of the special sciences and is compatible with all forms of special sciences laws, including probabilistic ones and even with cases of special sciences anomaly.

Fodor (1974) famously emphasized that one argument in favor of a non-reductive picture is the fact that the special sciences progress independently of physics, and the forms of their laws appear to be very distinct from that of the laws of physics; they are autonomous, in this sense. Our task is to show that these observations are completely compatible with Flat Physicalism, which can account for them. We will explain in what sense the special sciences are autonomous within reductive type-identity physicalism, and will explain how special sciences laws, that have forms very different from those of physics, can come about. One well known example is the time-directed second law of thermodynamics that obtains phenomenologically despite the temporal symmetry of fundamental physics; we shall not address this case in detail here (see

[7] Our argument (in Section 5) that multiple-realizability implies token-dualism is independent of other arguments (e.g., Churchland 1982; Churchland 1986; Kim 1992) against non-reductive approaches.

[8] See Bechtel (2017) for other arguments as to why the mechanistic multi-leveled ontology should be flattened. But he retains levels of *explanation*.

e.g., Sklar 1993; Albert 2000; Uffink 2007; Frigg 2008; Hemmo and Shenker 2012, 2016). We will show, in general outline, how special sciences laws that are probabilistic, and even cases of anomalous phenomena (e.g., the anomaly of the mental conjectured by Davidson 1970) can come about within a Flat Physicalist picture (see also Hemmo and Shenker 2015; Shenker 2017a, Hemmo and Shenker 2019a, 2019b, 2019c).

The paper is structured as follows. In Section 3 we present our proposal called Flat Physicalism which is a full-fledge reductive type-identity theory *inspired* by recent results in the philosophy of physics, mainly the philosophy of statistical mechanics. In Section 4 we consider in detail the strict identity of physical kinds and the special sciences kinds, and show that our physical theory is reach enough to explain the appearance of multiple-realizability in our experience in terms of physical kinds that are *identical* to the special science kinds. In Section 5 we show that any view which is compatible with genuine multiple-realizability (such as all forms of functionalism including computational functionalism) entails that there is something *non-physical*, which is *present* in each and every (*token*-)occurrence of the multiply-realizable kind. In Section 6 we show explicitly on the basis of Flat Physicalism how the special sciences could be autonomous and even anomalous in the sense that the (physical) kinds they describe don't strictly satisfy any law. In Section 7 we show that non-reductive approaches are committed to a picture of reality which is *literally* of multiple levels. Section 8 is the conclusion.

3. Constructing Flat Physicalism: A novel theory of reductive type-identity physicalism.

As we said above, our first task in this paper is to present a reductive physicalist theory, in which the special sciences kinds are physical kinds by *strict iden*tity. The details of our approach are inspired by recent results in the philosophy of physics, mainly the philosophy of statistical mechanics (see Hemmo and Shenker 2012, 2016; Shenker 2017b, 2017c), but in order to follow the discussion here, one need *not* be acquainted with the details of that theory.[9] We take statistical mechanics to be *a general theory of physical kinds*, and accordingly, will present its principles in the most general way. There are, in the literature, ample discussions of *special science kinds*, and of the ways in which they can relate to the physical kinds; there is much less discussion of what a *physical kind* is, and since such an analysis is crucial for understanding reductive identity physicalism, we undertake it in this section.

The main idea of Flat Physicalism is just this: the world is as described by physics, and this is *everything* that there is. Nothing is left out. The starting point is the *tokens*: by assumption, every token state of affairs is *fully* described by physics.[10] Understanding the tokens is key to our ideas; types will come out of them as we show below.

[9] For the standard approaches to the foundations of classical statistical mechanics, see e.g., (Sklar 1993); (Albert 2000); (Uffink 2007); (Frigg 2008). For our approach, see (Hemmo and Shenker 2012, 2016); (Shenker 2017b, 2017c); for the quantum case, see Hemmo and Shenker 2012, Appendix, Hemmo and Shenker 2017b; Shenker 2018).

[10] Scientific realism doesn't imply this characterization of a token. "Realism is about what is real and not about what is fundamentally real." (Psillos 2009, p. 38).

Let us illustrate our ideas with an example. Our toy model will be a universe that can be coherently, fully and correctly described by classical mechanics.[11] In classical mechanics one can distinguish between two kinds of tokens: *token-states* and *token-sequences*. In classical mechanics a token state – called a *microstate* – consists of the exact positions and velocities of all the particles in the universe at a point of time;[12] and the corresponding example of a token-sequence is a sequence of such microstates, generated by the equations of motion, and often called a *micro-trajectory*. According to classical mechanics, given a microstate of the universe, together with parameters such as the mass of each particle, constraints such as the volume available to them, and limitations such as the total energy of the universe, the equations of motion (ideally) yield a continuous infinite temporal sequence of microstates. The connection and distinction between token-states and token-sequences, illustrated by the distinction between microstates and micro-trajectories, is significant for our analysis of functionalism (below).

Suppose that we are given the full details of some physical token-state or token-sequence; and suppose, as assumed by Flat Physicalism, that physics is complete, so that the physical microstate or trajectory describes everything that there is. And then suppose that we want to talk about some special science, for example: we want to talk about some laws of biology or of geology. Many think that the terms, the properties and laws of physics don't capture those of these special sciences, and therefore in order to describe these features of the universe we need to *add* something to the physical microstate or trajectory, that is, add something which isn't part of what we took to be the complete (physical) description of what there is. The complete physical description of the world, so they say, *misses out* something. This claim is *rejected* by Flat Physicalism, that offers the following alternative.

According to Flat Physicalism, the microstate of the universe is everything that there is, so that one cannot say *more* about the special science kinds beyond specifying this microstate. The only thing that remains, if one wants to say something *different* from physics, is to say *less* than physics does. And this is the route taken by Flat Physicalism: when we talk about special science kinds we say l*ess not more*: we refer to an *aspect* of the token-state or token-sequence (that is, respectively, an aspect of a mechanical microstate of the world or of its micro-trajectory), and this aspect is given to us by a *partial description* of the token. So, according to Flat Physicalism the special sciences are about certain aspects of the physical tokens of the world, and *they cannot be anything else because there isn't anything else*.

A well-known example of an aspect in our sense is the identity statement "heat *is* molecular motion", or: "the temperature of an ideal gas in equilibrium *is* the average kinetic energy of its particles". In classical mechanics (which is our example), average kinetic energy of the particles of a sample of an ideal gas is only an aspect, given by a partial description, of a microstate of the universe, or more specifically, of the gas in question, which is a subsystem of the universe. There are other aspects (and other details about) the microstate of the universe

[11] We agree with Ladyman and Ross (2007) and Wallace (2001) that it is a mistake to carry out metaphysical investigations, assuming that classical mechanics is unrestrictedly true. At the same time, the use of classical terminology and laws is legitimate if they preserve essential features of the phenomena and fundamental facts being addressed, which is the case here, as it brings out the main ideas of Flat Physicalism.

[12] Of course, relativistic considerations should enter at this point, in non-classical physics. In standard quantum mechanics the microstate is the *pure* quantum state in Hilbert space; see (Shenker 2018) for our view concerning the foundations of quantum statistical mechanics.

at the moment of interest, and even other details of that subsystem, that aren't given by this partial description; for example, the positions of the gas particles, their total number, their specific velocities. Any such aspect is part of, or is literally in, the microstate: it is there in the same sense that the entire microstate is there (with the aspects on which we don't focus), even if we aren't looking at it, as it were. Once one (for example, Laplace's proverbial Demon) has access to the details of the microstate (or the micro-trajectory) of the universe, one can *derive* from it any aspect, by *ignoring* some details of the microstate or micro-trajectory (see also Portides 2019 on this idea). In that sense, there is nothing above and beyond the microstate (or micro-trajectory) described by physics; this is all that there is.

Flat Physicalism is an *identity theory* that is *not eliminativist*,[13] in the following sense. According to Flat Physicalism, the facts that are described by the special sciences are out there, as it were (subject to all the arguments for and against scientific realism). Science tells us that when referring to them we actually refer to the appropriate aspect of the fundamental physical microstate. Water *just is* H_2O (see Chang 2012 and Hoefer and Marti 2019); it isn't correlated with H_2O. Similarly, to use our example, temperature (of an ideal gas in equilibrium) is (identical to) average kinetic energy. It would be a *mistake* to say in statistical mechanics that, for example, average kinetic energy "gives rise to" temperature or "grounds" temperature, since according to statistical mechanics, there are no facts (any sort of facts) in the world beyond the actual microstate of the universe, its aspects and its sequence over time. In statistical mechanics there are no "mereological facts" which are about "combining Lego parts" as it were (see Chang 2012; Hemmo and Shenker 2015; 2016, 2017b; Shenker 2017a), since a *conceptually and physically inseparable* part of the physical description is the intermolecular interactions. In this sense, statistical mechanics (when successful) accounts for the thermodynamic phenomena in term of strict *identity statements* of the kind 'temperature is molecular motion.'

