The Second Law of Thermodynamics and the Psychological Arrow of Time

Meir Hemmo and Orly Shenker

Abstract

Can the second law of thermodynamics explain our mental experience of the direction of time? According to an influential approach, the past hypothesis of universal low entropy (required to explain the temporal directionality of the second law in terms of fundamental physics, which is time-symmetric) also explains how the psychological arrow comes about. We argue that although this approach has many attractive features, it cannot explain the psychological arrow after all. In particular, we show that the past hypothesis is neither necessary nor sufficient to explain the psychological arrow on the basis of current physics. We propose two necessary conditions on the workings of the brain that any account of the psychological arrow of time must satisfy. And we propose a new reductive physical account of the psychological arrow of time compatible with time-symmetric physics, according to which these two conditions are also sufficient. Our proposal has some radical implications, for example, that the psychological arrow of time is fundamental, whereas the temporal direction of entropy increase in the second law of thermodynamics and the past hypothesis is derived from it, rather than the other way around.

1 Introduction

> Our observations of physical things, our feelings and emotions, and our thinking processes extend through time and cannot escape the steady current that flows unhaltingly from the past by way of the present to the future. (Reichenbach [1956], p. 1)

Reichenbach aptly begins his book on the physical arrow of time by focusing on our experience, that is, our mental life. He draws attention to two (related) features of our temporal experience: flow (or passage) and direction. In this paper we focus only on the directionality per se of temporal experience. By this directionality we mean the fact that we have an experience of a temporal direction (that accompanies, for example, our experience of 'passage' (or 'flow') of time.[1] The explanation we seek for this fact is physical. The question we address is this: how is the psychological fact of a temporal direction to be accounted for by contemporary physics, which is time-symmetric?

Our working hypothesis is the standard one in this context, namely, that the mental is to be accounted for by the physical.[2] Some authors (for example,

[1] For other aspects of temporal experience (for example, passage, present, order, duration), see (Ismael [2010], [2017]; Maudlin [2005], Chapter 4; Dieks [2016]); for overviews, see (Le Poidevin [2015].

[2] We don't address here Hempel's dilemma concerning physicalism; see (Ney [2008]).

Albert ([2000], [2014]; Loewer [2012]; Smith [2014])[3] argue that the physical account of the psychological arrow is grounded in the entropic arrow, described by the second law of thermodynamics, including the past hypothesis, which is part of the statistical mechanical account of this law And since fundamental physics is time-symmetric,[4] and the second law is seen as the only (effective) temporal asymmetry that comes into play at levels of energy relevant for the brain,[5] it seems to many that there is no other option, but to explain the psychological arrow by the entropic arrow. But we show that: (a) this standard account doesn't work; and (b) there is a better way to explain the psychological arrow within time-symmetric fundamental physics.

Our approach is motivated by the seemingly innocent working hypothesis—which is fruitful and enjoys vast empirical support—that our mental experience is to be explained (somehow, by and large) by features of our brains, perhaps together with some other parts of the body; for simplicity we call all of them 'the brain'. But contemporary science (brain and cognitive science) is far from explaining how mental experience comes about. It is, therefore, extremely interesting that despite this lack of knowledge, there is a particular case—

[3] For various versions of the past hypothesis of low entropy, see (Sklar [1973]; Albert [2000], [2014]; Loewer [2001], [2012]; Callender [2010]; Wallace [forthcoming]; Hemmo and Shenker [2012]). Price [1996] argues that past hypothesis implicitly presupposes time-asymmetry. Earman ([2006], p. 410) criticises Feynman's dismissal of Boltzmann's remarks on this problem; see also (Goldstein, Tumulka and Zanghi [2016]).

[4] The meaning of time-reversal invariance is under debate; see (Allori [2015]). Asymmetries like the charge-parity asymmetry in quantum electrodynamics and classical electromagnetic radiation require special conditions that are compatible with the time-symmetry of the fundamental laws, but they are not explained by the laws.

[5] The charge-parity asymmetry in quantum field theory is considered irrelevant for the workings of the brain due to the high levels of energy in which it occurs; but see Section 6.

namely, the experience of temporal directionality—in which it is possible to come up with some minimal necessary conditions for the physical structure that gives rise to this experience. We shall point out these physical conditions.

We will show that the requirement that the psychological arrow is to be accounted for by contemporary fundamental physics entails that the locus of the psychological arrow is in a physical symmetry-breaking in the brain (in this sense we shall say that the symmetry-breaking is local), in the same physical degrees of freedom that are also responsible for giving rise to the experience of a temporal arrow. We call these two features 'asymmetry' and 'mental' (see Section 3). As we shall see, these conditions are rather weak and general, and hopefully un-controversial; the thing is to put them on the table and consider their implications; and those turn out to be non-trivial.

We shall argue that these two necessary conditions are also sufficient for a physical account of the psychological arrow. By this we don't mean that we have all the details of this account; far from it, of course, given the little that science knows at the moment concerning the physics of the mental. Our claim is, rather, that the asymmetry and mental conditions form the general structure of the matter, and the (vastly many) details that will be discovered in later research will fit in this framework, which is very different from the alternative frameworks in the literature based on the entropic arrow.

The implications of our new proposal are quite radical, but seem unavoidable. Here is briefly, an example: Suppose (on the one hand) that according to fundamental physics the world is symmetric under the reversal of the direction of time; and (on the other hand) that the psychological arrow is a feature of some asymmetric local degrees of freedom in the brain. If this is true, it may be that this local symmetry-breaking is projected by us onto the rest of the world, making the external world appear to us to be temporally directed. This result settles the tension between the temporal symmetry of fundamental physics and the temporal asymmetry of our experience. The interesting implication of this proposal is that the local-physical-psychological asymmetry determines the direction of the thermodynamic arrow given by the past hypothesis), and not (as is often believed) the other way around. We shall examine how our proposal and this consequence come about in some detail as we proceed.

Finally, our arguments don't depend on any specific answer to the mind-body question; they apply to any view in which the physical-to-mental correlations satisfy supervenience (logical or nomological) of (kinds of) mental states and processes on (kinds of) physical states and processes. In what follows we use a reductive terminology, but this is only for convenience.

And so, this paper has four central aims: we propose two necessary conditions for explaining the psychological arrow of time by fundamental physics; we argue that influential accounts of the psychological arrow of time based on the second law of thermodynamics and the past hypothesis are wanting; we propose a new alternative approach to the psychological arrow with the surprising consequence that the direction of entropy increase is fixed by the psychological arrow (rather than the other way around); and we explain why our approach leads to a research programme very different from the one implied by current approaches to the psychological arrow.

The paper is structured as follows: We start (Section 2) by characterizing some central features of Boltzmann's ([1964]) proposal for explaining the psychological arrow by the (local) entropic arrow. We introduce (Section 3) the two necessary conditions that a physical explanation of the psychological arrow should satisfy. We present (Section 4) what we see as the best way of understanding the prevalent attempts to explaining the psychological arrow of time on the basis of the second law of thermodynamics and the past hypothesis. We then argue that these attempts fail, since the past hypothesis of the low entropy of the universe is neither sufficient (Section 5) nor necessary (Section 6) for explaining the psychological arrow. Finally, in Section 7 we propose a new reductive approach to the psychological arrow and we examine some of its implications (for example, with respect to the foundations of statistical mechanics and brain research).

