

A Dilemma for Davidson's Anomalous Monism

Meir Hemmo[†], Orly Shenker[‡]

Abstract

Is freedom compatible with determinism? Davidson (in “Mental Events”) famously rephrased this question by replacing “freedom” with “anomaly of the mental”, that is, failure to fall under a law. In order to prove that the anomaly of the mental is compatible with other conjectures he makes, in particular that: (a) there is psycho-physical causation; (b) “where there is causality, there must be a law” (Davidson 1970, p. 208); and (c) the mental supervenes on the physical, Davidson proposed a model (i.e., an interpretation under which all these conjectures are true), that came to be known as anomalous monism. Accepting (as working hypotheses) all of Davidson’s conjectures, we compare the structure of Davidson’s argument with that of Einstein’s argument for the special theory of relativity. This leads us to an exposition of Davidson’s ontology in terms that are inspired by recent results in the philosophy of physics, that is, in terms of fundamental ontology and high-level coarse-grained descriptions. We explain in what sense Davidson’s model is a principle theory (in Einstein’s terms) and discuss some requirements that the constructive theory underlying Davidson’s principle approach must satisfy. We propose two constructive theories of description that may underlie Davidson’s approach and this deeper structure leads us to formulating a dilemma according to which Davidson’s approach entails either a non-physicalist type-identity reductive and monistic structure of events; or else it entails a structure of events that requires what we call token-substance dualism. We consider some issues which seem to suggest that the first horn of this dilemma collapses into a reductive type-identity physicalist theory, contrary to Davidson’s intent. Finally, we show how Davidson’s achievement of accounting for some freedom of the mental from the physical and the anomaly of the mental within anomalous monism can be achieved in a fully reductive type-identity physicalist theory.

Keywords: Anomalous Monism; description; dualism; non-reductive physicalism; principle vs. constructive theory; reduction; type-identity.

1. Introduction

Are we free? Notoriously, this depends on what you mean by “freedom”. Davidson (1970) understood freedom (or an essential element of it, or a necessary condition for it) as meaning that “mental events such as perceivings, rememberings, decisions, and actions, resist capture in the nomological net of physical theory”, in which the laws are strict. This idea came to be known as the *anomaly of the mental*. Whether or not the mental is indeed anomalous in this sense is a question of fact, and Davidson (1970) postulates that it is.¹ We do not purport to be able to either endorse nor deny this conjecture, and we accept it for the sake of this paper.

With Davidson our question is this: what kind of universe can give rise to an anomaly of the mental? The search for such a universe leads, first, to Davidson’s other assumptions.

[†] Philosophy Department, University of Haifa; meir@research.haifa.ac.il

[‡] The Hebrew University of Jerusalem; orly.shenker@mail.huji.ac.il

¹ Some explain this alleged fact by the normative nature of some mental processes; see e.g., (Child 1993). We do not address this line of thinking here.

One option of a universe in which the mental is anomalous is one in which there is no regularity at all; but it is not an interesting candidate. A more challenging one is a universe that exhibits regularity; more specifically, a world that is, by and large, as physics describes it. Davidson postulates that in the universe in which we are interested the laws of physics are strict.² Of course, given the strict laws of physics, one way to ensure an anomaly of the mental is to go for a universe in which the mental and the physical are uncorrelated. But again, this is not an interesting case. A more challenging case is one in which the mental and the physical are connected in the way that they seem to be in our world. Here Davidson demands two sorts of connections. First: assume that there is mental-mental and mental-physical bi-directional causation (where according to Davidson “events related as cause and effect must fall under strict laws”; we explain this below). Second: mental kinds supervene on physical kinds.

To show that at *all* these requirements or assumptions or postulates are consistent with each other, Davidson constructs a *model* in which (he argued) they are all true. Here, by “*model*” we refer to the notion in semantics, in which a set of formulae are mutually consistent if there is an interpretation – a model – under which they are all true. Davidson’s model came to be known as *anomalous monism*.³

While Davidson’s ideas are well known, in this paper we couch them in terms that make it easy to see that they are similar, in very interesting ways, to three major ideas in the philosophical foundations of physics. This way of looking at Davidson’s model enables us to expose some interesting, and even surprising, features of Davidson’s approach, as well as point out some counterpart ideas within reductive physicalism.

(I) First, we point out the structural similarity between Davidson’s (1970) argument for his anomalous monism and Einstein’s (1905a) argument for the special theory of relativity. Both, as we show, start out with the observation that several statements for which there is ample empirical and theoretical support, appear to be mutually incompatible, and continue to propose models in which all of them can be accepted as true.

(II) Second, Einstein (1919) saw his theory of relativity (1905a) as a “principle theory”, for which a “constructive theory” is still in need of exploration (see also Schilpp, 1949). He compared the status of the special theory of relativity to that of thermodynamics which is a paradigmatic “principle theory”, but which has a well known “constructive theory”: statistical mechanics. We show that Davidson’s anomalous monism is, too, a “principle theory”, and we explore what its “constructive theory” might be, and propose two such theories.

(III) Recent results in the foundations of statistical mechanics have clarified the distinction between “microstates” and “macrostates” as well as “macrovariables” (see e.g., Hemmo and Shenker 2012, 2016, Shenker 2017a,b and Ben-Menahem 1997, 2001, 2018.). We show that Davidson’s anomalous monism builds on the same kind of distinction, albeit in a different framework (see Shenker 2015). (For our present purpose only the general structure of these argument, and not their contents, is important. (See Hemmo and Shenker 2012, 2016, Shenker 2017a, 2017b.)

² By “strict” laws we include probabilistic laws.

³ See Malpas (2015) for an overview of the literature on Davidson’s work; and also Ludwig (2003).

(IV) Davidson's ontology is often classified as a *sui generis* member of the group of proposals that flourished around the same period, i.e., theories in which all the events are physical and still the mental is *not reducible* to the physical (e.g., Putnam 1967, Fodor 1974). The way we describe Davidson's model enables us to show that his anomalous monism has an important advantage in that it escapes some of the criticism mounted against the other non-reductive approaches.

The paper is structured as follows. We start in Section 2 by briefly presenting Einstein's distinction between theories of principles and constructive theories. In Section 3 we reconstruct Davidson's model of anomalous monism, showing that it is a principle theory in Einstein's sense. In Section 4 we expand on Davidson's principle theory. It is *well known* that indeed in this model the anomaly of the mental is possible despite the supervenience of mental predicates on physical predicates. *Our analysis contributes* in the following points. First, it enables us to show in more details under which conditions the mental becomes anomalous, thus exposing further assumptions made by Davidson (for example, assumptions concerning the precise sort of harmony between the causal net and the descriptions). Second, it enables us to show that the anomaly of the mental can be obtained *in exactly the same way* in a *reductive-type identity physicalist* theory. And third, our analysis enables us to clarify the difference between Davidson's non-reductive approach and those of (e.g.) Putnam and Fodor, and to point out why his model is not subject to problems faced by the others. In Sections 5 and 6 we address our main question: What kind of a constructive theory best underlies Davidson's principle theory, and we propose two different formulations of such a theory in the form of a dilemma: On one option the descriptions in Davidson's principle theory capture a *pre-existing* structure of events, and on the other option, the role of descriptions is *constitutive* in creating the kinds. Our dilemma reveals some non-trivial and surprising consequences concerning the ontological commitments of Davidson's approach. It turns out that Davidson may not be a monist physicalist after all.