As Smart wrote already in (1956), "you cannot correlate something with itself. You correlate footprints with burglars, but not Bill Sykes the burglar with Bill Sykes the burglar." (The same point is made by Papineau 1993.) The discovery that water is H2O doesn't eliminate the notion of water, but enriches it.

A question arises: Every token (state or sequence) has infinitely many aspects, each being a function of the complete microstate of the universe. All of them "exist" with the full token. Nevertheless, only a relatively small number of aspects appear in our experience and in our theories. Why is that so, and how the "preferred" (as it were) aspects are selected? Here is the Flat Physicalist explanation for this.

Consider two interacting systems (each is a sub-system of the universe), where the interaction between them is such that certain aspects of the microstate of each of them become correlated with certain aspects of the microstate of the other.[14] Although both are completely and fully on a par as being physical systems, and although physics is blind to any roles we might ascribe to any of the systems, it is convenient, for the purpose of our illustration, to call one of them a

---

[13] For an overview of eliminativism, see (Ramsy 2019).

[14] More precisely, the microstate of the universe evolves along its micro-trajectory such that if we focus on the two sets of the degrees of freedom associated with these two systems, and of certain aspects of the microstates of them, then we find that those aspects are correlated. We make here the assumption that the microstate of the universe is separable, which is standard in classical physics though not in quantum mechanics.

'measuring device', and the other a 'measured system', and we shall say that the measuring device is *sensitive* to certain aspects of the measured system, and not to others, in the case that from the end state of certain aspects of the measuring device, one can tell which was the value of the corresponding aspects of the measured system to which the measuring device is sensitive, at a certain time of interest, say the time of interaction. Our sense organs are such measuring devices, and (presumably) so are our brains: and they are sensitive to certain aspects of our environment (and not to others). Those aspects of our environment to which we are sensitive appear in the special sciences. If we are lucky (and we sometimes are) the aspects to which we are sensitive satisfy certain regularities, a fact that we can use to make predictions (and evolutionary explanations can be given for this fact). It is a task of the special sciences to identify the aspects that satisfy useful regularities and to formulate these regularities. Another task of the special sciences is this. We can extend the set of the aspects of our environment to which we are sensitive by building new measuring instruments, following scientific discoveries concerning regularities governing certain additional aspects of our environment, to which our sense organs were not initially sensitive. The instruments are sensitive to these additional aspects of the environment, to which we have no direct physical access; and to read these instruments we employ aspects of them to which we are sensitive. Identifying such further aspects of the environment that satisfy regularities, hitherto unobserved by our sense organs, and which may be interesting and useful, is also the job of the special sciences. In this process of discovery and selection the special sciences are *autonomous* (see Hemmo and Shenker 2015, 2017b, 2019b; Shenker 2017a); more on the autonomous status of the special science kinds and laws in Section 7). It turns out, then, that the selection of "preferred" aspects of the environment is relative to an observer, possibly aided by measuring instruments; and the special sciences identify them and their regularities. This role of the special sciences, which is of extreme importance in understanding our world, explains their autonomy; we return to this point in Section 7.

Consider, now, a case in which a measurement interaction is carried out between a measuring device and a certain system in the environment, by the end of which a certain aspect of the state of the device is correlated with a certain aspect of the measured system. For convenience of presentation suppose that the measuring device is *us*, and our brain state by the end of the interaction registers the aspect of the environment to which we are sensitive. At this point, *all* we know (following this interaction) about the microstate of our environment is that it has this aspect; we know nothing about the other (infinitely many) aspects of the environment. The important point is that there can be many *counterfactual* microstates that *share* this (known) aspect, but differ in the other (unknown) aspects; and they would all look to us the same as the actual microstate with which we have interacted. A *set* of microstates that *share* an aspect is called a *macrostate*. Corresponding to the two sorts of tokens: token-states and token-sequences, there are two sorts of macrostates: macrostates that are *sets of microstates* that share the same physical aspect, and macro-sequences that are *sets of trajectories,* or of sequences, that share suitable physical aspects.[15] (The latter will be important in understanding what a "function" is and what "functionalism" is in Flat Physicalism.)

The fact that two microstates of the universe (or of some sub-system of it) share the same aspect, and therefore belong to the same macrostate, is a fact in the world that obtains regardless of whether or not there is a measuring device that is sensitive to it or interacts with it. At the

[15] There are various notions of kinds and properties that come up in such discussions. Our arguments here don't depend on the details of these analyses, as long as the latter are compatible with physics.

same time, the fact that we – as measuring devices – are sensitive to certain aspects of our environment and not to others, explains the apparent preferred status that certain aspects of the environment have, in that they feature in our experience. To repeat and to emphasize, it is a conceptually distinct matter whether these aspects of our environment satisfy some regularity. However, it makes sense to assume that there would be an evolutionary advantage to creatures that are sensitive to aspects of the environment that also satisfy regularities, thus enabling predictions.

With respect to macrostates it is important to notice the following fact. Consider the example "temperature is molecular motion". The "molecular motion" in question is an aspect of the microstate of a sample of gas. When we carry out a measurement interaction with a particular sample of gas, in order to measure the temperature of that gas at that event, we in fact interact with the suitable aspect (namely, the "molecular motion") of the *particular* microstate that obtains at that *particular* moment. We emphasize that we *never* measure a *set* of microstates, since the set consists of the *actual* microstate as well as (infinitely) many *counterfactual* ones; and in the measurement we interact with *actual* matters of fact, not counterfactual ones, that don't obtain (and most of which will *never* obtain) in the universe. Indeed, when we look at the *particular* event of the *particular* sample of gas, we *directly* know (by measurement) what its temperature is: We don't need to know which *other* (counterfactual) microstates would give rise to the same temperature (for all we know, or sense, or measure, this could be the only member of the set). Sometimes people say that what we actually measure are "macrostates". If by "macrostate" they refer to an *aspect* of the *particular actual* microstate (also sometimes called "macrovariable"), then that is correct; but if the claim is that we measure *sets* of microstates, then this statement is *wrong*.[16]

In Flat Physicalism, then, a physical token belongs to a physical kind if that token has the suitable aspect, ideally given by a partial description of it; and all the special sciences kinds are physical kinds in this sense, just because the microstate and the micro-trajectory of the universe are everything that there is, *tout court*. But this point involves some further details, to which we now turn.

4. Special sciences kinds in Flat Physicalism

Here is the *central* question concerning reductive versus non-reductive physicalism. Consider three tokens A, B and C, such that A and B belong to the same special sciences kind M but C doesn't. *What is the fact in virtue of which this is the case?* Which facts in the world fix the partition of tokens into types? From now on we focus on this question, answering it in the framework of Flat Physicalism, comparing this answer to those of other frameworks, and studying its implications.

---

[16]Above we gave an example of an aspect, namely, the average kinetic energy of the molecules of an ideal gas in equilibrium. Some people say that "averages" are above and beyond the physical facts, since, possibly, most of the sets of molecules don't have this particular velocity, and possibly, none has. But when we have the full details of the microstate, we already have this average; it can be logically derived from the microstate, and in this sense, it is already *in* the microstate. We don't need to postulate any additional facts over and above the microstate, to derive this fact. It exists, our there, as it were, even if nobody is interested in calculating it and will never actually derive it.

According to Flat Physicalism the fact in the world that fixes the partition of tokens to types must be in the microstate (or micro-trajectory) *of the universe*, just because this is everything that there is. There are two options here:
(I) The fact that makes A and B (but not C) members of the same kind is in A and B (but not C).
(II) This fact is elsewhere in the microstate of the universe.

Let us now examine these options.

(I) In the first option, tokens A and B share an aspect which is M, while C doesn't have this aspect. We, as measuring devices, are sensitive to this shared aspect and register it.

An important result of this way of understanding special sciences kinds as physical kinds is that it provides an explanation for the following *fact*. In every particular event in which we encounter a particular token in our environment, say: A, we are able to say, directly and immediately, whether or not A is M, without having to know anything about B, and regardless of whether or not B ever obtains in the world. We detect the aspect M in the actual token A that obtains by physically interacting with it.