2 A Physical Account of the Psychological Arrow of Time

To see (somewhat intuitively) what a physical account of the psychological arrow of time amounts to, we shall begin by focusing on a famous proposal by Boltzmann ([1964]) according to which the psychological arrow of time is to be explained by the arrow of entropy increase. Boltzmann writes:

> […] in the universe, which is in thermal equilibrium throughout and therefore dead, there will occur here and there relatively small regions of the same size as our galaxy (we call them single worlds) which, during the relatively short time of eons, fluctuate noticeably from thermal equilibrium, and indeed the state probability in such cases will be equally likely to increase or decrease. For the universe, the two directions of time are indistinguishable, just as in space there is no up and down. However, just as at a particular place on the earth's surface we

> call 'down' the direction toward the center of the earth, so will a living being in a particular time interval of such a single world distinguish the direction of time toward the less probable state from the opposite direction (the former toward the past, the latter toward the future). (Boltzmann [1964], pp. 446–7)

Boltzmann addresses 'the two directions of time' with an analogy between the temporal arrow and the gravitational arrow; see (Sklar [1981]). The 'up–down' distinction refers to the local gradient of the gravitational field near the surface of Earth, and Boltzmann proposes understanding the 'past-future' distinction as referring to (what we might call) the local gradient of an 'entropy field' described by the second law of thermodynamics.[6]

Boltzmann's approach concerning the arrow of time is reductive, in the sense that the temporal asymmetry is explained in terms that are not temporal. It is crucial to realise that entropy is intrinsically non-temporal and it is not temporally oriented; the gradient of entropy (global or local) is not intrinsically correlated with a direction of time (with invariably the same side side of the sequence of states) more than, say, the distribution of particles in space is. For example, in (Boltzmannian) statistical mechanics the (relative) entropy of a system is the Lebesgue measure of a state space region, and in thermodynamics it is a function of properties like pressure and temperature: these concepts are distinct from the concepts of time or of temporal direction. If they were inherently temporal, the second law would be analytically true; but it is well

[6] The idea of a temporally orientable spacetime in which the direction of time is defined at every spacetime point is consistent with general relativity; see (Earman [1974]; Weingard [1977]).

known that fundamental mechanics is compatible with possible-worlds in which the second law is not satisfied and in which entropy 'zig-zags' in any imaginable way. The reductive account of the temporal arrow is, on the view we follow here, a non-trivial scientific discovery much like the discovery that water is H_2O. It is not *a priori* true, and calls for empirical support and proofs in much the same way.

However, while the local gravitational field is measured by the vestibular system, which then sends electrochemical signals to the brain, leading to the 'up–down' feeling,[7] Boltzmann doesn't worry about the way the local 'entropy field' can be measured or sensed by us to bring about our feeling of temporal directionality, that is, the psychological arrow of time. In this sense there is a lacuna in Boltzmann's approach to the psychological arrow of time that is therefore only partial.

Contemporary approaches addressing this question (for example, (Hawking [1988][8]; Albert [2000], [2014]; Loewer [2012]; Mlodinov and Brun [2013]; Smith [2014]) attempt to complete this missing link. Smith [(2014)], for example, appeals to processes of diffusion in memory formation and gene expression, which are locally entropy-increasing. He conjectures that these processes may be responsible for the psychological arrow of time, regardless of whether they are correlated with the entropy gradient in the rest of the universe.

There are two aspects of Smith's argument that are important for the

[7] We don't address the 'hard problem' (Chalmers [1996]).

[8] Hawking's ([1988]) explanation of the psychological arrow on the basis of thermodynamics of memory systems is flawed; it relies on mistaken ideas about the relation between entropy and information; see (Hemmo and Shenker [2012], [2013]).

account of the psychological arrow of time that we are seeking. The first is that Smith relies on generalizations of empirical findings concerning temporally asymmetric phenomena of dissipation (local entropy-increasing diffusion processes or gene expression) in the brain, and not on statistical-mechanical proofs of such phenomena (nor, more generally, on proofs of the second law of thermodynamics) from first principles of the low-level mechanical underpinning of such phenomena. The advantage here is that it is not exposed to various questions at the foundations of statistical mechanics concerning the mechanical origin of the time a-symmetry of the second law and the past hypothesis. However, the disadvantage is that it is not rooted in fundamental physics. The main challenge in providing a physical account of the psychological arrow is the fact that fundamental physics is time-symmetric,[9] and assuming that the high-level time-asymmetry of the second law is relevant here sidesteps this challenge.

The second point concerning Smith's approach is that although the experience of a temporal arrow might be explained by the entropy gradient in diffusion processes involved in memory formation, other relevant brain processes may be entropy-decreasing as the brain (which is an open system) becomes more complex with time, and they too could in principle account for the psychological arrow. It is thus a question of fact which of these kinds of processes actually gives rise to the experience of a temporal arrow. Given the current state of the art in brain science, Smith's attempt to settle this question

[9] By contrast, the fundamental asymmetry of the collapse of the wavefunction in quantum mechanics (as in the GRW theory) might explain the temporal arrow (see Section 7).

may be pre-mature.[10]

3 Necessary Conditions for the Psychological Arrow of Time

Following Boltzmann's proposal (see above), our first observation concerning the physical explanation of the arrow of time is the need to make a distinction between what happens in the brain (and is 'local' in this sense) and what happens in the environment (or the entire universe), and then make the suitable connections between them. Let us begin with some remarks concerning the brain part of the explanation, and then we shall address possible connections between the states of affairs in the brain and in the environment regarding the experience of temporal directionality.

To give rise to the psychological arrow of time, the relevant local degrees of freedom in the brain must satisfy two conditions, which we call 'mental' and 'asymmetry'. Let us now present and illustrate these two conditions with examples from the recent literature:

> Mental: Some degrees of freedom in the brain give rise to the mental experience of a temporal arrow.[11]

[10] We thank Ronen Golan for discussions on this issue.

[11] Contemporary physics describes the world by a continuous sequence of instantaneous states; see (Arntzenius [2000]). It seems to us that the 'specious present' ('integration window') proposed by Dennett and Kinsbourne ([1992]) to explain the temporal order in which we experience external stimuli is perfectly compatible with this description.

In a reductive approach, they should be fully described in terms of fundamental physics (physical laws and boundary conditions) and shown to be correlated with the experience of a temporal arrow. Presumably, not every physical system has experiences, but only systems with certain kinds of physical features: Human beings are such systems, and there might be others provided they have the right sort of physical features.

Here is a case in which the mental condition is not satisfied. Albert ([2014], p. 66) considers a pendulum clock and takes it to be a 'device which has no other business than distinguishing between what it has just done and what it is to do next—the paradigmatic distinguisher, the distinguisher *par excellence*, between what it has just done and what it is to do next'. Presumably pendulum clocks don't have any experience whatsoever, let alone the experience of temporal direction, and therefore they don't satisfy the mental condition.[12] Whether or not we have a mental experience of anything, and in particular of temporal direction, upon observing a pendulum clock depends on the physical features of our brains, not the features of the pendulum.

Obviously, searching for the kind of brain states that give rise to our experience, including the aspects involving the psychological arrow of time, is the subject matter of brain research. Contemporary science has no physical explanation of how mental states come about. Nevertheless, something can be said about the physical explanation of the particular mental experience that is the psychological arrow of time, and this is our second condition:

[12] A physical realization of a Turing machine is no different in this respect.

Asymmetry: There must be some physical symmetry-breaking in the degrees of freedom that instantiate the mental condition.