2. A common argumentative structure for two principle theories: Einstein's special theory of relativity and Davidson's anomalous monism

Einstein began his seminal (1905a) paper, in which the special theory of relativity was first presented, by considering two propositions, often called the "*relativity principle*" and the "*principle of the constancy of the speed of light*". Briefly, the former principle states that *the laws of nature have the same form* in all inertial frames of reference, while the latter states that the speed of light in vacuum is the same in all such frames. These two propositions had significant empirical support which made it reasonable to accept them as true, yet at first glance they seemed to contradict each other. Einstein attempted to solve this problem by giving both *empirical* generalizations the status of postulates, and searching for a *model* (in the abovementioned sense of the term) in which they are both true, thus proving that they are mutually consistent after all. In this structure, Einstein's (1905a) argument for the special theory of relativity is parallel to Davidson's (1970) argument for his anomalous monism: Davidson, too, seeks a model that will prove that certain statements, taken as postulates, are mutually consistent.

Of course, since Einstein's (1905a) argument concerns a physical theory, he hoped to show that his model, namely the special theory of relativity, is not only a proof of consistency, but is also true of our world. This hope is now universally taken to have been fulfilled, by

empirically confirming highly non-trivial predictions stemming from this model. The model that Einstein proposed is highly counter-intuitive, as it entails the relativity of simultaneity and of temporal and spatial intervals. For Einstein (1919), a theory of this kind that elevates empirical generalizations to the status of postulates is a “*principle theory*”. Within the framework of such a theory one does not ask *why* the postulates are correct: this is the task of another theory, that should be a counterpart of the principle theory, and is a “*constructive theory*”:

“There are several kinds of theory in physics. Most of them are constructive. These attempt to build a picture of complex phenomena out of some relatively simple propositions. ... When we say that we understand a group of natural phenomena, we mean that we have found a constructive theory which embraces them. But in addition to this most weighty group of theories, there is another group consisting of what I call theories of principle. ... Their starting point and foundation are not hypothetical constituents, but empirically observed general properties of phenomena, principles from which mathematical formulas are deduced of such a kind that they apply to every case which presents itself. ... The merit of constructive theories is their comprehensiveness, adaptability, and clarity; that of the theories of principle, their logical perfection and the security of their foundation.” (Einstein 1919)

In short, a theory of principle describes *very general* regularities in phenomena, while a constructive theory suggests a deep structure of the world that explains these phenomena and regularities. Howard (2010) considers this distinction to be Einstein’s most important contribution to the philosophy of science.

The paradigmatic pair of such theories that Einstein cites is that of thermodynamics and statistical mechanics. With Einstein, let us have a brief look at them, in order to clarify our concepts. Thermodynamics systematizes phenomena in terms of laws such as the Second Law of thermodynamics, according to which entropy cannot decrease in isolated systems. Any constructive theory that suggests a deep structure whose conclusions do not fit the phenomena as described by the principle theory of thermodynamics would be rejected as empirically inappropriate:

“[Classical thermodynamics] is the only theory of universal contents concerning which I am convinced that, within the framework of the applicability of its basic concepts, it will never be overthrown.” (Einstein 1970, 33).

However, as usual, things are not so simple. The constructive theory that explains thermodynamic phenomena is statistical mechanics. The scientists who accept this theory posit that, the universe consists of particles that obey the laws of mechanics, and that the phenomena described by thermodynamics are a manifestation of what occurs on the deeper level of particles. This approach is that of a strict, monistic, physical, reductive identity theory.⁴ Statistical mechanics is considered a successful theory since it explains and predicts major aspects of the phenomena described by the thermodynamic theory of principle.⁵ It does, however, more than this: it produces non-trivial predictions that go

⁴ This identity theory view of statistical mechanics is emphasized in Hemmo and Shenker (2012, 2016). For overviews of the standard state of art in the foundations of statistical mechanics; see Sklar (1993); Albert (2000); Uffink (2007); Frigg (2008).

⁵ An example for Einstein’s contribution to statistical mechanics is his *annus mirabilis* paper (1905b).

beyond the descriptions of thermodynamics, and that have been confirmed empirically.⁶ Even though correspondence with the theory of principle is a criterion for the empirical suitability of the constructive theory, it is important not to take the empirical generalizations of the thermodynamic theory of principle *too* seriously, since the constructive theory – mechanics – predicts phenomena that could not have been noticed at all without it, and these predictions are borne out by experience (Callender 2001). The success of these additional predictions supports the hypothesis about the mechanical deep structure of the world and of our understanding of the way this structure is manifested in our experience. The constructive theory thus not only allows us to understand known phenomena, but also improves our ability to discover and predict new ones.⁷ This is the source of its power.

Einstein stresses that despite the empirical success of the two great theories of principle that Einstein considers in (1919) -- the special theory of relativity and thermodynamics -- they leave us with a sense of discomfort, of not understanding why things are the way that they are (see Schilpp 1949). This is especially pressing if the model makes counter intuitive statements, as in the special theory of relativity. *Why* is the speed of light in a vacuum a constant? *Why* does the entropy of the universe never decrease? We feel that these phenomena are not the last word in the explanation of the universe, and that there is a deep structure that can explain them – that is, we feel that constructive theories in these cases are in place and should be informative and explanatory. The formulation of such theories is extremely non-trivial, and the attempts to do so are ongoing.⁸

We now turn to argue that Davidson’s anomalous monism is a theory of principle in this sense, on a par with thermodynamics and with the special theory of relativity (provided, of course, that the factual claims it makes, as well as its non-trivial predictions, are empirically supported; a point we do not address here).

3. Davidson’s theory of anomalous monism: the principle theory

In this section we *reconstruct* Davidson’s principle theory, namely, the model of *anomalous monism* within which all of his postulates are true. The structure of Davidson’s argument in “Mental events” (1970) mirrors that of Einstein in his (1905a) paper. Like Einstein, Davidson begins by presenting factual statements that have empirical support, but also seem mutually inconsistent, and goes on to suggest a model in which they are all true. In brief, Davidson’s postulates include (i) that there is a causal relation between some mental events and some physical events as well as between some mental events; (ii) that events related as cause and effect must fall under strict laws; and (iii) that there are no strict laws about the relations among mental events (the mental is anomalous). In addition to

⁶ One of the important results is the possibility of Maxwell’s demon, which is proved in Hemmo and Shenker (2012). Whether or not this is compatible with experience is still under debate.

⁷ We believe that this includes understanding of the entropy cost (or, rather, lack thereof) of measurement and information processing, and, as a result, Maxwell’s demon. See Hemmo and Shenker (2012, 2013).

⁸ One open debate in the foundations of statistical mechanics and its account of the second law of thermodynamics concerns the “typicality approach” to this account. The major line of thought on this issue is represented by Goldstein (2012), while a critical alternative is suggested in Hemmo and Shenker (2015a). One of the debates connected with the special theory of relativity concerns the question of just what are the empirical postulates underlying the theory, especially whether the constancy of the speed of light has to be a postulate or whether it can be deduced from other principles; see, e.g., Drory (2014).

these postulates Davidson (1970, 1993) also assumes supervenience of mental kinds (or predicates) on physical kinds.

The supervenience of the mental on the physical, especially in its relation with the anomaly of the mental, has been discussed extensively in the literature (see for example Davidson 1993; and more recently Kim 2012; Shagrir 2012). Davidson assumed from the outset (see his 1970) that mental kinds (sets of events that fall under certain mental descriptions) supervene on physical kinds (sets of events that fall under certain physical descriptions), where by supervenience he meant that: "there cannot be two events alike in all physical respects, but differing in some mental respect" (Davidson 1970, p. 214). In addition, with respect to the relation between mental kinds and physical kinds, Davidson also assumed that mental kinds are multiply realizable by physical kinds. This means, as Davidson says (1970, p. 215) that "no purely physical predicate, no matter how complex, has, as a matter of law, the same extension as a mental predicate" (see also Davidson 1973). Davidson emphasizes that it is this assumption which makes the mental irreducible to the physical in his model, because despite the supervenience of the mental on the physical, mental kinds are independent of the regularity satisfied by physical kinds. This idea ultimately allows him to derive some freedom of the mental from the physical.