As we already said, these facts concerning the aspects M of A and B and not-M of C obtains whether or not these microstates are being observed by a measuring device that is sensitive to that aspect, or not; although we are naturally interested in those aspects to which we, as measuring devices, are sensitive. The aspect that A and B (but not C) share may be very complex, and maybe we will never find it by scientific investigation; but by assumption, it is there, and it accounts for the fact that A and B are of the same kind, but C is not.

Here again our example is instructive. Today we are so used to saying that temperature (of an ideal gas in equilibrium) is average kinetic energy, that we tend to forget that this was a highly non-trivial scientific discovery: nothing in the pre-scientific notion of "temperature" prepared us for the scientific identification of temperature with this particular complex mechanical aspect. Possibly, the aspects that are identical with special science kinds in biology, for example, are even more complex and hard to discover. But since according to Flat Physicalism there is *nothing in the world* except the microstate, and regardless of the complexity and the prospect of discovery, it follows that this is nevertheless the case. Flat Physicalism is the idea that the reduction of thermodynamics to mechanics, by way of such commitments, is to be generalized to all the special sciences.[17]

(II) Suppose, however, that it is empirically discovered that the tokens that belong to a certain special sciences kind are *heterogeneous*, that is: suppose that they don't share any *relevant* aspect, that can be associated with the special science kind (they may share (infinitely) many other aspects that aren't M). Many people, starting for example from Putnam (1975) and Fodor (1974) (albeit with important exceptions such as Polger and Shapiro 2016) believe that this is indeed the case in our world. Whether or not this is the case, and whether or not the arguments for this are empirical, is debatable; but here we want to examine what Flat Physicalism has to say in case this is indeed true, and therefore we shall *assume* – at least to begin with – that there

[17] Whether or not "temperature" is of case (I) or of case (II) is under debate. Compere (Frigg and Hoefer 2015) and (List and Pivato 2015) with (Hemmo and Shenker 2019a).

are sets of tokens that belong to the same special science kind but don't share any relevant physical aspect.

For supporters of multiple realization, such a case seems mysterious. Fodor famously wrote:

> I am suggesting, roughly, that there are special sciences not because of the nature of our epistemic relation to the world, but because of the way the world is put together: not all natural kinds (not all the classes of things and events about which there are important, counterfactual supporting generalizations to make) are, or correspond to, physical natural kinds." (1974, p. 113).
>
> "The very *existence* of the special sciences testifies to reliable macrolevel regularities that are realized by mechanisms whose physical substance is quite typically heterogeneous. Does anybody really doubt that mountains are made of all sorts of stuff? Does anybody really think that, since they are, generalizations about mountains-as-such won't continue to serve geology in good stead? Damn near everything we know about the world suggests that unimaginably complicated to-ings and fro-ings of bits and pieces at the extreme *microlevel* manage somehow to converge on stable *macro-level* properties." (Fodor 1997, p.160)

What is the ontology in question? Which facts bring about special sciences kinds? Fodor famously continues in his "Conclusion (molto mysterioso)" of this paper, as follows.

> Why there should be (how there could be) macrolevel regularities at all in a world where, by common consent, macrolevel stabilities have to supervene on a buzzing, blooming confusion of microlevel interactions... why there should be (how there could be) unless, at a minimum, macrolevel kinds are homogeneous in respect of their microlevel constitution. Which, however, functionalists in psychology, biology, geology and elsewhere, keep claiming that they typically aren't. So, then, why is there anything except physics? ... Well, I admit that I don't know why. I don't even know how to think about why. I expect to figure out why there is anything except physics the day before I figure out why there is anything at all, another (and, presumably, related) metaphysical conundrum that I find perplexing. (Fodor 1997, p. 161)

For supporters of Flat Physicalism, phenomena such as the ones mentioned by Fodor aren't mysterious at all, and can be fully explained in the framework of reductive physicalism, in the following way.

Consider three *mutually equally heterogeneous* tokens: A, B, and C. And suppose that A and B belong to the same special sciences kind M, but C doesn't. Which fact makes it the case that this is the partition to the special science kinds? Why is B in the same set as A, but C – which is, by assumption, heterogeneous to the same degree! – isn't? The phrase "*to the same degree*" is very important here: B isn't "more similar" to A than C is, since if it were the case, A and B would share a physical coarse-grained aspect of their microstates, so that this second sort of special science kinds would collapse into the first sort. Therefore, we must preclude this case and suppose that C is as heterogeneous relative to A as B is; all three tokens are completely different from each other; they are all *mutually heterogeneous to the same degree*. This case is called, in contemporary literature, *multiple-realizability of special science kinds by physical kinds*. By assumption, the two following facts obtain: (a) tokens A and B (but not C) are of the same special sciences kind; (b) tokens A, B, and C are equally mutually heterogeneous. How can (a) and (b) be reconciled?

An influential example of this idea was given by Fodor (1974) in order to illustrate his non-reductive approach. He writes:

> Gresham's law says something about what will happen in monetary exchanges under certain conditions. I am willing to believe that physics is general in the sense that it implies that any event which consists of a monetary exchange (hence any event which falls under Gresham's law) has a true description in the vocabulary of physics and in virtue of which it falls under the laws of physics. But **banal considerations** suggest that a description which covers all such events must be **wildly disjunctive**. Some monetary exchanges involve strings of wampum. Some involve dollar bills. And some involve signing one's name to a check. What are the chances that a disjunction of physical predicates which covers all these events (i.e., a disjunctive predicate which can form the right hand side of a bridge law of the form 'x is a monetary exchanged...') expresses a physical natural kind? In particular, what are the chances that such a predicate forms the antecedent or consequent of some proper law of physics? The point is that monetary exchanges have interesting things in common; Gresham's law, if true, says what one of these interesting things is. But what is interesting about monetary exchanges is surely not their commonalities under physical description. A natural kind like a monetary exchange could turn out to be co-extensive with a physical natural kind; but if it did, that would be an accident on a cosmic scale. In fact, the situation for reductivism is still worse than the discussion thus far suggests. For, reductivism claims not only that all natural kinds are co-extensive with physical natural kinds, but that the co-extensions are nomologically necessary: bridge laws are laws. So, if Gresham's law is true, it follows that there is a (bridge) law of nature such that 'x is a monetary exchange iff x is P', where P is a term for a physical natural kind. But, **surely**, there is no such law. If there were, then P would have to cover not only all the systems of monetary exchange that there are, but also all the systems of monetary exchange that there could be; a law must succeed with the counterfactuals. What physical predicate is a candidate for 'P' in 'x is a nomologically possible monetary exchange iff Px'?
> To summarize: an immortal econophysicist might, when the whole show is over, find a predicate in physics that was, in brute fact, co-extensive with 'is a monetary exchange'. If physics is general - if the ontological biases of reductivism are true - then there must be such a predicate. But (a) to paraphrase a remark Donald Davidson made in a slightly different context, nothing but brute enumeration could convince us of this brute co-extensivity, and (b) there would seem to be no chance at all that the physical predicate employed in stating the co-extensivity is a natural kind term, and (c) there is still less chance that the co-extension would be lawful (i.e., that it would hold not only for the nomologically possible world that turned out to be real, but for any nomologically possible world at all). (Fodor 1974, pp. 103-104; our empheses)

Let us put to the side, for the moment, the fact that Fodor's argument in this paper is based on "*surely*" statements concerning unspecified "*banal considerations*". Since the entire argument is based on unexamined intuitions, we find it rather shocking that this paper (and its 1997 sequel) became so influential. But let us focus on the example itself which is this: things that are physically heterogeneous, for example, strings of wampum or dollar bills (today we would add states of certain electrical circuits) belong to the same economics kind of "monetary exchange". How can that be?