The asymmetry requirement is necessary because there cannot be any symmetry-breaking in the mental experience unless there is some symmetry-breaking in the physics, if the mental is to be explained by the physical.[13] This asymmetry should be expressed terms of the fundamental physical laws and boundary conditions. From an epistemological perspective this could mean that, ideally, by looking at the state of the brain at a time *t* and identifying the relevant asymmetry, one could tell on which 'side' of *t* the psychological past is.

4 The Psychological Arrow of Time via the Second Law of Thermodynamics and the Past Hypothesis

Essentially, all prevalent approaches to the psychological arrow of time attempt to show that the direction of the psychological arrow of time is aligned with (or even fixed by) the temporal direction of universal entropy increase described by the statistical mechanical counterparts of the second law of thermodynamics. This direction is, in turn, determined by the so-called past hypothesis, which is briefly this. Contemporary fundamental physics is time-symmetric, and therefore a proof that entropy is likely to increase towards the future entails that it is equally likely to increase towards the past. This entails that our memories of past events and other records, described and generalised by the second law of thermodynamics) as well as our belief that the second law holds at all times

[13] See, for example, (van Fraassen [1989]).

(so that entropy increases to the future but decreases to the past), are highly likely to be false.

To solve this problem, Feynman ([1965], p. 116) proposed to ‘add to the physical laws the hypothesis that in the past the universe was more ordered, in the technical sense, than it is today’.[14] Albert ([2000], p. 96) formulates this hypothesis, which he calls the past hypothesis, as follows[15]:

> […] the world first came into being in whatever particular low-entropy highly condensed big-bang sort of macrocondition it is that the normal inferential procedures of cosmology will eventually present to us.

There are two ways to understand the notion of ‘past’ in the past hypothesis:

(i) A non-reductive view: there is a fundamental temporal arrow, which is a fact above and beyond the facts described by contemporary time-symmetric physics, and entropy increases relative to this arrow (see Earman [1974]). A contemporary example of this approach is Maudlin ([2007]; see also discussion in Loewer [2012]). We don’t expand here on this line of thinking.

(ii) A reductive view: the fact that entropy is low only on one side of the sequence of states makes it the case that this side (of each time slice) has the

[14] See, for example, (Carroll and Chen [unpublished]; Carroll [2010]; Barbour *et al*. [2014]. For why relativistic and quantum considerations in cosmology might entail the empirical asymmetry of low entropy (so that a hypothesis to this effect is not needed), see (Goldstein, Tumulka and Zanghi [2016]); see criticism by Winsberg [2004]; Earman [2006]; Price [2010]; Callender [2004].

[15] Albert ([2000], Chapter 4) adds assumptions concerning the dynamics and the probabilistic algorithm of statistical mechanics.

characteristics of 'past' (for example, the states on this side can be known but cannot be changed), while the other side (in which entropy is highly likely to increase) has the characteristics of 'future' (for example, the states on this side are probabilistically predictable or, in a suitable sense, can be influenced).

This reductive approach is endorsed by Albert ([2000], [2014]) and Loewer ([2012]). According to this view the psychological arrow of time is a result of the past hypothesis, such that the following three statements turn out to be true:

(I) The universal entropy increase (which is reflected in the brain) is to be explained, as usual, from first principles (that is, without resorting to our experience of second law behavior), by a combination of (a) a proof that entropy is highly likely to increase towards the future, and (b) a proof of the past hypothesis concerning universal low entropy in the remote past (for example, along the lines of Carroll and Chen [unpublished]; Carroll [2010]).[16]

[16] For overviews of attempted proofs in the classical case, see (Uffink [2007]; Frigg [2008]; Werndl and Frigg ([2015]); in the quantum case, see (Linden *et al*. [2009]; Goldstein, Huse, Lebowitz, and Tumulka [2016]). A universal proof cannot be attained, since a Maxwellian Demon, is provably compatible with the principles of classical statistical mechanics; see (Albert [2000]; Hemmo and Shenker [2010], [2011], [2012], [2013]); for the quantum mechanical case, see (Hemmo and Shenker [forthcoming-a]). There are, however, important special cases of entropy, for example, Lanford's theorem, but these are subject to the minimum problem; see (Uffink and Valente [2010], [2015]).

The next task is to show that the 'mental arrow of time is aligned with the thermodynamic arrow' (Goldstein, Tumulka and Zanghi [2016], p. 5)—that is, that our experience reflects a fact about the external world.[17] (Towards the end of this paper we challenge this task.) Here is how this is supposed to be done by the past hypothesis, according to Loewer and Albert:

(II) The local entropy gradient in the brain is a consequence of the universal entropy increase induced by the past hypothesis. We shall consider two ways of establishing this claim that seem to appear in (Albert [2000], [2014]; Loewer [2012]).

(II) The psychological arrow is (reductively) explained by a local entropy gradient in the brain, thus satisfying the mental and asymmetry conditions, as for example in (Smith [2014; Mlodinov and Brun ([2014]).

This picture, however, doesn't work, for reasons we now describe, concerning points (II) and (III).

5 Why the Past Hypothesis is Insufficient for the Psychological Arrow of Time

Accounts of the psychological arrow of time that appeal to the past hypothesis have two stages that together should establish that the mental arrow of time is aligned with, or in fact determined by, the thermodynamic arrow:

[17] Whether the past hypothesis prevents causal influences on the past is an open question; see (Albert [2000], [2014]; Elga [2000]; Frisch [2010]).

> Stage 1: Connects the hypothesis of low entropy in the past of the universe, that (together with the proof of entropy increase to the future) brings about a universal entropy gradient, with a local entropy gradient (or some other appropriate local asymmetry) in the present in our brain.

> Stage 2: Connects the local entropy gradient in our brain at present with the psychological arrow of time at present.

Notice that our necessary conditions, asymmetry and mental, are relevant only for Stage 2. What, then, is the significance of Stage 1? The answer is this: If we identify the global entropy gradient with the notion of 'arrow of time', then Stage 1 ensures that the psychological experience, brought about by Stage 2, reflects this arrow, and in this sense provides us with information about a feature of the external world. Stage 2 without Stage 1 would, on this view, provide an internal experience that could be some kind of illusion of hallucination.[18] In Section 7 below we present a new physical account of the psychological arrow, which focuses on Stage 2, and provides a reductive substitute for Stage 1. But first, in this section, we describe the prevalent approaches, that consist of combining Stages 1 and 2.

In the literature there are two approaches regarding Stage 1 that are not always separated from each other. One approach attempts to connect directly

[18] Paul ([2015]) addresses the question of whether the psychological arrow can be illusory in some sense. We expand a bit on this issue in Section 7.

the hypothesis of low entropy in the past with the local gradient of entropy in the brain by typicality arguments. The second approach attempts to connect the hypothesis of low entropy in the past with the local gradient of entropy in the brain (or perhaps with some other asymmetry in the brain) by dynamical or causal considerations. Let us now show that in both Stage 1 approaches the past hypothesis is insufficient (even with high probability) to entail entropy gradient in the brain or any other asymmetry that might account for the psychological arrow of time. We shall start with the first approach via typicality, and we shall later show that also the second approach via a dynamical interaction relies indirectly on typicality arguments.

5.1 Stage 1, Version 1: From the past hypothesis to the psychological arrow via typicality

Recall that the task here is to satisfy the requirement in statement (II) of the previous section. In Version 1 the idea is to formulate a direct typicality argument for deriving the entropy gradient in (some degree of freedom) in the brain from the past hypothesis. An example of this strategy might be (Loewer [2012]; Albert [2014]), which is roughly this: The initial conditions of the universe, which are compatible with the past hypothesis, form two sets. One set gives rise to trajectories in which the entropy of the universe increases, effectively and to a good approximation in accordance with the second law, while the other gives rise to all sorts of other trajectories. The typicality assumption (which we don't dispute here[19]) is that according to 'standard

[19] See criticism in (Hemmo and Shenker [2012], [2014]).

