

Determinism and Chance from a Humean Perspective

Roman Frigg and Carl Hoefer¹

Published in Dennis Dieks, Wenceslao González, Stephan Hartmann, Marcel Weber, Friedrich Stadler and Thomas Uebel (eds.): *The Present Situation in the Philosophy of Science*. Berlin and New York: Springer, 2010, 351-372.

1. Introduction

On the face of it ‘deterministic chance’ is an oxymoron: either an event is chancy or deterministic, but not both. Nevertheless, the world is rife with events that seem to be exactly that: chancy and deterministic at once. Simple gambling devices like coins and dice are cases in point.² On the one hand they are governed by deterministic laws – the laws of classical mechanics – and hence given the initial condition of, say, a coin toss it is determined whether it will land heads or tails.³ On the other hand, we commonly assign probabilities to the different outcomes a coin toss, and doing so has proven successful in guiding our actions. The same dilemma also emerges in less mundane contexts. Classical statistical mechanics (which is still an important part of physics) assigns probabilities to the occurrence of certain events – for instance to the spreading of a gas that is originally confined to the left half of a container – but at the same time assumes that the relevant systems are deterministic. How can this apparent conflict be resolved?

One response to this problem would be to adopt an epistemic interpretation of probability and regard the probability we attach to events such as getting heads when flipping the coin or the spreading of the gas when opening the shutter as a reflection of our ignorance of about the particulars of the situation rather than the physical properties of the system itself. Outcomes really are determined, but we don’t know which outcome there will be and so we use probabilities to quantify our uncertainty about what will happen. There is no contradiction between determinism and probabilities thus understood.

However, this is unsatisfactory. There are fixed probabilities for certain events to occur, which are subjected to experimental test and which, in many cases, are governed by probabilistic laws (such as the laws of statistical mechanics). So these probabilities seem to

have nothing to do with the knowledge, or even the existence, of conscious creatures studying these systems: the chance of a coin to land heads is 0.5 and a gas is overwhelmingly likely to spread when the box is opened, no matter what anybody believes about these events. The values of these probabilities are determined by how things are not by what we believe about them.⁴ In other words, these probabilities are chances, not credences.

This is an unwelcome conclusion because chance and determinism seem to be incompatible. In this paper we argue that at least for a Humean this incompatibility is only apparent and that the problem can be resolved since Humean objective chances are compatible with there being underlying deterministic laws – Lewis’ own proclamation to the contrary notwithstanding.⁵

In our discussion we focus on a simple example, a coin toss, then develop a Humean account of chance, and then show that on this account there is a non-trivial sense in which the coin flips are chance events while at the same time being governed by deterministic laws. This is not lost labour. In fact, as we briefly indicate the last section, chances are introduced into statistical mechanics essentially in the same way as in the case of the coin and so the basic idea of deterministic chance developed here can be carried over to statistical mechanics without (much) further ado.

2. Flipping a Coin

Coin tossing is the most widely used example of a random process, and we are firmly convinced that the chance for getting either heads or tails is 0.5. At the same time we are also firmly convinced that coins obey the laws of mechanics and that therefore their flight as well as their landing heads or tails are determined by their initial conditions and the forces acting upon them. Can we consistently uphold both convictions?

This question has been discussed from a physics point of view by Keller (1986), and later, building on Keller’s work, by Diaconis (1998) and Diaconis, Holmes and Montgomery (2007). We believe that this approach provides all the ingredients need to explain why the chance for heads equals 0.5, and why there is no conflict between this and the fact that coins are governed by the laws of classical mechanics. However, the explanation we offer differs

from Keller's and Diaconis'. We now review in some detail their arguments since they serve as the springboard for our own discussion of chance in coin flips in Section 4.