Flat Physicalism offers the following account of this case. To combine our abstract notations with Fodor's example, suppose that token A is a string of wampum, token B is a dollar bill, and token C is a chair (assuming that chairs aren't cases of monetary exchange; in Fodor's approach they *could* be (as a brute fact), but suppose that (as a brute fact) they *aren't*); and M is the kind "monetary exchange". Consider Figure 1 in which the universe is partitioned into two sets of degrees of freedom: Along the vertical axis we depict the three systems with respect to which we ask whether their microstates are tokens of the kind M; and along the horizontal axis we depict the microstates of a system that we call The Device that starts out in some "ready state", and then according to the laws of physics evolves to a final pointer state that indicates “this is monetary exchange”, whereas if C obtains its final pointer state is “this wampums and dollar bills (but not chairs) are “monetary exchange”, despite the fact that it doesn’t measure (or otherwise reflect) any physical aspect of these tokens (which are equally mutually heterogeneous). The fact that The Device ends up in the same pointer state in cases A and B (but not C) isn’t in virtue of anything physical about A and B, but it is, of course, in virtue of the physics of the total universe, that takes the total microstate of the universe from the initial microstate at t0 to the final microstate at t_1. The fact that this final microstate of the universe happens to be such that The Device points at M in cases A and B but not C seems “*brute*” in the sense that it cannot be explained by any aspect shared by A and B, which are completely heterogeneous; but this perspective is misleading: the full physical explanation of matters of facts in the universe is given by the complete micro-trajectory of the universe, and is revealed by proper partition of the degrees of freedom of the universe into sub-systems. In our case, if we focus on the extended microstates of wampum+Device, dollar bill+Device, and chair+Device, then the special science kind “monetary exchange” turns out to be a feature shared by the first two extended microstates, that have the physical aspect M (but not with the third extended microstate, that has the aspect not-M). This fact is *fully explained* by fundamental physics, and *isn’t brute* at all. In other words: the kind “monetary exchange” is *not* a feature of wampums and dollar bills that is “*measured*” (in some appropriate sense of the term) by The Device: The Device is *not a “measuring device”* for measuring the property “monetary exchange”, and Figure 1 does*n’t* depict a measurement interaction; rather, “monetary exchange” is a feature of The Device itself, together with these other systems.

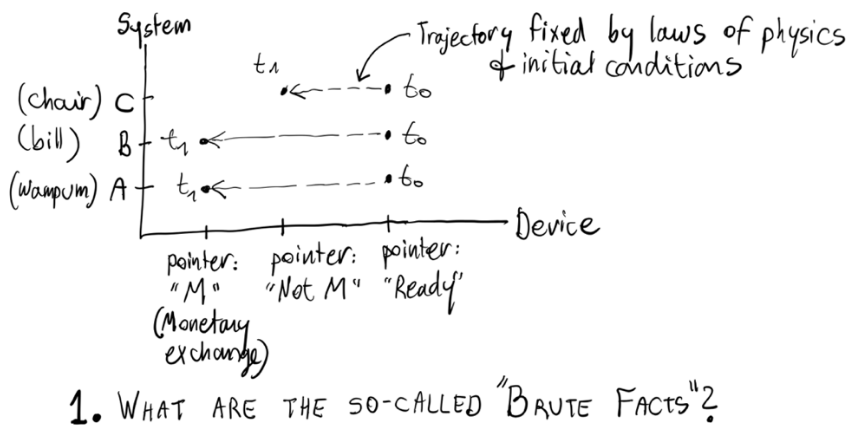

Figure 1: What are the so-called “brute facts”?

Constructing such a device isn’t trivial, and the evolution described by the trajectories depicted in Figure1 may seem strange or even conspiratorial, since one cannot explain the motion of The Device by appealing to a measurement interaction that is sensitive to a shared aspect of A and B. But however strange this case is perfectly compatible with fundamental physics.

If indeed wampums and dollar bills don’t share any physical aspect (as we assume here), then it turns out that nature herself has built an incredible device that operates as in Figure 1, namely: us. Our brains evolve, following our interactions with society, in such a way that certain brain structures are what we call “conventions” and so on, in such a way that when we encounter a coin we *immediately and directly* enter the mental state of “entertaining the thought that this is monetary exchange”. Regardless of what “conventions” (etc.,) are, in order for us to be able to recognize coins (or wampums, in the appropriate case) as monetary exchanges and act accordingly, in our daily lives, without constantly consulting with endless lists of tokens that might fall under this kind, our brains must be prepared in the physical (e.g., neuronal) state that assimilates these "conventions", and acts like The Device in Figure 1.

Since the special science kind "monetary exchange" is a feature not of wampums or dollar bills, but of wampum+Device or dollar bill+Device, it is *not genuinely multiply realized* by physical kinds. It nevertheless gives *the impression* of multiple-realization, if we ignore The Device. *Ignoring* The Device makes our scenario a case of *apparent, non-genuine*, multiple-realization: If we think of ourselves as The Device, then it is natural to expect that we should observe empirically the appearance of multiple-realization (as indeed many think that this is the case), since it is natural to treat *only* the external environment as "the system", ignoring the role of ourselves as The Device in the "extended system".[18]

In option (I) for constructing physical types, we stressed that the fact that two tokens share an aspect is a feature of the world regardless of whether or not any measuring device is sensitive to this aspect and can register it. This is different in option (II) for constructing physical types: here, The Device is essential, since the physical kind is a feature of it. Lacking a Device, there is no partition of the state space into the kinds M and not-M; there is no "monetary exchange", but only heterogeneous tokens that have nothing to do with each other.

Suppose, however, that one insists on the following ontological claim: A, B and C are mutually equally heterogeneous, and A and B are M (but C isn't) even if there is no Device. Flat Physicalism has no resources to account for this case, and we are led to the conclusion that there is something non-physical in each and every token microstate, a fact that makes it belong to either one of these two sets, *a non-physical fact that we somehow perceive* in the presence of that token. We shall come back in more detail to this point in the next section.

5. Why "non-reductive physicalism" is a form of token-dualism

The case in which the physically heterogeneous tokens A and B are of the same special science kind M, where M may be, for example, a biological kind or an economics kind, can – as we have just seen – be accounted for in Flat Physicalism, by extending the token and bringing in The Device, whose aspects are shared by the A+Device token and the B+Device token. This route is, however, not available for psychological or mental kinds. Our main argument in this section is this. If The Device is us, and its pointer states "M" or "not M" are our psychological or mental states in which we entertain the thought "this is monetary exchange" or "this isn't monetary exchange" (or have some other corresponding mental state), then – according to Flat Physicalism – the set of (for example) "having M thoughts" cannot be multiply realized by heterogeneous physical kinds; all the "having M thoughts" cases must share a physical aspect, and the kind "has M thoughts" would be identical with the physical kind of having that physical aspect. This has the following consequence: The popular idea called "non-reductive physicalism" in which the tokens are physical but the kinds are (somehow) not reducible to physics, is *inconsistent*. *Either* psychological kinds *are* (reducible to, identical to, nothing over and above) physical kinds and tokens are physical, *or* (exclusive or) psychological kinds are *not* (reducible to, identical to, nothing over and above) physical kinds, and *token-dualism* obtains in the world. By "*token dualism"* we mean that every token of the multiply-realizable kind contains a non-physical element, that may be a non-physical property of the token or some

---

[18] We thank Carl Hoefer for discussing this point with us extensively.

non-physical substance. This entails that *all* forms of functionalism (including computational functionalism), even if they require supervenience, as long as they allow for multiple realizability (even if they accept that in our world there is no multiple realization), are forms of token dualism. *Supervenience is, then, not sufficient for physicalism, and is compatible with property or substance dualism*. Let us see why this is the case.

Suppose that the tokens A and B are microstates of two systems that share the same mental state. They could for example be the microstates of a human being and an octopus both of which feel the exact same kind of pain (to use Putnam's 1967 famous example; see discussion in Polger and Shapiro 2016). And suppose that when The Device interacts with either A or B, it ends up in a microstate that has the physical aspect P, that indicates "being in pain". So far, the case is the same as the one of non-mental kinds (compare Figure 1), as described in the previous section.

But here are the difficulties.
(i) If this case is indeed similar to the above one, in which the kinds are non-mental, then – as we have seen – the property P (here: being in pain) isn't a property of A or B (here: the human being or the octopus) but of A+Device or B+Device, or even a property of The Device. So if we say that you, the reader, are in pain, we don't speak about you, and there is no fact of the matter concerning this mental state of yours, until some external Device is brought in to measure your brain state. This result seems to us unacceptable and contrary to the empirical starting point.

(ii) The alternative is to leave out The Device, and focus on A and B themselves. But A and B are – by assumption – completely genuinely heterogeneous: they share nothing physical. And if, in every token, the physical is everything that there is, then A and B share no relevant feature at all, and don't form a kind. In Figure 1, recall, if we omit the horizonal axis (of The Device) the tokens on the vertical axis don't form sets or kinds.

Let us remark that the idea of self-measurement will not work here, since it will collapse to either case (i) or to case (ii).

(iii) If, nevertheless (as friends of multiple realizability insist), the genuinely heterogeneous A and B share some fact, namely: the fact of "being pain", that is: if *genuine multiple realizability* of the mental kind P by physical kinds obtains, then A and B must share a non-physical fact. Each of them has this non-physical fact, independently of whether or not the other ever obtains (indeed, we feel pains directly in individual cases regardless of whether or not other pains obtain in the world). Thus, each *token* contains a non-physical element. It is immaterial for our argument whether this non-physical element is a property or a substance (and therefore we don't address this point here).