We find it useful to think about Davidson's ontological picture as consisting of two layers. In the first *fundamental layer* there are the basic or primitive elements of the ontology: *events* and *causal relations*. At the *second layer* the central notion is that of *descriptions* of events. The events are individuated by their spatio-temporal indices (see 1963, 1969, 2001). Events gain properties (except for spatio-temporal indices) or become instances of *kinds* only *under descriptions*.⁹ Most importantly, the descriptions induce partitions of the events into sets, so that the events in each set are equivalent under some description, and form a kind.

As an example for the relation between events and descriptions, consider now the mental and the physical sorts of descriptions, which are the most important ones for Davidson. In Davidson's model, the physical description of events is such that every event is physical (see Figure 1), that is, every event falls under some physical description. By contrast, not every event is mental. As we mentioned above, the relations between the mental and physical descriptions satisfy supervenience and multiple realizability.

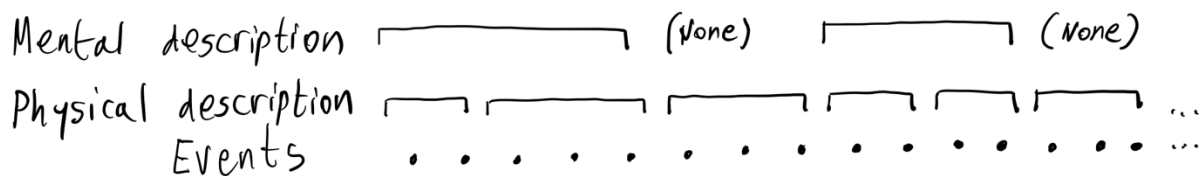


Figure 1:
Physical description covers all the events, mental description covers only some of the events

⁹ Davidson's thought on description has been influenced by Anscombe 1957, via the theory of action; see Davidson (1963); see also Steiner 1983.

There is another characteristic of the physical descriptions: the physical kinds satisfy strict laws (deterministic or probabilistic). This immediately raises the question of what is the connection between causal relations that hold at the fundamental layer between events and physical laws, that hold at the second layer between sets of events. Davidson postulate (ii) says that “where there is causality, there must be a law”¹⁰ (Davidson 1970, p. 208), and this leads to two ways of understanding the relation between the causal network and the physical laws.

(i) **A bottom-up reading:** One reading starts out by emphasizing that, in Davidson’s ontology, causal relations exist in the fundamental layer between the events *independently of description*, and in this sense the causal relations (and the spatio-temporal structure) are primitive and conceptually prior to the classification into sets, which comes about only at the second layer. To appreciate this point, it may be useful to recall that originally Davidson thought that events should be individuated by their position in the causal network, and later (following some criticism) changed his mind, taking space and time to be the individuating parameters. Since the laws of physics obtain between kinds that come about only under the physical description, causal relations are (by construction) conceptually prior to the laws of physics. This makes it possible to explain *why* the physical sets exhibit strict laws, as Davidson requires. On this reading, the physical partitioning of the events into sets *captures* a pre-existing causal structure. It is also possible in principle that as a matter of fact there will be no partitioning into sets that captures the causal relations and satisfies some strict laws, and it is also possible that there is more than one such partition. On this reading, whether or not for a given causal network there is a description (or more than one such description) that gives rise to strict laws is a question of fact; we don’t see any general theorem to this effect, one way or the other. Notice that in general such why-questions are to be answered by a constructive theory, not by the principle theory, so we already have a glimpse of the former here.

(ii) **A top-down reading:** On the other hand, and this is the second way of understanding the relation between the causal structure of the events and the physical laws, it seems that Davidson was *not* a reductionist about the relation between causation and laws. On this reading, falling under a strict law is a *pre-condition* for causation, so that it is the *second layer* that determines the causal relations (rather than the other way around, as in the first reading). That is: on this top-down reading one has at the fundamental layer spatio-temporal events which satisfy some connections between individual events. If one has in the second layer a partition of all possible events into sets that satisfies some strict laws, then and only then, the connections between the events at this bottom layer are called causal.

Henceforth we shall assume (for both readings) that there are partitions into sets which exhibit strict regularities, as assumed by Davidson with respect to the physical sets.

We are not sure about which reading fits better Davidson’s ideas about causation and laws. It seems to us that both appear in different places of his writings.

We now turn to the *mental* description. Two main features of the mental description are quite unlike the physical one. First, as we said, not all events are mental. The mental partition consists of a major partition and sub-partitions. The major partition is to events that are mental and events that are non-mental; and the former set is partitioned into sets

¹⁰ For Davidson’s approach to laws, see his *Causal relations* (1967). We don’t address this issue here.

corresponding to various kinds of mental states and events. Because of the major partition, and because (as we saw above) all the events have a physical description, we can say that some of the physical events are also mental, and other physical events are not mental. The second feature of the mental description is the anomaly: Davidson conjectures¹¹ that the mental sets (i.e., the sets of events in each of which all the events fall under the same mental description) do not exhibit regularities; this is his claim concerning the *anomaly of the mental*, postulate (iii). We will now describe how the anomaly of the mental comes about in this model.

4. Harmonies and disharmonies between the dynamics and the partition to sets

The different sorts of *harmonies* between the events and causal relations of the fundamental layer and the descriptions of the second layer (that give rise to various kinds), bring about different regularities pertaining to these kinds. Consider Figure 2 in which the space of all possible events is depicted, in three cases. In each case, the descriptions that induce the second layer partition the events into sets. In principle, the events can be partitioned to sets in infinitely many ways (since presumably there are infinitely many spatio-temporal events), and if we take each such partition to correspond to a description (under which the events in each set are of the same kind), then there are infinitely many descriptions. Since under different descriptions each individual event belongs (in general) to different sets, every individual event belongs to an infinite number of kinds at the same time (and indeed in Davidson's ontology an event can be both physical and mental).

For convenience we shall assume now the first reading of Davidson's approach. We shall later describe the figure in terms of the second reading. Given the primitive structure of the causal network between events, the different partitions to sets give rise to different sorts of regularities of the corresponding kinds. Here is an example of three sorts of regularities, that are the result of different sorts of descriptions, given the *same* underlying causal network of events which include *all* the actual events as well as the possible ones.

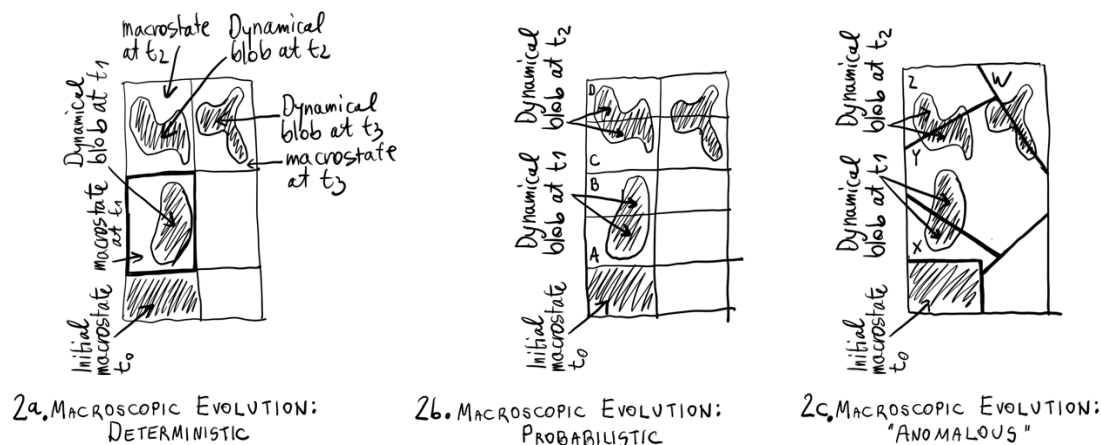


Figure 2: Regularities as expressing harmonies between descriptions and causal network

¹¹ This is sometimes explained as a consequence of normative and teleological notions like rationality. We do not address this point here.