Boltzmannian arguments' (Albert [2014], p. 65), the former set is much larger (by some appropriate measure). Its trajectories are, in this sense, 'typical'.

This set of typical entropy-increasing trajectories, in turn, has two subsets. In the first subset most subsystems are by and large entropy-increasing, and in the second this is not the case (for instance, in the second subset entropy may increase substantially in only a few subsystems and yet, due to the additivity of entropy, the global entropy of the universe would increase). The first subset, in which most subsystems are entropy-increasing (so the argument goes), is much larger (by some appropriate measure). In this sense its members are 'typical'.

There are two levels of typicality considerations involved here: globally entropy-increasing universes are typical and, amongst them, universes in which most subsystems are entropy increasing are typical. If our universe belongs to the typical set and typical subset, then this entails the empirically verified prediction that subsystems (such degrees of freedom in our brains) are entropy-increasing. So, by this double typicality argument, the past hypothesis is a sufficient condition for a local psychological arrow of time.

The question is, however, why should one accept the latter 'if' clause, that is that our universe is typical in the two (abovementioned) levels. Two reasons are presented in the literature. The first one is this: (1) The spatiotemporal part of the universe that we observe obeys the second law of thermodynamics; this statement has ample empirical support. (2) There is no reason to think that the spatiotemporal part of the universe that we observe is special, that is, atypical (see debate on this matter between Feynman ([1965]) and Earman ([2006])); hence we can assume that it is typical, in the two levels. (3) Therefore, the rest of the universe that we observe is highly likely to be the same as the part of the universe that we observe. (4) Therefore, our universe and most of its subsystems obey the second law of thermodynamics. It seems to us that the proposal by Loewer ([2012]) can be understood along these lines. (5) It is tempting to continue by saying that statement (4) explains why our observed environment also obeys the second law, but this would be circular, given the above argument. We started off from the actual case that we observe and inferred that typical cases are like it; and therefore, since the actual dictates the typical, the typical cannot explain the actual.

The second reason for accepting the 'if' clause (of Stage 1 of the typicality argument (which states that our universe belongs to the typical set and its typical subset) is that the measure used to determine 'typicality' in the above two senses is dictated by the dynamics of the universe (see, for example, Dürr [2001]; Goldstein [2012]). We have argued elsewhere (Hemmo and Shenker [2014]) that none of the dynamical arguments presented in the literature for the choice of the measure of typicality are compelling.

Perhaps Loewer ([2012], p. 125) acknowledges the weakness of typicality arguments (in both lines of thinking) when he writes: 'But this is as should be. The job is to get the second law from the Mentaculus *in so far as* the second law is correct'. Our point is that to complete this version of Stage 1, one must accept the typicality argument in at least one of its versions sketched above. Suppose that one could show by direct dynamical considerations (regardless of typicality) that the universe evolved from its low entropy state at the time of the big bang (stipulated by the past hypothesis) in accordance with the second law of thermodynamics, and (in particular) the universe evolved in such a way that as a matter of fact also subsystems of the universe increase their entropies; see

(Winsberg [2004]; Frisch [2010]). Here there would be no appeal to counterfactual universes but rather to the dynamics of our actual universe. This would be, indeed, a way to derive the entropy gradient in the brain from first principles of physics. However, the stipulated dynamics does all the work here and the past hypothesis becomes explanatorily redundant. One could equally derive by similar direct dynamical considerations the entropy gradient in the brain in a universe for which the past hypothesis is false. And so, on this reasoning the past hypothesis by itself is neither necessary nor sufficient for generating a local entropy gradient in the brain. So, this dynamical argument entails exactly what we wish to show. (We consider below a different version of a dynamical or causal argument according to which the past hypothesis is essential to the way in which the local entropy gradient in the brain is generated.)

One last point concerning the typicality argument implicit in Albert and Loewer's approach. Suppose that a local entropy gradient in the brain turns out to be responsible for the psychological arrow of time. Then for the explanation of our actual experience of this arrow (which is, after all, a matter of fact) it is totally irrelevant whether or not this actual state of affairs is also likely in some sense. Such likelihood is important only as part of a justification for a belief that this local entropy gradient obtains or will continue to obtain (in the absence of empirical evidence for that effect).

5.2 Stage 1, Version 2: From the past hypothesis to the psychological arrow via dynamical or causal effects

The second approach for explaining how the past hypothesis gives rise to local asymmetry in the brain appeals to some sort of causal considerations. Here (unlike the argument in Version 1) the past hypothesis of low entropy is meant to be essential to the way in which the local entropy gradient in the brain is causally (or dynamically) generated. Albert ([2014], p. 16) suggests that an appropriate correlation between our brain state at the present and the low-entropy past 'will have been hard-wired into us as far back as when we were fish'. Notice that although Albert seems to have in mind a local entropic asymmetry in the brain as recording or reflecting the past low entropy of the universe, it seems to us that this proposal is not committed to the idea that it must be some entropy gradient in the brain that satisfies the asymmetry and mental conditions described above. For all we know, the feature of the brain, that is causally correlated with the entropy gradient in the world induced by the past hypothesis, could be some other local asymmetry in the brain, not necessarily entropic asymmetry. We return to this point in the next section; for now, we focus on the case in which the 'hard wiring' of the past low entropy occurs in the entropy gradient in some degree of freedom in the brain.

What is this 'hard wiring'? The 'hard-wiring' hypothesis assumes that at some point in the remote past, when the hard-wiring developed, there was in us an entropy-gradient sensor that measured the entropy gradient of its environment (which by assumption reflected the entropy gradient of the universe at that time), such that the outcome of this ancient measurement is (possibly following a chain of intermediate stages) registered in precisely those degrees of freedom in the brain (that satisfy the asymmetry condition) that also bring about the psychological temporal experience (as required by the mental condition). What can the entropy-gradient sensor be?

Recall Boltzmann's proposal for explaining the psychological arrow in a way that is analogous to explaining the 'up–down' sensation by referring to the local gravitational field. We know that the vestibular system is the sensor for the external gravitational field. But what could measure the field that leads to the psychological arrow of time? Suppose that the relevant external field is an entropy gradient (and let us assume that this gradient satisfies the second law of thermodynamics). Since entropy is a function of the measure over the state space, that is, a function of a continuous set of micro-states, only one of which is actual (while the rest are counterfactual at every moment), entropy is a number without units, that is a function of the macro-variables (state functions) of the sets of micro-states (given some partition of the state space into sets).[20] Therefore, an entropy sensor should (a) be sensitive to all the macro-variables of which entropy is a function; and (b) carry out the calculation of this function.[21] The fact that we don't have 'entropy meters', despite the importance of this quantity, raises the suspicion that our physiology doesn't contain such a sensor, in which case an entropy gradient (either at present or in the remote past) cannot account for the psychological arrow of time. This is only a suspicion,

[20] See (Hemmo and Shenker [2016], [forthcoming-b]) for the role of macro-variables in statistical mechanics.