This is why, according to Flat Physicalism, the special sciences of psychology are radically different from the other special sciences (like biology or economics) in that the psychological kinds cannot even *apparently* be multiply realizable. We emphasize that this special nature of psychology has nothing to do with the so-called hard problem of consciousness or related issues.

In their defense of the idea of how high-level kinds can be multiply realizable by physical kinds, Davidson (1970), Fodor (1974, 1997), and Putnam (1967) seem to argue that multiple realizability is a consequence of the fact that the high-level special sciences sets (or kinds) are formed by "brute enumeration" (see Fodor 1974 who follows Davidson on this point). Fordor (1974), for example, argues that it is a brute fact that tokens of heterogeneous physical kinds form a high-level set (or kind). Why should this be problematic? Our main point here is *not* that this idea by itself is logically incoherent; perhaps it is logically coherent (we don't take a stand on this question). Our point is that this idea is *incompatible* with another idea in physicalism, according to which the set of *all* possible tokens (e.g., the set of all possible microstates according to contemporary physics) is *all* that can possibly exist in the world, so that the sets (or the kinds) are determined completely by the tokens and nothing else, because there is *nothing* else. The "brute enumeration" (or facts) idea is that the sets are *not* determined by the tokens, but rather (somehow) by the enumeration of their members. But this is misleading, since the critical question is: *how* these sets are formed, not how they are enumerated after they are formed (as it were). That is: given the set of all possible tokens, what in the *tokens* (or which facts about the tokens, which, again, are all there is) make it the case that some partitions into the sets obtain and some do not obtain. Sets determined by shared features of the tokens are formed by the facts about the tokens and nothing else. By contrast, sets formed by "brute enumeration" require *an enumerator*, that is, they require some *thing* or some *feature* or some (brute) facts external to (that is, over and above) the set of all possible tokens, which determine the enumeration, which in turn gives rise to the sets. And this (as we argued above) is token-dualism (substance or property) in disguise, since one has here *both* the tokens and the enumerator or the fact that determine that the enumeration is such-and-such.

If genuine multiple-realizability is a coherent idea (as many believe it is), then it could be true of psychological and cognitive states in our world; but then, importantly, mind-body dualism would be true of the world. In that sense, the empirical discovery of genuine multiple-realization would amount to an empirical discovery that psycho-physical dualism is true of the world.

6. Functionalism as token-dualism about token-sequences

Functionalism is sometimes taken to solve the main problem that we presented in Section 4 above. Recall: *The central question concerning reductive (versus non-reductive) physicalism* is this. Consider three tokens A, B and C, such that A and B belong to the same special sciences kind M but C doesn't. What is the fact in virtue of which this is the case? Which facts in the world fix the partition of tokens into types? Accepting that it may happen that A, B and C might be equally physically heterogeneous, so that the fact that A and B (but not C) are both M cannot be explained by a shared physical aspect, functionalists opted for the following explanation: A and B share the same "functional role". It may happen that two physical histories have the same functional structure, they emphasize, that is "realized" (this is a popular term in this context) in heterogeneous physical ways. Figure 2 illustrates this case: the states on the left hand side, indicated by numbers, are physically heterogeneous from the states on the right hand side, indicated by letters; but the structures are the same. Two tokens are deemed of the same "functional kind" if they occupy the same role in the functional structure, for example: State 1 and State A.[19]

[19] Whether the input and output states can be multiply realized is a subject of debate, that we don't address here.

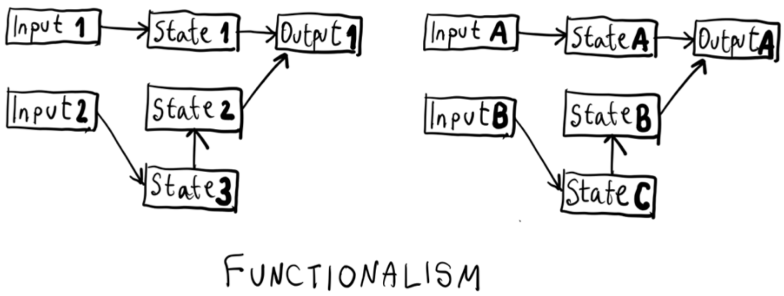

Figure 2: Functionalism

We would like to very briefly note that in this sense functionalism doesn't solve the problem but repeats it. As we said above, there are two sorts of tokens: state tokens, and sequence tokens; and both are fully describable (ideally, of course) by physics. From the perspective of physics, a "function" is nothing but a set of token-sequences. Consider three token-sequences: A', B', and C'; and suppose that A' and B' are of the same "functional kind" but C' isn't. What fact makes this the case? If token-sequences are physical, then there are two options. Either (I) sequences A' and B' share a physical aspect (that C' doesn't share with them); this is a reductive type-identity understanding of this case, analogous to the reductive type physicalist understanding of kinds of token-states, discussed above; or (II) A' and B' are heterogeneous, in which case in order to subsume them under the same "functional kind" we need a Sequences-Device, analogous to the one used above in the case of token-states. So Functionalism (like any approach that allows for multiple realizability) is either a version of reductive type-identity physicalism described by Flat Physicalism, or a version of token dualism.

We conclude with Kim (2012, pp. 177-8):

> Token physicalism... is no physicalism - unless of course one lets it collapse to type physicalism.

7. The autonomy of the special sciences

We have shown that the full-fledged reductive physicalist identity theory of Flat Physicalism is the only theory compatible with physicalism, according to which every token occurrence in the world is completely physical. If the special sciences kinds aren't strictly identical to physical kinds (however complex), as in functionalism of all versions, then there is something non-physical in every individual token, so that token-dualism follows. Fodor (1974, 1997), wishing to ensure autonomy for the special sciences, opted for token dualism, which underlies his ontology of multiple realizability. But this move was unnecessary, since *Flat Physicalism guarantees an autonomy for the special sciences*, which is in line with actual scientific practice.

As we said above, every token (state or sequence) has infinitely many aspects, but only some of them appear in our experience and in our theories. The proverbial Laplacian Demon, who knows the complete micro-trajectory of the universe and – if told which degrees of freedom of the universe are "us" – can calculate which are the aspects of our environment to which we are sensitive. Moreover, let us add that, presumably, only some of the aspects satisfy regularities; and Laplace's Demon can tell us which those are, thus predicting the macroscopic laws of nature that we shall experience, that is: what would be our special sciences, the special sciences kinds, and the special sciences laws. Laplace's Demon is in fact a personification of the fundamental theories of physics, and so the above is a way of saying that the nature of the world, as described by physics, determines what would be our special sciences, the special sciences kinds, and the special sciences laws. These facts are, then, not mysterious, as far as physicalism is concerned.

However, importantly, and as a matter of principle, we are *unable* to discover these facts, and not only because the immense complexity of the world, but also because we ourselves are part of it, and the very act of making predictions is a *physical* process in our brain, that is, in the world that is to be predicted. Thus, the only way in which we can learn which are the aspects of the world to which we are sensitive and what are their regularities, and what other aspects – accessible only via complex measuring instruments – satisfy interesting and useful regularities, is the job of the special sciences. Physics cannot do this. This fact underlies the autonomy of the special sciences.

Unlike the mysterious nature of this autonomy as described by, for example, Fodor (1974, 1997), according to Flat Physicalism the autonomy of the special sciences is a fact that has a straightforward physical explanation. The fact that the special sciences laws are autonomous in the above sense supports our view that a psycho-physical identity theory doesn't entail eliminativism with respect to the special sciences kinds.

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 their regularities 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 Flat Physicalism.

(1) The second law of thermodynamics is time directed, despite the time symmetry of *all* the laws of fundamental physical.[20] This directedness can be accounted for in terms of non-temporal local asymmetries (see Hemmo and Shenker 2019a).

(2) 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., Loewer 2001; Frigg and Hoefer 2015; List and Pivato; Hemmo and Shenker 2019b).[21] This can be accounted for in terms of the harmony between the micro-trajectories and the partition of the state space into sets according to the aspects to which the measuring devices are sensitive in the sense explained above; see Hemmo and Shenker 2015; Shenker 2017a). This is illustrated in Figures 3a and 3b. 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).[22] The state space is further partitioned to sets of microstates that share an aspect to which a given measuring device is

[20] The 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!

[21] On 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 (Albert 2000); (Hemmo and Shenker 2012, Appendix).