Figure 2a depicts some description, inducing a partition of all (possible) events into six sets. Suppose that the universe (or a system) starts out, at time t_0 , in some event that is a member of one such set under this description, say the bottom left set (call it the *initial set*). In other cases, the universe might start out in a different event in the same kind, that is, in a different event that is within that same initial set. Each event in the initial set is connected to other events by the primitive causal relations, and – for every event in the initial set – the causal relations fix the event that will obtain after some time, say an event that will obtain at time t_1 (we assume determinism for simplicity, it seems to us that Davidson’s strict laws include probabilistic or chance laws). If we map all the events that start out at t_0 in the initial set, to the events that will obtain – according to the causal net – at t_1 , we obtain the set as marked in Figure 2a. And so on for t_2 , etc. Given the underlying causal network between events, what is important here is that the *sets* brought about under the description depicted in Figure 2a, exhibit a deterministic regularity. If the universe starts at t_0 in some event in the initial set, then given the regularity depicted in Figure 2a, at t_1 the event will belong to the left middle set. By contrast, the *sets* given by the description illustrated in Figure 2b, exhibit a *probabilistic* behavior: at time t_1 the event will be either of kind A or of kind B, under that description, with some probabilities, and at time t_2 the event will be of either kind C or kind D, with some probabilities. This shows that probabilistic laws describing the regularities of the kinds in the second layer are compatible with the determinism at the first layer (or more generally with different chance laws at the first layer). The description in case 2c cuts the events into sets in a more complex way, and it may be that following the different individual histories of the different events no regularity will be exhibited. These kinds would appear to be *anomalous* (as in Davidson’s mental kinds, to which we return below). Nomological relations between events are, on this picture, relations between them *in virtue of* their being elements in sets, under given descriptions, given the primitive causal relations in the fundamental layer. Given a certain primitive causal structure, different regularities in the world are expressions of different descriptions, that pick out different sets of events. (Ben-Menahem 1997, 2001, 2018 shows how these different harmonies give rise to a new notion of necessity that allows for degrees.)

The same analysis (*mutatis mutandi*) applies in the second top-down reading. On this reading the connections between the events as given by the time evolution from one event to the next (depicted by the shaded regions) are considered causal only if the descriptions in the second layer satisfy strict laws. This means that only the relations between the sets of events in Figure 2a and 2b are causal, while the relations between the sets in Figure 2c are not causal, and as we said, anomalous (i.e. don’t fall under strict laws).

Very importantly: there is a striking similarity between our construal of Davidson’s model of anomalous monism and statistical mechanics. The events in the first layer of Davidson’s model structurally parallel the notion of mechanical microstates, and the sets of events parallel the notion of macrostates, which are in statistical mechanics sets of microstates. We describe important dis-analogies between the two pictures in Section 6. Here is an interesting consequence of this analogy. The same structure of analysis of how anomalous mental behavior is compatible with strict physical laws (and the causal efficacy of the mental) applies if we replace Davidsonian events and kinds (sets) with physical *microstates* and *macrostates* (respectively) within a straightforward *physicalist type-identity* theory. According to such a theory there are only physical predicates and *a fortiori* supervenience of mental kinds on physical kinds is trivially satisfied. This is a novel and surprising result

which implies that all of Davidson's conjectures may be satisfied not only in some monistic universe in which the descriptions of the events given by physics are ontologically on a par with (say) the descriptions given by psychology (as Davidson has thought), but also by a strict type-identity physicalist universe in which there is only physics, and every predicate is reducible to a physical predicate (contrary to Davidson)! In particular, in such a universe multiple realization does not hold, and nevertheless the mental may be anomalous. We describe these ideas in much more detail in Hemmo and Shenker (2015b, 2019a, 2019b); Shenker (2015, 2017c).

Out of the many possible descriptions, that induce many possible partitions of the events into kinds, Davidson focuses on two, that he calls the *physical description* and the *mental description*: for him, events become “physical” or “mental”, or both, only *under descriptions*. At the fundamental layer, events are neither inherently physical nor inherently mental, and they become so only under the corresponding descriptions. (In Section 5 we address the question of why the focus of inquiry should indeed be (as it is) the physical and mental descriptions out of the many possible other partitions of the events into sets.)

Some features of the physical and the mental descriptions are necessary in order to ensure that Davidson's three postulates are mutually compatible; for this reason, those features of the physical description and the mental description are part of the *model*, they are part of the *principle theory* of anomalous monism. As can be seen in Figure 2, there are two senses in which mental kinds may be anomalous (given all of Davidson's assumptions, namely his three postulates, plus supervenience of mental kinds on physical kinds plus multiple realizability of mental kinds on physical kinds. First, mental kinds may be anomalous in the sense of Figure 2c above, that is: the underlying causal network is such that under the mental description no regularity arises. One has here a partition into the mental sets, which does not fit the time evolution as described by the physical laws.

The other sense of anomaly of the mental, which is perhaps what many have in mind, and which was of main interest to Davidson, is that (in our terms) there is no *harmony* between the physical description and the mental description in the following sense: If one follows the *evolution* of events under their mental description, but at each point of time (e.g. t_0 , t_1 , t_2 in Figure 2) describes each of them under its *physical* description, then (if there is no harmony between the two descriptions) *no physical regularity* will be exhibited. The total lack of harmony between the two partitions is such that the mental sets can be multiply realized by the physical sets and vice versa, so that none of them supervenes on the other. In Davidson's (1970) model, however, supervenience of mental kinds on physical kinds is assumed; Davidson (1970, 1993) rightly argued that this is *consistent* with the anomaly of the mental; that is, in our terms, supervenience of the mental sets on the physical sets is consistent with a *lack of harmony* between the mental and physical sets of the sort that appears in Figure 2c in both of these two senses of anomaly.

To sum up, let us see how Davidson's postulates turn out to be mutually compatible in the sense that they are all true in his model. We have just seen how the postulate of the anomaly of the mental is satisfied in Davidson's model (in both readings). Regarding postulate (i) of the causal efficacy of the mental: on the first reading of the relation between causation and laws (see above) postulate (i) is satisfied provided that indeed the two events (the cause and the effect) are causally connected in the *first* layer, which is a primitive fact about events. On the second reading, the connections (or correlations) between the two events are causal in virtue of the fact that all the events are physical and physics satisfies strict

laws. Similarly, postulate (ii) that events related as cause and effect must fall under strict laws is satisfied on both readings of the relation between causation and laws, again, in virtue of the fact that all the events are physical and physics satisfies strict laws.

The sense in which freedom (or its pre-condition of anomaly) is obtained in Davidson's model is this. Since mental kinds are assumed to supervene on physical kinds, the mental is *not* completely independent of the physical. But still the supervenience assumption does not say anything about which change in the physical sets is brought about upon a change in kind in the mental event. As we said, Davidson also assumes multiple realizability of mental kinds by physical kinds in the sense that mental predicates are in principle irreducible to physical predicates (see Davidson 1970, p. 215; Davidson 1973). Given multiple realizability, the freedom of the mental from the physical means that if one starts off from a given mental event, it is genuinely *indeterminate* on which physical kind the mental kind supervenes. That is, the physical set of events that realize a mental kind is fixed only by brute fact (see Fodor 1974, pp. 103-104 who refers to a remark by Davidson concerning the co-extension of mental and physical predicates). Moreover, freedom also means given Davidson's model that events belonging to the *same* mental kind may evolve over time (depending on the equations of motion) to disparate regions of the state space, so that they may in general lead to radically different mental kinds which will not be predictable. This sort of unpredictability applies even to Laplace's Demon precisely because of the multiple realizability: The complete information about the physical description and the Demon's unlimited computational capabilities are not enough to derive the partition of the events into the mental kinds since multiple realizability means that the mental predicates are not reducible to the physical predicates (see Hemmo and Shenker 2019b). However this sort of Davidsonian freedom is not full, because changes in the physical kind *are* predictable, so that if one knows the physical kind at a given time, one might be able to predict the way in which it will evolve over time to the next physical kind, and then by the supervenience relation one might be able to predict the way in which the mental kind may evolve (although only if enough details are given about the supervenience relation).