[21] There is an open debate on the status of the measure needed to determine the size of sets of micro-states (or regions in phase space). It is not analytic that the function for calculating entropy should be the Lebesgue measure; see for example, (Hemmo and Shenker [2012], Chapter 7) on this point.

however. Perhaps there is an entropy sensor—perhaps nature came up with an entropy gradient sensor that we, at the moment, cannot understand or imitate.[22]

Nevertheless, this suspicion brings to mind the option that there is no sensor: the asymmetry responsible for the arrow of time is generated within the brain, and it need be correlated with an external field outside it. This option has some radical consequences, which we explicitly consider in the next sections. But it also raises another reading of Albert and Loewer's account of how the psychological asymmetry is brought about dynamically by some sort of sensitivity to an already existing external asymmetry in our environment. It goes roughly as follows[23]: Let us suppose (for a moment) that the past hypothesis of a low entropy macro-state, together with the uniform probability distribution over the micro-states in this macro-state, ground the asymmetry of entropy increase, which in turn grounds, for example, both the asymmetry of records and influence. The idea is that these asymmetries play a crucial role in accounting for the psychological arrow in that humans have (presumably) evolved so that their experience keeps track of the temporal order of events in their environment. This is because there can be records (for example, memories) of events only in the states of systems that occur at times that are further away from (the time of the) low entropy state, and we can at the present significantly influence (or have handles of influencing) events only if they occur at times that are further away from the time of the low entropy state; see (Albert [2000], [2014]). How the brain implements this tracking, namely how the brain

[22] We thank Yemima Ben-Menahem for discussions on this point.
[23] We thank an anonymous reviewer for proposing this reading.

becomes correlated with these external asymmetries, is a question for psychology and neuroscience (as we said above), and in particular it need not be implemented in some entropy gradient in the brain. It is enough that these external asymmetries be correlated with some asymmetry in the brain, not necessarily an in-brain asymmetry of entropy. So: also this version of Stage 1 suggests that the thermodynamic arrow, the arrows of records and influence, and the psychological arrow are all aligned, pointing away from the time of the low entropy past assumed by the past hypothesis. And the crucial point is that the role played by the past hypothesis in both the way that the asymmetries of records and influence are brought about, as well as in the way that they are 'tracked' by humans, is crucial in accounting for our psychological arrow of time.

However, this reading of Albert and Loewer's account of the psychological arrow is subject to the same challenges spelled out above. First of all, the claim that humans have evolved such that their experience 'tracks' the external asymmetries of records and influence need be unpacked. If some degrees of freedom in the human brain become correlated with the asymmetry of records exhibited in the states of some local systems in our environment, then there must be a physical account of how precisely this is done: in particular, this idea requires that humans implement some sort of a 'measurement' on the temporally asymmetric physical properties of, for example, the recording systems in their environment, in which case our criticism in this section re-arises: which physical properties are 'tracked' by human's brains such that they become correlated with the asymmetry of records?

Second, on Albert and Loewer's account, the asymmetry of records and influence are brought about by the asymmetry of entropy increase in local systems in our environment, which exhibit these asymmetries; while the past hypothesis of low entropy is crucial in bringing about the entropic asymmetry in these local subsystems. But, by exactly the same argument spelled out in Section 5.1 with respect to the in-brain entropy gradient, the past hypothesis of low entropy (together with the uniform probability distribution over this macro-state) is insufficient to derive local entropy increase for any subsystem of the universe, not only the brain. Here, in the current reading of Albert and Loewer's account (Version 2 of Stage 1), as in the previous reading (Version 1 of Stage 1) one implicitly relies on the double-typicality assumption. And so: this current reading already fails in its first step in showing how the local asymmetries of records and influence are brought about by the universal past hypothesis of low entropy.

The third problem for this reading of Albert and Loewer arises with respect to the role of the asymmetry of records, which, indeed seems to us faithful to Albert and Loewer's account of the psychological arrow. However, the asymmetry of records depends not only on the hypothesis of low entropy in the past but crucially on anther independent hypothesis that the macro-state of low entropy of the universe is also what Albert ([2014], p. 38) calls 'the mother (as it were) of all ready conditions'. However, this additional hypothesis, which is crucial for establishing the asymmetry of records, is highly implausible.[24] It

[24] The role of this 'ready-state' hypothesis is to make it the case that our memories (and other records) in the present of events at intermediate times between the present and the low entropy past are probably reliable.

requires not only that we shall be prepared to assume without appealing to any records whatsoever (on pain of indefinite regress) that the low entropy state of the universe is literally a 'ready-to-measure' state, but also and crucially, that our inferences concerning past events by means of local records in the present depend on our having detailed knowledge of the past macro-state cited by the past hypothesis that go much beyond what contemporary cosmology presents to us. And the trouble is that without knowing these additional details of the past state of the universe, we can never be in a position to discern which features of the states of the record systems in our environment in the present are correlated with past events and which features are not. Since this point strongly undermines Albert's ([2000]) claim that the asymmetry of records is a consequence of the asymmetry of entropy, it also undermines the claim that our psychological arrow is a consequence of the past hypothesis via our sensing the asymmetry of records.

5.3 Stage 2: From entropy gradient in the brain to the psychological arrow of time

Once Stage 1 is completed and we have an account of how the past hypothesis brings about an entropy gradient in the brain, the task is to proceed to Stage 2, and connect the local entropy gradient in our brain with the psychological arrow. This task is to be completed by arguments such as those of Mlodinov and Brun ([2014]; Smith [2014]), who attempt to satisfy the asymmetry and mental conditions (implicitly, of course, as they don't discuss these conditions explicitly).

It is worth repeating here that there are various entropy gradients in the brain: the brain's entropy increases in some sense and decrease in other senses. It is an open system, and it is not even approximately quasi-isolated, so that application of second law considerations are highly irrelevant for the brain. But suppose that, as a matter of fact, some thermodynamic processes; see for example (Smith [2014]) will turn out to be candidates for the physics behind the psychological temporal experience. The second law of thermodynamics, and its statistical mechanical counterparts, are not analytic, that is, the entropic arrow is not analytically an arrow of time. The discovery of a correlation between these arrows is a huge scientific achievement, which may come in one of two ways: (1) If we are given a temporal direction, prove that entropy increases in that direction; this is the standard way in which things are done in statistical mechanics. (2) Show that the entropic arrow gives rise to the fact we call a temporal arrow, for example, a psychological fact concerning our experience. These are two distinct lines of research, and in the present context keeping them apart is crucial, in order to avoid circularity.

The macro-variables described by statistical mechanics, like average kinetic energy (which is the counterpart of temperature of an ideal gas in equilibrium) or coarse-grained distribution of positions (the counterpart of volume) are time-symmetric. This implies that within the set of micro-states that share the same thermodynamic macro-variables, some micro-states sit on trajectories that are entropy-increasing (at t) and others sit on trajectories that are entropy-decreasing (at t). So a sensor that would measure such a set of micro-states (by measuring the macro-variables shared by the micro-states in this set) would have to be sensitive to a symmetric property. But the entropy-gradient sensor would be sensitive (by assumption) to an asymmetric property of trajectories, which is shared by a set of micro-states that is a subset of the set that shares the abovementioned thermodynamic macro-variables. In the approaches we consider here the psychological arrow is brought about by sensing this 'finer-grained' property, and not by sensing the thermodynamic properties that appear in statistical mechanics and thermodynamics. And so, the entropy-gradient sensor, that is crucial for all the approaches based on Stage 1 and Stage 2 described in this section, doesn't sense standard thermodynamic properties, but some other, more fine-grained properties. How exactly this could be done is—by the very nature of these properties—beyond the scope of standard thermodynamics and statistical mechanics. As we shall see in Section 7, our own proposal is not subject to this difficulty.