[22] The sequence of grey rectangles in Figure3 are the Poincaré sections of the bundle of trajectories of the time evolution of that initial set of microstates, at several points of time.

sensitive. In Figure 3a the aspects to which the device is sensitive happen to be such that the device will detect a deterministic behavior, whereas in the case of Figure 3b the measuring device is more sensitive, i.e., it can distinguish between finer values of the aspects of the microstates, and the result is that this device will detect a probabilistic behavior: 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.

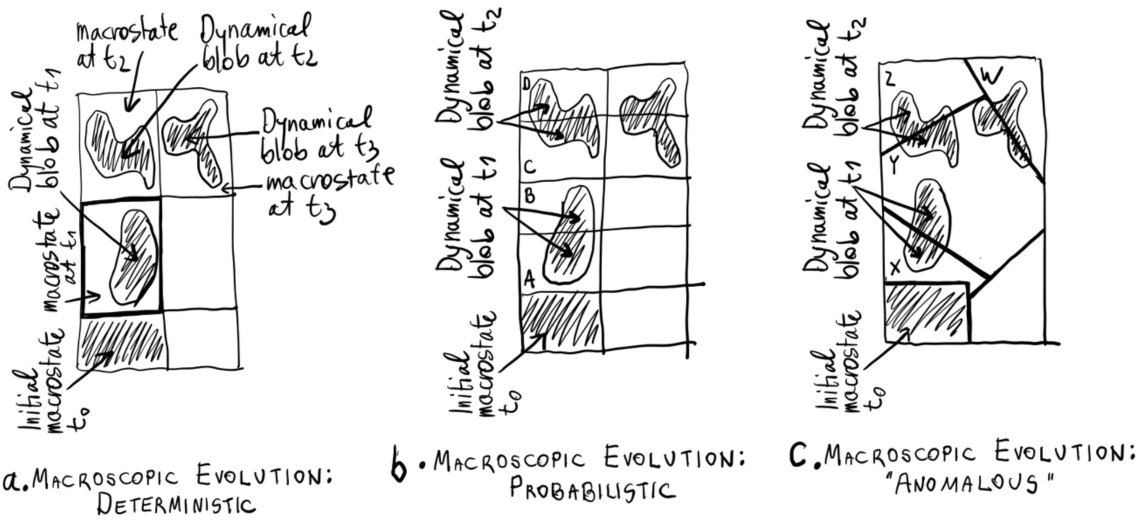

Figure 3: Dependence of macroscopic evolution on the interplay between dynamics and partition of the state space

(3) 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 and Shenker 2019d). Whether or not the mental is anomalous is a question of fact, and we don't purport to be able to answer it. Suppose, for the sake of the argument, that the mental is anomalous. The important point is that Flat Physicalism can account for this in a very simple way, along the following lines, illustrated in Figure 3c. For simplicity as well as for emphasizing our point, the starting macrostate and the dynamics in all three cases 3a, 3b and 3c are the same, and only the partitions (illustrated by the thin lines in the figure) of the microstates into regions or subsets of the state space are different in the three cases. The difference between the partitions in 3a and 3b is, as we already said, one of coarser graining and finer graining of the state space. In the third case (Figure 3c), there is even less harmony between the micro-evolution 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 in, for example, Figure 3a, but the dynamics is different (call it the 'alternative dynamics'), so that the result is as in case 3c: the alternative dynamics in this case may lack any macro-regularity of the aspects described by the 3a partition. Interestingly, given the alternative dynamics, which appears to result, under the 3a 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 3c one, or even a more complex-looking partitioning) under which the alternative dynamics will exhibit an interesting regularity. Discovering such highly non-trivial cases (of physical kinds) is the autonomous job of the special sciences.

Of course, Davidson (1970) wanted something stronger: he wanted some deep sort of anomaly of the mental. And so, since he argued (1970, p. 215) "no purely physical predicate, no matter how complex, has, as a matter of law, the same extension as a mental predicate," and since (he assumed) mental kinds supervene on physical kinds (see his 1970, p. 214, and his 1993)[23], he allowed for *genuine* multiple realization of mental kinds by physical kinds. But importantly, Davidson's ontology *isn't physicalist at all*: his fundamental events are neutral, and are physical only under a description. In (Hemmo and Shenker 2019d) we show that there are two possible interpretations of Davidson's picture, in one of which his notion of "description" assumes a rather strong form of substance token-dualism, because it requires adding a *descriptor* to the event structure. If that is the ontology, deep anomaly shouldn't be surprising. However, we propose another way of interpreting Davidson's approach which seems to be in line with his attempts to come up with a *non-physical* but *monistic* approach, but on this interpretation the mental and physical features of the events that fix the partitions to sets (or kinds) are fixed by the spatio-temporal and causal structure of the events, and in this sense (unlike the first interpretation) they are inherent to the events in the low level. We consider in what sense why such an approach is non-physical and in what ways it differs from a reductive physicalist approach, such as Flat Physicalism; and we argue that contrary to Davidson's intentions this approach in fact collapses to a type-identity physicalist theory.

8. Levels of reality

According to Flat Physicalism (as its name indicates) the world is absolutely, categorically, unequivocally, *flat*: there are no so-called "*levels of reality"* ("levels of explanation" is only a *façon de parler* and has no bearing on reality). In this paper we have assumed that: (i) reality is as described by physics, the toy model being a world fully and correctly described by classical mechanics; and that (ii) this is all that there is in reality, that is, the description of reality as given by microphysics is complete. Physical aspects of the microstate of the universe are parts of this complete reality, they don't stand for additional structure or facts over and above the complete physical microstructure. Sets (in the state space postulated by physics) of microstates that share aspects (called macrostates) don't exist in the world, for according to our best theories of contemporary physics, each such set consists of one *actual* microstate and (infinitely) many *counterfactual* ones, which aren't in the world.

The denial of levels in Flat Physicalism is quite radical. The prevalent sentiment in contemporary philosophical literature on this matter is represented in the following quotation (brought in a different context - that of asking whether the relation of "grounding" is fundamental).

> I have no knockdown argument against the claim that the world is flat. But every fiber of my being cries out in protest. ... The true flatworlder... denies that there are *any* non-fundamental properties, and, indeed, ... she denies that there are states of affairs, she denies that there are sets, she denies that there are people. ... any version of flatworldism will be radically revisionary. I repeat that I have no real argument against it. I will simply say that flatworldism is, to borrow a colorful word from a friend, "crazypants". An imprecise complaint to be sure, but it is my complaint nonetheless. It is a cousin of the incredulous stare. (Bennett 2011, p. 28)

[23] By supervenience of the mental on the physical, Davidson (1970, p. 214) assumed that: "there cannot be two events alike in all physical respects, but differing in some mental respect".

Bennett (2011) is very clear that her rejection of "flatworldism" is (at this preliminary stage, at least) a matter of intuition (unlike, for example, the argument in Fodor 1974, 1997, mentioned above). Our intuition, that guides us in this paper, is the opposite one: we find it hard to see how levels of reality can come to exist, and what sort of existence is meant by existence in different levels. The very fact that opposing intuitions are available makes it easier to discuss the matter. Here we shall not argue for (nor against) flat reality: our main task is to show what a flat reality amounts to and how it is a consequence of Flat Physicalism as described above.

By "*levels of reality*" we mean (in this paper) this: A multi-level structure of reality is one in which there are matters of fact at a relatively high level that aren't part of a relatively low level. Not vice versa: a case in which there are matters of fact at a low level that aren't part of a higher level will not be treated, in this paper, as one in which reality is multi-leveled; below we explain why this is so.

Terminological remarks: (1) In the literature there are various other notions denoted by the term 'levels'; for example, there are 'levels of explanation', and other notions. We don't address them (see, for example, Albert 2000, 2014; Loewer 2001, 2012; Craver 2007; Craver and Bechtel 2007; Frigg and Hoefer 2015, List and Pivato 2015; Bechtel 2016; List 2019). (2) The terms '*matters of fact'* or '*facts'* are intended (in this paper) to be wide and inclusive, and the way we think of them will hopefully become clear as we proceed. (3) The term '*relatively'* is intended to put to the side, in this paper, the question of whether or not there is a fundamental level of reality. Hereafter we shall omit the term 'relatively' in this context, for simplicity. To make things even simpler we shall assume for the sake of the argument that there is a fundamental level, and call that level "physical".

Often when people discuss the multi-levels structure of reality (either in the above sense or in other senses) they do so in the context of characterizing reality in terms of two concepts: whether or not *multiple-realizability* of special science kinds by physical kinds obtains, and whether or not *supervenience* of special science kinds on physical kinds obtains.