There is an important difference with respect to multiple realizability between Davidson's approach and other so-called non-reductive physicalist approaches, such as Fodor's 1974, and functionalists such as Lewis (1966) and Putnam (1967). On Davidson's approach the fact that every event is physical means that every event belongs to some physical set under the physical description, but otherwise physical descriptions are on a par in his model with mental descriptions, and both are high level descriptions. By contrast, in non-reductive physicalist approaches the *elementary* events are *physical*, and mental predicates are considered to be higher-level descriptions of the elementary *physical* events. We will get back to this point in the next sections.

5. The role of the constructive theory in Davidson's anomalous monism

In this section we wish to motivate the search for a constructive theory of descriptions that may underlie Davidson's principle theory of anomalous monism. The first reason is expressed by Einstein (1919). He conceded that his special theory of relativity is a theory of principle, but he was certainly worried about which constructive theory might underlie special relativity. For example, according to Einstein (Schilpp, 1949, p. 59), the use of rods and clocks in the standard principle view of special relativity introduces "two kinds of physical things, i.e., (1) measuring rods and clocks, (2) all other things, e.g., the electro-

magnetic field, the material point, etc.” This, Einstein says, “is inconsistent; strictly speaking measuring rods and clocks would have to be represented as solutions of the basic equations [of motion], not as it were as theoretically self-sufficient entities.” Indeed, Einstein, in Schilpp (1949, p. 61) calls this inconsistency ‘a sin’. Brown (2005) can be seen as an attempt in this direction in providing a constructive *dynamical* theory of special relativity; taking the geometrical structure of Minkowski spacetime as fundamental (excluding gravity) is another attempt in this direction (see Hagar and Hemmo 2013 for more on this point).

The second reason is this. A constructive theory that underlies Davidson’s anomalous monism would be a metaphysical picture that explains the fundamental layer of events and their primitive causal relations, and more importantly, explains *what are descriptions* and *how they come about*, and in particular how they come about in a way that gives rise to the various kinds and regularities that govern them. The account of descriptions and how they come about is not part of Davidson’s model; and this element of his principle theory calls for an expansion or a deepening of the picture. (Davidson’s thought on description has been influenced by Anscombe 1957, via the theory of action; see Davidson (1963). See also Steiner 1983. Some important implications of Davidsonian descriptions, and their connection to statistical mechanics and other scientific theories, are studied in Ben Menahem 1997, 2001, 2018; see also Shenker 2015.)

In their defense of the idea of how mental (and other special sciences) kinds can be multiply realizable by physical kinds, Davidson (1970), Fodor (1974, 1997), and Putnam (1967) have argued that multiple realizability is a consequence of the fact that the special sciences sets (or kinds) are formed by “brute enumeration.” Here is Fodor (1974) who follows Davidson on this point:

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 emphases). (Fodor 1974, pp. 103-4)

Why is the idea of brute co-extensivity problematic? If Fodor’s points in (b) and (c) in the quote above do not hold, then the question arises (a-c) what fixes the brute co-extensivity. In other words: What makes it the case that a particular event belongs to one set rather than another? The question is how the sets are formed, and not how do we know to which set an event belongs after the set is given. In statistical mechanics the sets are formed either by shared properties of the microstates belonging to the same set (called macrovariables), or by a correlation between the microstates in the same set and a physical measuring device. In this sense the enumeration of the sets in statistical mechanics is not brute.

What does it mean to say that we *experience* some types of events? To understand this, let us distinguish between an event of *experiencing*, and other events. In particular, we shall focus on the difference between ways in which these two sorts of events are given to us:

between *experiencing* an event as belonging to a type, and *recognizing* an event as belonging to a type. To make this distinction, consider two sets of events. One set consists of all the events in which *we* (or *you*, the reader) are in pain of some particular easily recognizable kind. Call this set “P”. Another set, call it “A”, is formed by collecting certain particular events, at random in the sense that we do not identify any common property that is shared by the members of A, except the very fact that they are all members of this set. In this sense the set A is purely extensional, and suppose that this is possible, just for the sake of the argument¹². Many agree that this is possible: in fact, extensional sets such as A underly the very idea of multiple realizability of high level kinds by a low level kind in contemporary non-reductive physicalist approaches (including functionalism), in particular the multiple realizability of mental kinds by physical kinds. This extensional character of a set is perhaps what Fodor (1974) means by “brute enumeration” of the physical events that fall under the extension of a mental predicate that is multiply realizable by different physical predicates.

Suppose now that a particular event takes place, and we need to say whether or not it belongs to A, and whether or not it belongs to P. To determine whether or not our event is a member of A we need to consult the *list* of events that established the set A. By contrast, to determine whether or not it is a member of P we do not need to do anything: we immediately and directly have the *experience* characteristic of events in that set. We know directly, without consulting any list, whether or not *we* are in this particular sort of pain. This is what we mean by knowing this *by experience*. The same thing happens when *we* (or *you*, the reader) experience an intentional event: we know directly, by experience, whether or not we now have some particular sort of thought. In this sense, the type A is not experienced, and therefore it cannot be a mental type, whereas P is (or can be) a mental type.

How can that be? What determines which sets of events will form mental types (since, recall, not all sets are mental!), and what determines what sort of a mental type will each set be, and – perhaps more importantly – when an event occurs, that belongs to a certain mental type, how can we *experience*, immediately and directly, to which type it belongs, without consulting any list? What makes such mental sets “P-like” rather than “A-like”?

Let us stress immediately, that an event cannot belong to a certain mental type in virtue of its properties – since (according to Davidson) events do not have properties (apart from spacetime indices and causal relations), except under a description. Recall: the fundamental ontology consists of events, individuated only by spacetime indices and interconnected by primitive causal connections, and the idea is that these events gain *all* the (other) properties only at the higher level, *in virtue of descriptions* (or under descriptions). Before these descriptions are given, events are neither mental nor non-mental. And now the question is this. Consider an event, of which we do not know whether or not it is mental. What fact will determine whether or not it is mental, and if it is – what sort of mental event it is? Can we derive this fact from the low level facts, concerning the spacetime indices and causal relations of the events? Especially: *how do descriptions come about? What are they?*

¹² It is debated in the literature whether or not multiple realization holds as a matter of fact in our world; see e.g. Polger and Shapiro (2015). But even those who think that it does not hold, take it that the very idea is coherent. Rejecting the possibility of purely extensional sets would undermine the coherence of the idea of multiple realizability.

Recall that we said in Section 4 that there is a striking similarity between statistical mechanics and our construal of Davidson's model. One of the most salient features of statistical mechanics is that the question of how the macrostates come about has a straightforward answer: namely, by what is called *macrovariables*, which are shared mechanical properties of all the microstates belonging to the same macrostate, and an observer that interacts with this particular macrovariable. We now proceed to examine what might be the constructive theory of Davidson's anomalous monism, and in particular what is the counterpart of the notion of macrovariables in such a constructive theory.

6. Constructive theory of anomalous monism: A dilemma for Davidson

We now propose two constructive theories of descriptions that may unrelie Davidson's principle theory of anomalous monism. These two theories are compatible with both readings of Davidson's approach to the relation between causation and laws (see Section 4). It seems to us that they are by and large exhaustive:

6.1 Option I for a constructive theory: Anomalous monism as a reductive *type-identity but non-physicalist* theory, that is (nevertheless) compatible with *psycho-physical multiple-realization*, regardless of whether or not supervenience of mental kinds on physical kinds holds.