6 Why the Past Hypothesis Is Unnecessary for the Psychological Arrow of Time

We saw that the past hypothesis is not sufficient for explaining the psychological arrow of time. But is it necessary for explanation the arrow? It would be a substantial discovery of physics if the past hypothesis, which involves an entropic asymmetry in the history of the universe, would turn out to be necessary for generating the psychological arrow of time. However, we shall now argue that the past hypothesis (in either the typicality approach (Section 5.1) or the dynamical-causal approach (Section 5.2)) is not necessary for a physical account of the psychological arrow of time.

To see this notice (to start with) that there are functions of the system's micro-state other than entropy that evolve in a variety of ways; and these other state functions provide other possibilities for relevant local asymmetries in the brain that will satisfy the asymmetry requirement and could also satisfy the mental requirement. As a matter of principle, any effective local symmetry-breaking compatible with the time-symmetry of fundamental physics could satisfy the mental condition. Discovering this asymmetry is a question to be settled empirically by brain research.

But it is crucial to note here that entropy gradient is a function of the micro-state which is as good (for this purpose) as any other function of the brain's micro-state, for example, the distance of a certain molecule from our left ear.[25] Of course, effective local asymmetries, if relevant to the workings of the brain, such as classical electromagnetic radiation asymmetry or even the charge-parity asymmetry from quantum electrodynamics, could—for all we know—give rise (locally) to the psychological experience of an arrow of time. But a universal account of the psychological arrow by such asymmetries will not work for essentially the same reasons (*mutatis mutandis*) we have given in Section 5. Also, it is conjectured that the arrow of time induced by the charge-parity- time symmetry is itself a consequence of the past hypothesis of low entropy (together with the probability distribution over the low entropy macro-state in quantum statistical mechanics); see (Atkinson [2006]). If this is true, our argument in Section 5 that the universal past hypothesis (and the probability distribution) is

[25] Memories and expectations are stored in two distinct brain regions. However, this correlation may be explained in two opposite ways: It could be that an event occurred 'in the past' or is expected to happen 'in the future' makes it the case that it is stored in these regions. But it could also be the other way around: the fact that an event is stored in these regions makes it the case that it is felt to be 'in the past' or 'in the future'.

insufficient to bring about the psychological arrow of time will apply directly to a universal account based on the CPT theorem.

Another local alternative for accounting for the psychological arrow of time might be the fundamental time-asymmetric collapse of the wavefunction in quantum interactions,[26] as in the spontaneous localization theory by Ghirardi, Rimini and Weber (GRW) [1986].[27] Which of these (and other) local effective asymmetries actually gives rise to the psychological arrow of time, is a question of fact, to be settled by brain research (of course once the physical theory is given). This is the main point of our new proposal, to which we now turn.

7 A New Reductive Account of the Psychological Arrow of Time

In this section we present a new reductive physicalist account of the psychological arrow of time, which is not subject to the difficulties faced by the approaches described so far. The schematic framework of our proposal is this: Above we presented two necessary conditions for physically explaining the psychological arrow of time: asymmetry and mental. We shall explain in what sense these two conditions are also sufficient for explaining the psychological arrow. Despite the fact that our proposal (that the asymmetry and mental conditions are sufficient to explain the psychological arrow of time) is highly

[26] If there is a fundamental collapse of the quantum state (as in the GRW theory), then physics is fundamentally not time-reversal invariant; in this case the symmetry-breaking that would explain the psychological arrow need not be merely effective.

[27] von Neumann ([1955], p. 358) argued in reverse, that the second law of thermodynamics grounds the asymmetry of the projection (collapse) postulate in quantum mechanics: 'It is desirable to utilize the thermodynamic method of analysis, because it alone makes it possible for us to understand correctly the difference between [Schrödinger's unitary transformation] and [the measurement transformation], into which reversibility questions obviously enter'.

schematic, and only provides general guideline for where to look for the details on this matter; and despite the fact that all we say is that (once the details of the asymmetry and mental conditions are filled in), we shall argue that nothing more need be sought, in order to provide a full account of the physics of the psychological arrow. As we shall see, our proposal turns out to be radically different from the alternatives, and to have highly non-trivial implications.

It may be best to explore our proposal by comparing it with the views described in the previous sections. While the asymmetry and mental conditions are not presented explicitly in the existing literature, the various proposals we examined above are compatible with them. What they do, in a nutshell, is make further requirements from a physical account of the psychological arrow, and all that we do is reject those further requirements as irrelevant and as unsupported by science (but this tiny step has radical implications). In Section 5 above we distinguished between two kinds of proposals: Stage 1 proposals connect the universal entropy gradient with a local entropy gradient (or some other appropriate local asymmetry) in the brain, and Stage 2 proposals connect the local entropy gradient in our brain with the psychological arrow. We begin with Stage 2 proposals and then address the significance of Stage 1 proposals.

Our proposal differs from the prevalent Stage 2 proposals (see Section 5.3) in that—following the discussion in Section 6—the local degrees of freedom in the brain that satisfy the asymmetry requirement need not be entropy gradient, and moreover need not even be correlated with entropy gradient: they may turn out to have nothing to do, whatsoever, with entropy gradient in the brain (or elsewhere). The feature of the brain that instantiates the asymmetry and mental conditions is a matter of fact, to be discovered by (far into the future) brain research. This is one important aspect of our proposal. Contra to the prevalent habit of thought, entropy gradient in the brain is not the sole option for grounding in physics the psychological temporal directionality.

We stress that entropy gradient has no conceptual nor methodological advantage over other possible local asymmetries in the brain. And brain science should be open to this possibility. Thus, by the way, our proposal is not subject to the difficulties facing the standard Stage 2 proposals, pointed out in Section 5.3 above.

Turning now to Stage 1 proposals (see Sections 5.1 and 5.2): Stage 1 is the attempt to show that the global entropy gradient is responsible for the local entropy gradient in the brain (or for other local asymmetries; we can allow this now), which is—in turn (in Stage 2)—responsible for the psychological experience of temporal directionality. The advantage and significance of this move is that it seems to ensure that the psychological experience reflects the universal entropic arrow, and in this sense provides us with information about a feature of the external world. Stage 2 without Stage 1 would, on this view, provide an internal experience that could be some kind of illusion of hallucination: it would give us a feeling of temporal directionality, while no such fact obtains out there. This is a very strong and important argument for endorsing (or seeking) Stage 1 arguments. Since on our view there is no requirement for Stage 1—because the asymmetry and mental conditions are sufficient for explaining the psychological arrow of time—we need to explain how we can do without it. And what our proposal does, in this context, is reverse the explanatory direction. Here is how.

The local asymmetry in the brain, and the fact that it gives rise to a mental experience of temporal directionality, is the basic fact that we learn from studying the brain (and not from studying other, so-called thermodynamic, external systems). It is this local fact—and only this local fact—that breaks the temporal symmetry. According to fundamental physics, the world is time-symmetric, and indeed there is no 'real' temporal asymmetry 'out there'. And since—as physics tells us—there is no temporal asymmetry 'out there' for us to grasp, our experience of temporal asymmetry is not a case of grasping anything external. Our mental experience of temporal asymmetry indeed doesn't reflect any truth about the world 'out there'. The temporal directionality is—on this reductive view—a fact about our mental experience only, to be accounted for by some physical and local (non-temporal!) asymmetry in the brain. Before we turn to the implications of this view, it is important to stress that—if one takes fundamental physics seriously—then there cannot be any external temporal asymmetry for us to grasp, and the mental experience of temporal asymmetry brought about by non-temporal physical asymmetries in the brain (that may or may not be correlated with other non-temporal asymmetries outside the brain) is all that can be.