Consider Figure 4: in the two cases described there supervenience obtains, but whereas in case 2 multiple-realization is allowed or possible (the special kind X is realized by the physical kinds A and B), in case 1 it isn't. Supervenience is often taken to be the hallmark of physicalism, giving physics a preferred status over all the special sciences (see for example, Kim 2012). In terms of Flat Physicalism, in case 1, tokens 1 and 2 are both of the kind X in virtue of their sharing a physical fact, specifically: they share the aspect A of their microstate. They are partially identical, and the aspect in which they are identical *is* the kind X. Tokens 3 and 4 don't have this aspect, and it is therefore that they don't belong to the kind X. To make our point clear, here is an example.

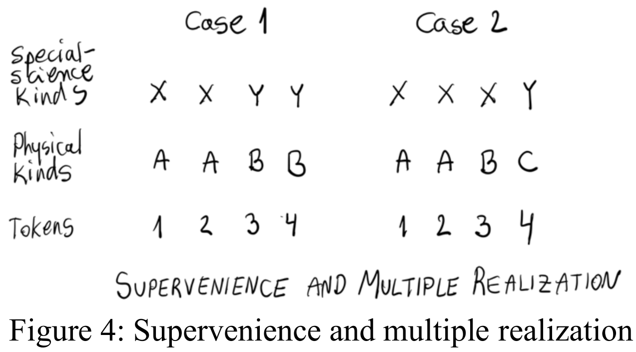

Figure 4: Supervenience and multiple realization

(Terminological remark: Some people call case 1 "multiple-realization" by which they mean that the property X is shared by the different tokens 1 and 2 that, although they share the aspect A, differ in other aspects. There is no point in arguing about terminology, as long as the core of the matter is understood, confusion between cases 1 and 2 is avoided, and the way in which genuine multiple realization is a form of dualism is understood. Since *we* find this terminology confusing, we avoid it, and emphasize that 1 is a case of strict identity (between the kind X and the aspect A that is present in each of the tokens 1 and 2 but not 3 and 4), whereas Case 2 is a case of genuine multiple-realization.)

Case 1 is the *only* case in which reality can be flat, that is, without multiple levels (in the above sense of this term).[24] In case 2, by contrast, reality is leveled, since in order to make tokens 1, 2, and 3 belong to the same kind X despite the fact that token 3 doesn't share a physical aspect with tokens 1 and 2, another fact – beyond that of the complete physical tokens – needs to be brought in. This additional fact might be The Device mentioned above; but if The Device is us, and its pointer states are our brain-cum-mental states, then the extra non-physical fact has to be external to physics. Adding a non-physical fact brings about another layer, so to speak, of reality.

8. Conclusion: functionalism as a form of dualism, and some consequences

Are there facts in the world over and above those described by physics? This is a question of fact, and we don't know what is the truth about the world; hence we don't attempt to prove reductive physicalism nor to disprove functionalism or any other form of dualism. Our task is, rather, *to identify which views are forms of dualism and which aren't*, and in particular, to point out that views that call themselves "non-reductive physicalism" aren't physicalist *at all*, since they posit the existence, as part of each individual token, of certain non-physical facts, that fix the special-sciences sets to which this token belongs.

The conclusion from what we have seen so far is that if multiple-realization is allowed, *even if supervenience is required*, then the matters of fact that make it the case that a certain tokens (but not others) belong to certain kinds (rather than others) aren't, and cannot be, physical matters of fact, but must be non-physical matters of fact. Importantly, in the case in which both multiple-realization and supervenience hold, the non-physical matters of fact must be *part of every token*, so that it would be wrong to say that each token is physical and nevertheless the (so-called high-level) kinds cannot be fully accounted for by the physics of the token. We called this *token-dualism*. And so, our claim is that *allowing for multiple-realization, even with supervenience, entails (or just is) token-dualism*.

The term *dualism* isn't meant pejoratively; for all we know dualism might be true about the world! It is very rich and has a variety of meanings in the literature. Our aim is to show that some approaches which present themselves as forms of physicalism are, in fact, forms of dualism. This may lead to conceptual confusion. And so, *our task is a clarificatory one: to clarify the metaphysical commitments of some contemporary approaches*: which views are dualistic and which aren't in the following sense: In this paper we understand *dualism* along the lines of the slogan often used to describe *physicalism*: if physicalism is the view that *everything is physical*, then dualism is the view that *some things aren't physical*.

[24] See also arguments for a flat picture in (Bechtel 2016).

While we didn't argue for reductive physicalism in this paper, it may be significant to add a word of motivation here (but our argument is totally independent of this motivation). The view that accepts, seriously and authentically, the idea of multiple-realizability, is a hindrance to the advancement of science, because it says that there is no point in searching for the common physical feature of cases that appear to be of multiple-realization. Fortunately, working scientists don't accept this idea and continue with their research, and whenever they find shared feature they see it as progress – a judgement that would be meaningless for friends of multiple-realizability (see for example, cases described in Polger and Shapiro 2016).

Our main target, in addressing forms of non-reductive physicalism, is the various forms of functionalism, including computational functionalism. Since functionalism requires, or is at least compatible with, multiple-realizability of functional kinds (for example, causal kinds) by physical kinds (including biological and psychological kinds), our proof shows that *functionalism is a form of token-dualism*, in the above sense of the terms. This argument only depends on multiple-realizability. It is independent of other features of functionalism, in particular computational functionalism, which are also in tension with physicalism, such as the multiple-computations thesis (also originally due to Putnam 1988). According to this well-established thesis (even only in its weakest form[25]), the time evolution of every open macroscopic system implements more than a single computation. This result means that functionalism implies a violation of *supervenience*. There is a debate about whether or not this kind of failure of supervenience is problematic (see, for example, Chalmers 2011, Shagrir 2012); many contemporary thinkers in this field take this result as seriously challenging the functionalist theory of mind. It has been recently shown (see Hemmo and Shenker 2019b) that the only solution to this problem that is compatible with physicalism is a *type-identity* theory in which the question of which function (or computation) is implemented by a given system is *uniquely* determined by the sequence of brain states which are identical to the mental states of an observer. But our argument here that functionalism implies token-dualism doesn't depend on any considerations related to multiple-functions or multiple-computations, but rather only on the assumption that mental kinds are multiply-realizable by heterogeneous physical kinds.

We have shown that this assumption alone implies that in each and every token-sequence of microstates realizing a high-level functional kind (or implementing a certain computation), there must be some non-physical sequence of facts in virtue of which it belongs to that functional kind. This conclusion pertains to *all* forms of functionalism, including computational functionalism. Putnam (1975) famously said that "the functional-state hypothesis is *not incompatible with* dualism!".[26] We have argued in this paper for a much stronger conclusion: the "functional-state hypothesis" *entails* token-dualism or *is a form of* token-dualism, and is therefore *incompatible with physicalism*.

Acknowledgement

References

---

[25] This result holds even if one adds various counterfactual and other constraints on what counts as a physical implementation of a function; see, for example, (Godfrey-Smith 2009), (Schuetz 2012), (Piccinini 2015), and (Piccinini 2017) for a recent overview of this issue.
[26] The reason for this is, in his words, that: "Although it goes without saying that the hypothesis is "mechanistic" in its inspiration, it is a slightly remarkable fact that a system consisting of a body and a "soul," if such things there be, can perfectly well be a Probabilistic Automaton."