On this constructive theory, descriptions do *not* add anything to the fundamental layer of the events, given by the spatio-temporal and causal relations: they are merely features of the fundamental-layer of events and causal relations. *Epistemologically*, when we want to talk about these features, we are giving the appropriate partial descriptions of the fundamental layer of events and causal relations; but *ontologically* the features exist regardless of whether or not anyone (or anything) is "describing" (or reflecting or representing, or enumerating, etc.) the state of affairs, so that the term "description" is only a *façon de parler* in order to make it easier for us to think about the ontology. There is no real act of description (as it were), there is no *descriptor*. Since the descriptions, in this sense, are nothing but features of the fundamental-layer of spatio-temporal events and their causal relations, the partitions to the mental and physical kinds are *entirely* fixed by the spacetime indices and causal relations.¹³

For this reason the relation between the first layer (of the primitive events) and the second layer (of the descriptions) on this picture is of *type-identity*: each (mental or physical) kind is *identical* with some spatio-temporal kind that satisfies certain (characteristic) causal relations.

Whether or not an event is mental, and which kind of mental event it is, *is a feature of the spacetime indices and the place of that event in the causal network*. Thus, ideally, by looking at an individual event and considering its spatio-temporal and causal features, one can see whether or not it is mental and what kind of a mental event it is. This way of thinking is analogous to the idea of Laplace's Demon in classical mechanics that sees the entire precise microstate of the universe and calculates its evolution, without being part of the universe; here the Demon is (nothing but) a personification of mechanics and its

¹³ Compare with the notion of "*aspects*" in the Flat Physicalism approach of Hemmo and Shenker 2019a, Shenker 2017c, and the notion of macrovariables in their work, see Hemmo and Shenker 2012, 2016, 2019b, see also Portides 2019, Ben Menahem 1997, 2001, 2018).

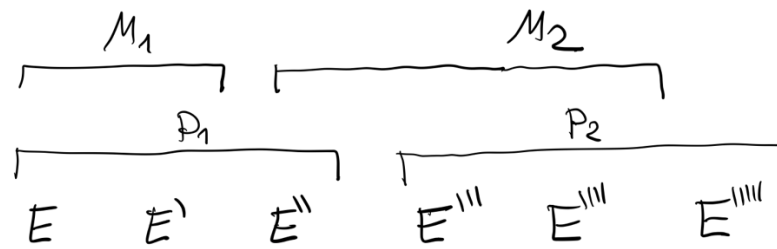
ontology. Likewise, the Davidsonian descriptions, on this constructive theory, are nothing but a sort of personification of the ontology, so that it is possible to understand why it is that (in the above example), upon encountering an event we are able to recognize whether or not it is of the experienced kind P or the extensional kind A that can be recognized only by consulting with a list.

Although the physical descriptions of the events are exhaustive (all the events fall under physical descriptions; see Figure 1), while the mental descriptions aren't exhaustive (only some but not all events are mental), this picture is *non-physicalist* (arguably; we challenge this point below). Since all possible sorts of descriptions (e.g., the physical and the mental) are equally (entirely) fixed by the fundamental layer, Davidson's idea that the primitive events at the fundamental layer are not *inherently* physical is retained: the events are physical *only* under a description. In this sense, unlike physicalist approaches, physics has *no* ontological priority in this picture.

This non-physicalist character may clarify the idea of multiple realizability in Davidson's anomalous monism (see the end of Section 4). For many non-reductive physicalists the idea that mental kinds are multiply realizable by physical kinds seems "molto misterioso", as Fodor (1974) has put it, despite the fact that it is logically consistent with supervenience of mental kinds on physical kinds. The reason is that in such approaches the fundamental level of events is given by physics, but if multiple realizability holds, the fundamental physical level *cannot* account for the high-level partitions of the events into the mental kinds (hence Fodor's "brute fact" brute enumeration answer). We have shown elsewhere (see Hemmo and Shenker 2019, 2019b; Shenker 2017c) that multiple realizability in non-reductive physicalist approaches (including functionalism of all sorts) implies that every individual mental event (or token-state) has some primitive *non-physical* feature in virtue of which it is of a certain mental kind. This conclusion does *not* apply to Davidson's approach on the constructive theory of descriptions we propose here, since the latter theory is indeed a type-identity reductive theory, according to which the kinds (under the various descriptions) are fully reducible to the spacetime parameters and the causal relations between individual events. Hence in this approach multiple realizability of mental kinds by physical kinds (assuming it is true) is not mysterious at all (contrary to Fodor 1974): it is a relation between *two higher-level* descriptions, where both levels can be fully explained by the fundamental spatio-temporal events and their causal relations. Since on this picture, all the kinds are reducible to spatio-temporal indices and causal relations, there can be *no* multiple realizability of any kind by genuinely different spatio-temporal kinds and their causal relations. Obviously, this route is not open in a physicalist approach where the fundamental level is given by physics.

This deeper structure gives Davidson the genuine freedom of the *mental* from the *physical* that he looked for (as we explained in the previous section). Still, the relation between the mental sets and the inherent (spatio-temporal and causal) features of the events that fix the mental partition is of *type-identity*; and similarly, the relation between the physical sets and the inherent features of the events that fix the physical partitions is also of type-identity (see Figure 3). In this way, the mental kinds and the physical kinds can exhibit *any* relation whatsoever, without breaking the *monism* of the event spatio-temporal and causal structure. In Figure 3 we depict the most general relation between the mental sets and the physical sets in which multiple realization of the mental sets on physical sets holds and supervenience fails. But note that on this constructive theory of descriptions Davidson's anomalous monism turns out to be non-physical even if one assumes supervenience of the

mental on the physical, since (on this theory) the fundamental layer of events is non-physical.



E, E', E'' share feature P_1 .
 E'', E''', E'''' share feature P_2 .
 E, E' share feature M_1 .
 E'', E''', E'''' share feature M_2 .
 M_2 is multiply realized by P_1, P_2 .

 P_1 realizes both M_1 and M_2 .
 E'''' is physical but not mental.

Figure 3: Davidson's monistic deeper structure

There are two major problems with this reductive proposal of a constructive theory of descriptions for Davidson's approach. First, this proposal does not seem to be in accord with Davidson's ideas about the *constitutive* role of descriptions: Davidson (1993) stressed that the events in the fundamental layer are neutral: they are neither inherently physical nor inherently mental; and if so, the partitions to kinds are induced (rather than expressed) by the descriptions. In other words, the partitions to kinds under descriptions are not redundant or short-hand expressions of a ready-made partitioned ontology, in which the kinds under the descriptions are fully reducible to the spacetime structure and causal relations. We reiterate that we do not find explicit textual support for this constructive theory in Davidson's writings, which focus on formulating the principle theory, not the constructive theory that explains it. We propose this constructive picture as one way of *making sense* of the notion of descriptions in his approach. As we said, we take our *inspiration* for this constructive theory from the structure of statistical mechanics.¹⁴ We make no conjecture as to whether Davidson would have approved of this version of his ontology; if he wouldn't, he would have to go for the alternative picture that we describe below.

The second problem is a serious threat that this constructive theory collapses to a reductive type-identity physicalist theory. Here is why. As we saw, the events in this theory are individuated by space and time indices, and a primitive causal network between them. But both elements are, essentially, the subject matter investigated by physics. Space and time, that individuate the events in anomalous monism, are non-trivial *physical* magnitudes. Our

¹⁴ Our own approach to the mind-body question, and more generally to the relation between physics and the special sciences is based on statistical mechanics as the paradigm theory of the special sciences. It is a reductive type-identity flat physicalist theory; see our 2019a).

understanding of the spatio-temporal structure has drastically changed with relativity theory, and some believe that it will further dramatically change in the future theories that will provide a consistent account of both quantum theory and general relativity. What some people take to be intuitive pre-theoretical concepts of space and time are, on this picture, merely concepts of old physical theories (see Ladyman and Ross 2007). So the spatio-temporal characterization of events is either dependent on physics on Option I or is so vague as to be useless. Similarly, for causation: to the extent that physics can make sense of causation (we don't argue that this question is fully settled; but see Ben Menahem 2018), in Davidson's approach causation requires strict regularity, and if strict regularity is to be found *only* in physics (although we don't think that there is an *a priori* argument to this effect), then causation in Davidson's approach would be dependent on physics, in which case Option I would be reduced to type-identity *physicalist* theory. As we saw above, one may change the conceptual order of priority and explain the strict nature of physical laws as emerging from the causal network and the particular partition of the events into so-called "physical" sets, if the latter exhibit strict regularities (see the top-down reading in Section 3), but on this view it is not a priori necessary that there should be a partition of the events into sets that would give rise to strict laws (and if so, it is not a priori necessary that that partition would be unique). We are not sure that this is compatible with Davidson's understanding of what causal relations are; it seems that for him the connection between the laws of physics and the causal nature of relations is much stronger, and if so, physics seems to enter the fundamental layer of events in Davidson's ontology.