One may object to our (Stage 2) reductive approach, in which we follow Boltzmann's analogy between the case of gravity and the case of time in that it is uncontroversial that the brain's degrees of freedom that give rise to the up–down experience need not have anything to do with gravity. But with respect to time, this is under debate; see (Phillips [2017]; Dennett and Kinsbourne [1992]). Our proposal doesn't presuppose any relation of representation. The in-brain asymmetry responsible for the psychological arrow of time may be spatial or any other non-temporal asymmetry, precisely because it doesn't represent anything. Here it is also crucial to realise (as we said) that entropy gradient (local or global) is not intrinsically future-directed, or more generally, it is not temporally oriented at all. The fact that there are mechanical trajectories along which entropy changes in almost anyway one likes (increases, decreases, zigzags) supports this point. If entropy were inherently temporal, proofs of the second law from the underlying mechanical theory (classical or quantum) would be analytic and trivial. We take this to mean that to explain the psychological arrow of time, entropy gradient (local or global) is conceptually on a par with any other spatial asymmetry, so it has no *a priori* status in this respect.

Once we have the symmetry-breaking in the brain, we can explain the rest. Fundamental physics attempts to describe in full the sequence of instantaneous states, but doesn't give it a temporal direction (since the sequence of physical states has no temporal direction in a time-symmetric physics). The direction is a reflection of our psychological arrow, which is explained by the physical asymmetry and mental conditions. We say that the low entropy assumed by the past hypothesis is on the 'side' of the (physically temporally non-directed) sequence of states that we feel to be 'past' due to the local (non-temporal) asymmetry in our brains. And so, it turns out that (on our proposal) the past hypothesis—the fact that the low entropy is 'in the past'—is derivative, and not primitive; it is explained by the psychological arrow, and doesn't explain it. Thus, our way of solving the tension between the temporal symmetry of fundamental physics and the temporal asymmetry of our experience is by reversing the explanatory direction: a local non-temporal asymmetry in the brain is the origin of our experience as-if there is a temporal direction out there; not the other way around.

Thus, Stage 1 views are irrelevant for explaining the psychological arrow, and—since on our view the psychological arrow determines the thermodynamic arrow—Stage 1 views don't explain the latter as well. It turns out that what they take to be the explanandum is—on our view—the explanans, and vice versa: the asymmetrical past hypothesis doesn't explain the psychological arrow, but rather the fact that it is an asymmetric boundary condition is explained by the psychological asymmetry.

Let us conclude with the main implication of our proposal with respect to the origin of all the so-called temporal arrows, in science and elsewhere, which is this. The psychological arrow is the basis for all other temporal arrows (in and outside of physics). In particular, if we consider the direction of entropy increase in the second law of thermodynamics, the 'side' on which we place the past hypothesis is derivative and determined by the degrees of freedom in the brain that satisfy the asymmetry and mental conditions. If it is in fact the case that some local asymmetry gives rise to the psychological arrow, this means that we have a direction of time before postulating the past hypothesis and regardless of its truth. The past hypothesis is placed on the 'side' of the sequence of states that we already experience as 'past', and the explanation of the latter fact may have nothing to do with entropy at all; it may be based on other effective local asymmetries compatible with the time-symmetry of fundamental physics. This means that the psychological arrow of time fixes the temporal direction of the past hypothesis, rather than the other way around. Therefore, if there were no creatures like us with this local asymmetry, the sequence of states of the universe would simply be directionless. Of course, there is no question of reversing the direction of thermodynamic processes, since on our approach there is no such direction in the absence of our asymmetric brains.

Here, however, another question arises, concerning the way in which 'arrows of time' are aligned relative to each other: how can we explain the fact that we place the low entropy state in the psychological past (which is given to us by the psychological arrow of time) rather than in the psychological future? The answer is that the two arrows are (contingently) aligned because this might be evolutionary advantageous. By contrast, on Albert and Lower's approach the two arrows cannot be dis-aligned, because the psychological arrow is derived from the entropic arrow. Therefore, the two approaches don't agree on which facts account for the psychological arrow of time, they don't agree on various possible counterfactuals, and (moreover) they in fact lead to very different research programs. On Albert and Loewer's approach, the psychological asymmetry consists only in the correlations between the universal entropy increase and human's brains, where such correlations may equally exist also in pendulum clocks or in thermometers; in particular, the psychological arrow has nothing to do in particular with brain states that are also mental states. Therefore, it will never be found in some asymmetry intrinsic to human's brains. By contrast, on our approach, the psychological arrow is literally in human heads (and maybe bodies), because: (i) it is grounded in some asymmetry in the brain; and (ii) this asymmetry resides in the same degree of freedom that also grounds the experience of direction. These two conditions make the psychological arrow a subject matter for brain science.

Acknowledgements

We thank David Albert, Heather Demarest and Barry Loewer for comments on issues that come up in this paper. Also, we thank two anonymous reviewers for raising valuable points and questions. This research has been supported by the Israel Science Foundation (grant number 1148/2018).

Meir Hemmo
University of Haifa
Haifa, Israel
meir@research.haifa.ac.il

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

References

Albert, D. [2000]: *Time and Chance*, Cambridge, MA: Harvard University Press.

Albert, D. [2014]: *After Physics*, Cambridge, MA: Harvard University Press.

Allori, V. [2015]: 'Maxwell's Paradox: The Metaphysics of Classical Electrodynamics and Its Time-Reversal Invariance', *Analytica,* **1**, pp. 1–19.

Arntzenius, F. [2000]: 'Are There Really Instantaneous Velocities?', *The Monist*, **83**, pp. 187–208.

Atkinson, D. [2006]: 'Does Quantum Electrodynamics Have an Arrow of Time?', *Studies in History and Philosophy of Modern Physics*, **37**, pp. 528–54.

Barbour, J., Koslowski, T., and Mercati, F. [2014]: 'Identification of a Gravitational Arrow of Time', *Physical Review Letters,* **113**, p.181101.

Boltzmann, L. [1964]: *Lectures on Gas Theory*, Berkeley, CA: University of California Press.

Callender, C. [2004]: 'Measures, Explanation, and the Past: Should "Special" Initial Conditions Be Explained?', *British Journal for the Philosophy of Science,* **55**, pp. 195–217.

Callender, C. [2010]: 'The Past Hypothesis Meets Gravity', in G. Ernst and A. Hüttemann (*eds*), *Time, Chance, and Reduction: Philosophical Aspects of Statistical Mechanics*, Cambridge: Cambridge University Press, pp. 34–58.

Carroll, S. [2010]: *From Eternity to Here*, New York: Dutton.

Carroll, S. and Chen, J. [unpublished]: 'Spontaneous Inflation and the Origin of the Arrow of Time', available at <arxiv.org/abs/hep-th/0410270>.

Dennett, C. D. and Kinsbourne, M. [1992]: 'Time and the Observer: The Where and When of Consciousness in the Brain', *Behavioral and Brain Sciences,* **15**, pp. 183–201.

Dieks, D. [2016]: 'Physical Time and Experienced Time', in Y. Dolev and M. Roubach (*eds*), *Cosmological and Psychological Time*, Chem: Springer, pp. 3–20.

Dürr, D. [2001]: 'Bohmain Mechanics', in J. Bricmont, D. Dürr, M. C. Galavotti, G. Ghirardi, F. Petruccione, N. Zanghi (*eds*), *Chance in Physics: Foundations and Perspectives*, Berlin: Springer, pp. 115–32.