Albert, David (2000) *Time and Chance*, Cambridge, MA: Harvard University Press.
Albert, David (2014) *After Physics*, Cambridge, MA: Harvard University Press.
Atkinson, David (2006) "Does Quantum Electrodynamics Have an Arrow of Time?" *Studies in History and Philosophy of Modern Physics*, 37 (3), 528-54.
Bechtel, William (2016) "Explicating Top-Down Causation Using Networks and Dynamics," *Philosophy of Science* 84, 253–274.
Bennett, Karen (2011) "By our Bootstraps," *Philos Perspect* 25, 27–41.
Bickle, John (2010) "Has the Last Decade of Challenges to the Multiple Realization Argument Given Aid and Comfort to Psychoneural Reductionists," *Synthese*, 177, 247–260.
Brown, Robin Gordon and Ladyman, James (2019) *Materialism: A Philosophical Inquiry*, London and New York: Routledge.
Chalmers, David (1996) *The Conscious Mind*, Oxford University Press.
Chalmers, David (2012) "The Varieties of Computation: A Reply," *Journal of Cognitive Science* 13, 211-248
Chang, Hasok (2012) *Is Water H2O?* Dordrecht: Springer.
Crane, Tim (2017) "How We Can Be," *Times Literary Supplement* 5956, 7–8.
Craver, Carl (2007) *Explaining the Brain: Mechanisms and the Mosaic Unity of Neuroscience*, New York: Oxford University Press.
Craver, Carl and Bechtel, William (2007) "Top-Down Causation Without Top-Down Causes," *Biology and Philosophy* 22, 547-563.
Davidson, Donald (1970) "Mental Events," in Davidson, D. (1980) *Essays on Actions and Events*, University of California Press, pp. 207– 227.
Davidson, Donald (1993) "Thinking Causes," in John Heil and Alfred Mele (eds.), *Mental Causation*, pp. 3-18, Oxford: Clarendon Press; reprinted in Davidson 2005 *Truth, Language and History: Philosophical Essays,* Oxford: Clarendon Press.
Elpidrou Andreas (2018) "The Character of Physicalism," *Topoi* 37, 435-455
Frigg, Roman (2008) "A Field Guide to Recent Work on the Foundations of Statistical Mechanics," in D. Rickles (ed.), *The Ashgate Companion to Contemporary Philosophy of Physics*, London: Ashgate, pp. 99-196.
Frigg, Roman and Hoefer, Carl (2015) "The Best Humean System for Statistical Mechanics," *Erkenntnis* 80: 551–574.
Fodor, Jerry (1974) "Special Sciences: Or the Disunity of Science as a Working Hypothesis," *Synthese* 28, 97–115.
Fodor, Jerry (1997) "Special Sciences: Still Autonomous After All These Years," *Nous* 31, 149-163.
Gillett, Carl and Loewer, Barry (2001) (eds.) *Physicalism and its Discontents*, Cambridge University Press.
Godfrey-Smith, Peter (2009) "Triviality Arguments Against Functionalism," *Philosophical Studies*, 145, 273-295.
Hemmo, Meir and Shenker, Orly (2012) *The Road to Maxwell's Demon*, Cambridge University Press.
Hemmo, Meir and Shenker, Orly (2015)
Hemmo, Meir and Shenker, Orly (2016) "Maxwell's Demon," *Oxford Handbooks Online*, Oxford University Press.
Hemmo, Meir and Shenker, Orly (2017a) "A Quantum Mechanical Maxwellian Demon", in B. Loewer, B. Weslake and A. Winsberg (eds), *Time's Arrow and the Origin of the Universe: Reflections on Time and Chance: Essays in Honor*

of David Albert's Work, Cambridge, Mass.: Harvard University Press (forthcoming).
Hemmo, Meir and Shenker, Orly (2017b) "A Psychophysical Identity Theory: Flat Physicalism", paper read at the PSA2016 meeting in Atlanta, unpublished.
Hemmo, Meir and Shenker, Orly (2019a) "The Past Hypothesis and the Psychological Arrow of Time," *British Journal for the Philosophy of Science*, axz038, https://doi.org/10.1093/bjps/axz038
Hemmo, Meir and Shenker, Orly (2019b) "Two Kinds of High-Level Probability," *The Monist* 102, 458–477.
Hemmo, Meir and Shenker, Orly (2019c) "The Physics of Implementing Logic: Landauer's Principle and the Multiple-Computations Theorem," *Studies in History and Philosophy of Modern Physics*, forthcoming, https://doi.org/10.1016/j.shpsb.2019.07.001
Hemmo, Meir and Shenker, Orly (2019d) "A Dilemma for Davidson's Anomalous Monism", forthcoming.
Hoefer, Carl and Marti, Genoveva (2019) "Water Has a Microstructural Essence After All," *European Journal for Philosophy of Science* 9 (12), https://doi.org/10.1007/s13194-018-0236-2
Kim, Jaegwon (2012) "The Very Idea of Token Physicalism," in Gozzano Simone and Hill Christopher (eds.) New Perspectives on Type-identity, pp. 167-185, Cambridge University Press.
Kripke, Saul (1980) *Naming and Necessity*, Wiley.
Ladyman, James, and Ross, Don (2007) *Every Thing Must Go*, Oxford, Oxford University Press (with D. Spurrett and J. Collier).
List, Christian (2019) "Levels: Descriptive, Explanatory, and Ontological," *Noûs*, forthcoming.
List, Christian and and Menzies, Peter (2009) "Non-Reductive Physicalism and the Limits of the Exclusion Principle," *Journal of Philosophy* 106 (9), 475–502.
List, Christian and Pivato, Marcus (2015) "Emergent Chance," *Philosophical Review*, 124(1), 119-152.
Loewer, Barry (2001) "Determinism and Chance," *Studies in History and Philosophy of Modern Physics*, 32, 609-620
Loewer, Barry (2012) "Two Accounts of Laws and Time," *Philosophical Studies* 160 (1), 115-137.
Malcolm, Norman (1968) "The Conceivability of Mechanism," *Philosophical Review*, 77, 45–72.
Ney, Alyssa (2008) "Physicalism as an Attitude," *Philosophical Studies* 138, 1-15.
Papineau, David (1993) "Physicalism, Consciousness and the Antipathetic Fallacy," *Australasian Journal of Philosophy* 71, 169-183.
Piccinini, Gulatiero (2015) *Physical Computation*, Oxford University Press.
Piccinini, Gualtiero (2017) "Computation in Physical Systems," *The Stanford Encyclopedia of Philosophy* (Summer 2017 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/sum2017/entries/computation-physicalsystems/>.
Polger, Thomas and Shapiro, Lawrence (2016) *The Multiple-Realization Book*, Oxford University Press.
Portides, Demetris (2019) "Idealizations and Abstraction in Scientific Modelling," *Synthese*, DOI https://doi.org/10.1007/s11229-018-01919-7
Psillos, Stathis (2009) *Knowing the Structure of Nature*, Macmillan.

Putnam, Hilary (1975) “The Nature of Mental States,” in Putnam, H., *Mind, Language and Reality*, pp. 429–40, Cambridge: Cambridge University Press (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).
Putnam, Hilary (1988) *Representation and Reality*, Cambridge, MA: MIT Press.
Ramsey, William (2019) “Eliminative Materialism,” *The Stanford Encyclopedia of Philosophy* (Spring 2019 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2019/entries/materialism-eliminative/>
Russell, Bertrand (1925) Preface. In: Lange, F. The History of Materialism and Criticism of its Present Importance, Trans. E.C.Thomas. London: Routledge.
Schuetz, Mathias (2012) “What is it Not to Implement a Computation: A Critical Analysis of Chalmers’ Notion of Implementation,” *Journal of Cognitive Science* 13, 75-106.
Searle, John (1992) *The Rediscovery of the Mind*, Cambridge, Mass: MIT Press.
Shagrir Oron (2005) “The Rise and Fall of Computational Functionalism, in Y. Ben-Menahem (ed.), *Hilary Putnam: Contemporary Philosophy in Focus*, Cambridge: Cambridge University Press.
Shagrir, Oron (2001) “Content, Computation and Externalism,” *Mind*, 110, 369-400.
Shagrir, Oron (2012) “Can a Brain Possess Two Minds?” *Journal of Cognitive Science* 13, 145-165.
Shenker, Orly (2017a) Flat Physicalism: Some Consequences. *Iyyun: The Jerusalem Philosophical Quarterly*, 66, 211-225.
Shenker, Orly (2017b) “Foundations of Statistical Mechanics: Mechanics by itself”, Philosophy Compass, https://doi.org/10.1111/phc3.12465
Shenker, Orly (2017c) “Foundations of Statistical mechanics: The Auxiliary Hypotheses”, *Philosophy Compass*, https://doi.org/10.1111/phc3.12464
Shenker, Orly (2018) “Foundations of Quantum Statistical Mechanics”, in E. Knox, and A. Wilson (eds.), *Routledge Companion to the Philosophy of Physics*, Oxford:Routledge (in press).
Sklar, Lawrence (1993) *Physics and Chance*, Cambridge: Cambridge University Press.
Smart, John Jamieson Carswell (1959) “Sensations and Brain Processes,” The Philosophical Review 68, 141-156.
Smart, John Jamieson Carswell (2017) “The Mind/Brain Identity Theory”, The Stanford Encyclopedia of Philosophy (Spring 2017 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2017/entries/mind-identity/>.
Tiehen Justin (2018) “Physicalism,” *Analysis* 78, 537-551.
Uffink, Jos (2007) “Compendium to the Foundations of Classical Statistical Physics,” in J. Butterfield and J. Earman (eds.), *Handbook for the Philosophy of Physics*, Part B, pp. 923-1074.
Wallace, David (2001) “Implications of Quantum Theory in the Foundations of Statistical Mechanics. Unpublished manuscript. Preprint available at http://philsci-archive.pitt.edu/410/