The upshot is this: If, as proposed by Option I, the spatio-temporal and causal structures determine the partition of the set of events into kinds, then the partition is determined by events which seem to be *inherently* physical. That is, the events on this option are, for example, microstates in classical and quantum mechanics¹⁵, in which case the partition to sets is fixed by physical features (e.g., aspects called macrovariables) of each and every microstate in the state space. And so, the partition to sets will never give rise to multiple realizability of mental sets (kinds) by physical sets (kinds). In all known cases in physics, the partitioning of microstates into sets is determined by features (i.e., macrovariables) which are *inherent* to the microstates, such that sets of microstates share a common macrovariable! This leads immediately to a full blown reductive identity theory. So it seems to us that if the events on Option I are indeed physical – as we suspect they are (given that they are individuated by spatio-temporal and causal roles) – then Option I in fact collapses into a full blown reductive type-identity *physicalist* theory in which the mental is *reduced* to the physical, contrary to Davidson's intent.

6.2 Option II for a constructive theory: Anomalous "monism" as mind-body dualism.

The second option for understanding descriptions in a constructive theory is that they *create* the properties, that is descriptions create the partitioning of the events into sets, in a way that has *nothing* to do with the details of these events – in a way that *cannot be read off the events*. The idea here is to stick to the Davidsonian idea of leaving the events *neutral*: nothing in or about an event fixes to which kinds (physical, mental or what have you) it belongs; this is fixed solely by the descriptions. But if so, what are descriptions, and how

¹⁵ In classical mechanics the microstate of a system is given by a point in the system's phase space. In standard quantum mechanics the microstate of a system is given by a vector in the Hilbert space associated with the system. In both theories, the microstate is assumed to give the *complete* physical state of the system.

do they determine which event belongs to which set? In particular, which fact determines that a certain event is mental, and what kind of mental event it is?

It seems to us that the only available possibility is that the partitioning of the events to sets is determined by an *external* source, which we call here *the descriptor*. The descriptor describes each event as it chooses, without being constrained by anything about the events or their causal structure. When we experience an event *as* belonging to a certain mental type (e.g., we experience an event *as* a case of mental kind P), we either have *access* to the description by the descriptor, or *we are the descriptors*; in any case we must have access to something that is external to the event. Whether the descriptor is inside us, or even part of our mind or our body, or whether it is outside us, is a question which for our purposes we can set aside: the important point is that for each and every event for which one wishes to fix the kinds, the descriptor is not part of the event. And this is *substance dualism*: we have both the structure of neutral events and a descriptor which makes it the case that the events are partitioned into kinds in one way or another. This substance dualism is, moreover, a case of *token dualism*: to each and every token event an extra element of a “descriptor” must be added, in order to determine to which kinds that event belongs.

One may want to say (with Fodor 1974) that the extension of a mental predicate is fixed by some “brute fact” or by some facts about “enumerating” events in a way that determines to which kinds they belong. But this would be a far cry from avoiding the dualistic ontology: in this way of speaking we have facts about the spacetime indices and (perhaps, depending on the relation between laws and causation) facts about causation, which leave the events neutral, and *in addition* we have “brute facts” (about the enumeration; or an enumerator), which result in partitioning the events into kinds.

Ben-Menahem (1997, 2001, 2018) discusses “symbolic” description: for example, an object is a stop sign under such a description. We submit that such descriptions are, too, explained in the above two ways. By option I, “being a stop sign” is a feature shared by all the events in which there are stop signs, or in which there are people with mental states that recognize objects as stop signs. But Ben-Menahem believes that individual events are neutral in the sense that there is nothing in them that dictates that they should be described as stop signs. Therefore, she must opt for option II of the constructive theory, and *add* descriptors to the ontology of events.

7. Conclusion: why the dilemma is destructive

We have shown in this paper that Davidson's principle theory of anomalous monism faces a dilemma with respect to the constructive theory of descriptions that underlies his principle approach. It seems to us that Options I and II for constructive theories of description are the only coherent ways to explain the way that descriptions come about and how they produce physical kinds and mental kinds in Davidson's approach. In short, a constructive theory underlying Davidson's approach should belong to either Option I or Option II.

However, it seems to us that this dilemma is destructive for Davidson's anomalous monism for the following reasons.

As we have shown only Option I might be monistic (since Option II collapses to a dualistic theory), precisely because the partition of the events into sets is fixed completely by the intrinsic spatio-temporal features of the events and their causal relations. In this theory,

descriptions capture the spatio-temporal structure of events and do not induce in any interesting sense the partition of the events into sets, quite unlike Option II, in which description induce the partitions to sets, but this has the consequence that the structure of events entails substance-token-dualism, not monism. As we have explained, on Option I, however, there is a threat that anomalous monism collapses into a type-identity *physicalist* theory, because the spatio-temporal structure and causal relations seem to be exactly what physics is about. But from our construction of Davidson's anomalous monism one can see that if Option I does not collapse to type identity physicalism, then Davidson's conjecture about the relationship between causation and strict physical laws is not general and may or may not be true of our world. Moreover, we have also indicated that the freedom that Davidson achieves in his model can be accounted for also within a *type-identity reductive physicalist* theory in which multiple realizability is only apparent.

The two options for a constructive theory of descriptions *mirror*, or are *analogous* to, the approaches of physicalism vs. mind-body dualism in theories where the fundamental layer of events is physical (and not of events that are neutral in this sense). Davidson may have attempted to build something that is *neither* a physical type-identity theory *nor* a substance dualistic theory. His attempt reflects the then-contemporary attempts that resulted in the idea of so-called *non-reductive* physicalism (including functionalism of all forms) that allows for multiple realization of special sciences kinds by physical kinds. These attempts have failed: multiple realization is either a form of physical-type-identity – if the tokens are appropriately identified, i.e. expanded as physics requires; or else – if multiple realizability is genuine, it is a form of mind-body dualism (see Hemmo and Shenker 2019a, 2019b; Shenker 2017c). The fate of Davidson's event ontology on Option II (discussed above) is the same, and there is a serious threat that it is the same also on Option I.

Acknowledgment

Some of the ideas in this paper have been presented in a conference in honor of Yemima Ben-Menahem's work. We are grateful to Yemima for her insightful comments, and for discussing these matters with us for years. This research is supported by the Israel Science Foundation, grant number 1148/2018.

References

- Albert, D. (2000) *Time and Chance*, Cambridge, Mass.: Harvard University Press.
- Anscombe, G.E.M. (1957) *Intention*. Harvard University Press.
- Atkinson, D. (2006) "Does Quantum Electrodynamics Have an Arrow of Time?" *Studies in History and Philosophy of Modern Physics*, 37 (3), 528-54.
- Ben-Menahem, Y. (1997) "Historical Contingency," *Ratio* 10/2:99–107.
- Ben-Menahem, Y. (2001) "Direction and Description," *Studies in the History and Philosophy of Modern Physics* 32:621–635.
- Ben-Menahem, Y. (2018) *Causation in Science*, Princeton University Press.
- Brown, H.: Physical Relativity (2005) *Spacetime Structure from a Dynamical Perspective*, Oxford, Oxford University Press.
- Callender, C. (2001) "Taking Thermodynamics Too Seriously," *Studies in the History and Philosophy of Modern Physics* 32:539–553.
- Chang, H. (2012) *Is Water H2O?* Dordrecht: Springer.