Earman, J. [1974]: 'An Attempt to Add a Little Direction to "The Problem of the Direction of Time"', *Philosophy of Science,* **41**, pp. 15–47.

Earman, J. [2006]: 'The Past Hypothesis: Not Even False', *Studies in History and Philosophy of Modern Physics,* **37**, pp. 399–430.

Elga, A. [2001]: 'Statistical Mechanics and the Asymmetry of Counterfactual Dependence', *Philosophy of Science,* **68**, pp. S313–24.

Feynman, R. [1965]: *The Character of Physical Law*, Cambridge, MA: MIT Press.

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*, London: Ashgate, pp. 99–196.

Frisch, M. [2010]: 'Does a Low-Entropy Constraint Prevent Us from Influencing the Past?', in G. Ernst and A. Hüttermann (*eds*), *Time, Chance,* *and Reduction: Philosophical Aspects of Statistical Mechanics*, Cambridge: Cambridge University Press, pp. 13–33.

Frisch, M. [2011]: ‘From Arbuthnot to Boltzmann: The Past Hypothesis, the Best System, and the Special Sciences’, *Philosophy of Science,* **78**, pp. 1001–11.

Ghirardi, G., Rimini, A. and Weber, T. [1986]: ‘Unified Dynamics for Microscopic and Macroscopic Systems’, *Physical Review D*, **34**, pp. 470–9.

Goldstein, S. [2012]: ‘Typicality and Notions of Probability in Physics’, in Y. Ben-Menahem and M. Hemmo (*eds*), *Probability in Physics*, Berlin: Springer, pp. 59–72.

Goldstein, S., Huse, D. A., Lebowitz, J. L. and Tumulka, R. [unpublished]: ‘Macroscopic and Microscopic Thermal Equilibrium’, available at <arxiv.org/abs/1610.02312v1>.

Goldstein, S., Tumulka, R. and Zanghi, N. [unpublished]: ‘Is the Hypothesis about a Low Entropy Initial State of the Universe Necessary for Explaining the Arrow of Time?’.

Hawking, S. W. [1988]: *A Brief History of Time*, London: Bentam Press.

Hemmo, M. and Shenker, O. [2010]: ‘Maxwell's Demon’, *The Journal of Philosophy*, **107**, pp. 389–411.

Hemmo, M. and Shenker, O. [2011]: ‘Szilard's Perpetuum Mobile’, *Philosophy of Science*, **78**, pp. 264–83.

Hemmo, M. and Shenker, O. [2012]: *The Road to Maxwell’s Demon*, Cambridge: Cambridge University Press.

Hemmo, M. and Shenker, O. [2013]: ‘Entropy and Computation: The Landauer–Bennett Thesis Reexamined’, *Entropy*, **15**, pp. 3387–401.

Hemmo, M. and Shenker, O. [2015]: ‘Probability and Typicality in Deterministic Physics’, *Erkenntnis*, **80**, pp. 575–86.

Hemmo, M. and Shenker, O. [2016]: ‘Maxwell's Demon’, *Oxford Handbooks Online*, available at <doi.org/ 10.1093/oxfordhb/9780199935314.013.63>.

Hemmo, M. and Shenker, O. [forthcoming-a]: ‘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.

Hemmo, M. and Shenker, O. [forthcoming-b]: 'Two Kinds of High Level Probability', *The Monist*.

Ismael, J. [2010]: 'Temporal Experience', in C. Callender (*ed*.), *Oxford Handbook on Time*, Oxford: Oxford University Press.

Ismael, J. [2017]: 'Passage, Flow, and the Logic of Temporal Perspectives', in C. Bouton and P. Huneman (*eds*), *Time of Nature and the Nature of Time*, Chem: Springer, pp. 23–38

Le Poidevin, R. [2015]: 'The Experience and Perception of Time', in E. N. Zalta (*ed*.), *Stanford Encyclopedia of Philosophy*, available at <plato.stanford.edu/archives/sum2015/entries/time-experience/>.

Linden, N., Popescu, S., Short, A. J. and Winter, A. [2009]: 'Quantum Mechanical Evolution towards Thermal Equilibrium', *Physical Review E*, **79**, p. 061103, available at <doi.org/10.1103/PhysRevE.79.061103>.

Loewer, B. [2001]: 'Determinism and Chance', *Studies in History and Philosophy of Modern Physics*, **32**, pp. 609–20.

Loewer, B. [2012]: 'Two Accounts of Laws and Time', *Philosophical Studies,* **160**, pp. 115–37.

Maudlin, T. [2007]: *The Metaphysics within Physics*, Oxford: Oxford University Press.

Mlodinow, L. and Brun, T. [2014]: 'On the Relation between the Psychological and the Thermodynamic Arrows of Time', *Physical Review, E,* **89**, p. 052102.

Ney, A. [2008]: 'Physicalism as an Attitude', *Philosophical Studies*, **138**, pp. 1–15.

Paul, L. A. [2015] 'Experience and the Arrow', in A. Wilson (*ed*.), *Chance and Temporal Asymmetry*, Oxford: Oxford University Press.

Phillips, I. [2017]: *The Routledge Handbook of Philosophy of Temporal Experience*, London: Routledge.

Price, H. [1996]: *Time's Arrow and Archimedes' Point*, Oxford: Oxford University Press.

Price, H. [2010]: 'Time's Arrow and Eddington's Challenge', *Seminaire Poincare*, **15**, pp. 115–40.

Reichenbach, H. [1956]: *The Direction of Time*, Berkeley, CA: University of California Press.

Sklar, L. [1973]: 'Statistical Explanation and Ergodic Theory', *Philosophy of Science*, **40**, pp. 194–212.

Sklar, L. [1981]: 'Up and Down, Right and Left, Past and Future', *Noûs*, **15**, pp. 111–29.

Sklar, L. [1993]: *Physics and Chance*, Cambridge: Cambridge University Press.

Smith, R. [2014]: 'Do Brains Have an Arrow of Time?', *Philosophy of Science*, **81**, pp. 265–75.

Uffink, J. [2007]: 'Compendium to the Foundations of Classical Statistical Physics', in J. Butterfield and J. Earman (*eds*), *Philosophy of Physics*, Amsterdam: North Holland, pp. 923–1074.

Uffink J. and Valente, G. [2010]: 'Time's Arrow and Lanford's Theorem', *Seminaire Poincare*, **15**, pp. 141–73.

Uffink J. and Valente, G. [2015]: 'Lanford's Theorem and the Emergence of Irreversibility', *Foundations of Physics*, **45**, pp. 404–38.

Van Fraassen, B. [1989]: *Laws and Symmetry*, Oxford: Clarendon.

Von Neumann, J. [1955]: *Mathematical Foundations of Quantum Mechanics*, Princeton, NJ: Princeton University Press.

Wallace, D. [forthcoming]: 'The Logic of the Past Hypothesis', 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.

Weingard, R. [1977]: 'Space-Time and the Direction of Time', *Noûs* **11**, pp. 119–31.

Werndl, C. and Frigg, R. [2015]: 'Rethinking Boltzmannian Equilibrium', *Philosophy of Science*, **82**, pp. 1224–35.

Winsberg, E. [2004]: 'Can Conditioning on the "Past Hypothesis" Militate against the Reversibility Objections?', *Philosophy of Science*, **71**, pp. 489–504.

Winsberg, E. [2012]: 'Bumps on the Road to Here (from Eternity)', *Entropy*, **14**, pp. 390–406.