- Child, W. (1993) "Anomalism, Uncodifiability, and Psychophysical Relations," *The Philosophical Review*, Vol. 102, No. 2, 215-245.
- Davidson, D. (1963) "Actions, Reasons and Causes", *Journal of Philosophy*, 60, 685–700; reprinted in D. Davidson, *Essays on Actions and Events*, pp. 12–25 University of California, Berkeley.
- Davidson, D. (1967) "Causal Relations" in Donald Davidson, (2001) *Essays on Actions and Events*, 126–137, University of California Berkeley (2nd Edition).
- Davidson, D. (1970) "Mental Events," in Donald Davidson, D. (2001) *Essays on Actions and Events*, 170–187, University of California, Berkeley.
- Davidson, D. (1973) "The Material Mind", in Donald Davidson (2001) *Essays on Actions and Events*, pp. 200–211, University of California, Berkeley.
- Davidson, D. (1993) "Thinking Causes," in John Heil and Alfred R. Mele (eds.), *Mental Causation*, pp. 3-18, Oxford University Press.
- Drory, A. (2014) "The Necessity of the Second Postulate in Special Relativity," *Studies in the History and Philosophy of Modern Physics*. Printed version forthcoming, available online doi:10.1016/j.shpsb.2014.08.015 and <http://arxiv.org/pdf/1412.4018.pdf>
- Einstein, A. (1905a) "On the Electrodynamics of Moving Bodies," *Annalen der Physik* 1905. English translation in Einstein, *The Principle of Relativity*. London: Methuen, 1923.
- Einstein, A. (1905b) "On the Movement of Small Particles Suspended in a Stationary Liquid Demanded by the Molecular-Kinetic Theory of Heat," *Annalen der Physik* 1905. English translation in Einstein, *Investigation on the Theory of the Brownian Movement*. London: Dover, 1956.
- Einstein, A. (1919) "Time, Space, and Gravitation," *Times* London, 28 November 1919, 13–14. Reprinted as "What is the Theory of Relativity?" In Einstein, *Ideas and Opinions*, 227–232. New York: Bonanza Books, 1954.
- Einstein, A. (1970) "Autobiographical Notes," in P. A. Schilpp, ed., *Albert Einstein: Philosopher-Scientist*, vol. 2. Cambridge: Cambridge University Press.
- Fodor, J. (1974) "Special Sciences, or: The Disunity of Science as a Working Hypothesis," *Synthese* 28, 97–115.
- Fodor, J. (1997) "Special Sciences: Still Autonomous After All These Years," *Nous* 31, 149-163.
- Frigg, R. (2008) "A Field Guide to Recent Work on the Foundations of Statistical Mechanics," in D. Rickles (ed), *The Ashgate Companion to Contemporary Philosophy of Physics*, pp. 99–196, London: Ashgate.
- Goldstein, S. (2012) "Typicality and Notions of Probability in Physics." In Y. Ben-Menahem and M. Hemmo, *Probability in Physics*, 59–72. Berlin: Springer.
- Hagar, A. and Hemmo, M. (2013) "The Primacy of Geometry," *Studies in History and Philosophy of Modern Physics* 44, 357-364.
- Hemmo, M. and Shenker, S. (2011) "Introduction to the Philosophy of Statistical Mechanics: Can Probability Explain the Arrow of Time in the Second Law of Thermodynamics?" *Philosophy Compass* 69, 640–651.

Hemmo & Shenker, Davidson

- Hemmo, M. and Shenker, S. (2012) *The Road to Maxwell's Demon*, Cambridge, Cambridge University Press.
- Hemmo, M. and Shenker, S. (2013) "Entropy and Computation: The Landauer-Bennett Thesis Reexamined," *Entropy* 15, 3387–3401.
- Hemmo, M. and Shenker, S. (2015a) "Probability and Typicality in Deterministic Physics", *Erkenntnis*, 80 (3), 575–86.
- Hemmo, M. and Shenker, O. (2015b) "The Emergence of Macroscopic Regularity," *Mind and Society*, 14 (2), 221-244.
- Hemmo, M. and Shenker, O. (2016) "Maxwell's Demon," Oxford Handbooks Online, Oxford University Press.
<http://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199935314.001.001/oxfordhb-9780199935314-e-63?rskkey=plUI7T&result=1>
- Hemmo, M. and Shenker, O. (2017) "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, Ma: Harvard University Press, forthcoming.
- Hemmo, M. and Shenker, O. (2019a) "On Why Functionalism is a Form of 'Token-Dualism in Disguise,'" forthcoming.
- Hemmo, M. and Shenker, O. (2019b) "Two Kinds of High-Level Probability," *The Monist* 102, 458–477.
- Hemmo, M. and Shenker, O. (2019c) "The Past Hypothesis and the Psychological Arrow of Time," *British Journal for the Philosophy of Science*, axz038, <https://doi.org/10.1093/bjps/axz038>
- Howard, A.D. (2010) "Einstein's Philosophy of Science." The Stanford Encyclopedia of Philosophy. Summer 2010 Edition. Edward N. Zalta, ed. URL = <http://plato.stanford.edu/archives/sum2010/entries/einstein-philsience/>, Section 6.
- Kim, J. (2012) "The Very Idea of Token Physicalism" in Gozzano Simone and Hill Christopher (eds.) *New Perspectives on Type Identity*, Cambridge University Press, pp. 167-185.
- Ladyman, J., and Ross, D. (2007) *Every Thing Must Go*, Oxford: Oxford University Press (with D. Spurrett and J. Collier).
- Ludwig, Kirk (2003) (ed.) *Donald Davidson*, Cambridge University Press.
- Mermin, N.D (2009) *It's about time*. Princeton University Press.
- Ney, A. (2008) "Physicalism as an Attitude," *Philosophical Studies* 138, 1-15.
- Orilia, F. and Swoyer, C. (2016) "Properties." The Stanford Encyclopedia of Philosophy (Winter 2016 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2016/entries/properties/>.
- Polger, T. and Shapiro, L. (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>

- Schilpp, P. A. (1949) *Albert Einstein, Philosopher–Scientist*, New York: Open Court.
- Shagrir, O. (2011) “Supervenience and Anomalism are Compatible,” *Dialectica* 65, 241-266.
- Shenker, O. (2015). “Davidsonian Descriptions as a Principle Theory,” *Iyyun: The Jerusalem Philosophical Quarterly* 64, 171-190. In Hebrew
- Shenker, O. (2017a) “Foundations of Statistical Mechanics: Mechanics by Itself,” *Philosophy Compass* 12 (12), <https://doi.org/10.1111/phc3.12465>
- Shenker, O. (2017b) “Foundations of Statistical Mechanics: The Auxiliary Hypotheses,” *Philosophy Compass*, 12 (12) <https://doi.org/10.1111/phc3.12464>
- Shenker, O. (2017c) “Flat Physicalism: Some Implications,” *Iyyun: The Jerusalem Philosophical Quarterly* 66, 1–15.
- Shenker, O. (2018) “Foundations of Quantum Statistical Mechanics,” in Eleanor Knox and Alastair Wilson (eds.), *Routledge Companion to the Philosophy of Physics*, Oxford: Routledge, forthcoming.
- Sklar, L. (1993) *Physics and Chance*, Cambridge: Cambridge University Press.
- Steiner, M. (1983) “Under a Description,” in L. Cauman *et al.* (Eds.), *How Many Questions? Essays in Honor of Sidney Morgenbesser*, 120–131. Hackett.
- Uffink, J. (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.