Keller introduces the following mechanical model of the coin flip. Consider a circular coin of radius r , negligible thickness, and with homogeneous mass distribution. The only force acting on the coin after being tossed is linear gravity, and the surface on which it lands is mushy so that the coin does not bounce. Furthermore the coin is flipped upwards in vertical direction with velocity v at initial height h (above the surface on which it eventually lands) so that it rotates with angular velocity ω around a horizontal axis along the diameter of the coin (i.e. we rule out precession). Solving Newton's equations for this situation (and assuming that the coin is flipped in horizontal position) yields

$$x(t) = vt - \frac{gt^2}{2} + h \tag{1}$$

$$\phi(t) = \omega t, \tag{2}$$

where $x(t)$ is the coin's height at time and $\phi(t)$ the coin's angle relative to the plane. Using the coin's radius one can then determine which point of the coin touches the surface first, and together with the assumption that the coin does not bounce this determines whether the coin lands head or tails. These calculations then allow us to determine which initial conditions result in the coin landing head and tails respectively; that is, they allow us to determine for every quadruple (x_0, v, ϕ_0, ω) whether the coin having this initial condition lands head or tails. We have assumed that all coin tosses start at height h and that all coins leave the hand in horizontal position: $x_0 = h$ and $\phi_0 = 0$); hence different tosses vary in their vertical velocity v and their angular velocity ω . Assuming that the coin starts heads up, initial conditions lying in the black areas of the graph shown in Figure 1 come up heads, while those lying in white areas come up tails. For this reason Keller calls the hyperbolic black and white stripes in Figure 1 the 'pre-images' of head and tails.

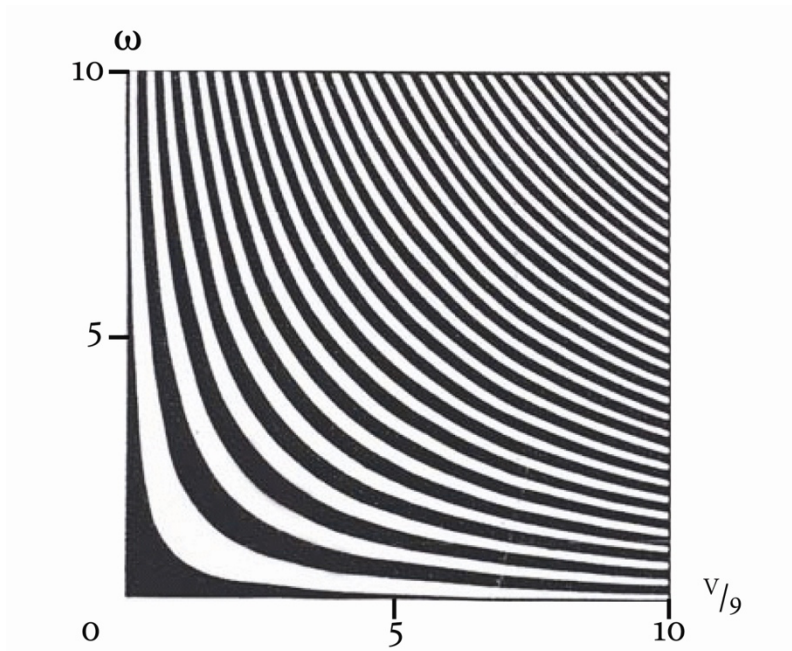


Figure 1. The pre-images of heads and tails (Diaconis 1998, p.803).

What follows from these considerations about the chance of getting heads? Keller presents an argument in two steps. The first is to regard the initial condition as a random variable with a continuous probability distribution $\rho(v, \omega)$ with support in the region shown in Figure 1 (i.e. $\omega \geq 0$ and $v \geq 0$). Then the probability for heads, $p(H)$, is given by

$$p(H) = \int_B \rho(v, \omega) dv d\omega, \quad (3)$$

where B denotes the black regions in $\omega \geq 0$ and $v \geq 0$. Mutatis mutandis Equation (3) also gives the probability for tails, $p(T)$. The second step consists in showing that $p(H) = p(T) = 0.5$. To this end Keller proves a limiting theorem, basically showing that if the boundaries of the region over which the integral in Equation (3) is calculated is shifted towards infinity (i.e. if we integrate over $B' = \{(v, \omega) : v \geq k, \omega \geq k\}$ and let k tend towards infinity), then $p(H) = 0.5$ *no matter what* distribution $\rho(v, \omega)$ we choose. This result becomes intuitively plausible when we realise that the stripes get thinner as the values of ω and v increase (see Figure 1), and so the integral becomes less sensitive to fluctuations in $\rho(v, \omega)$. Hence, in this limit there is a unique probability for heads.⁶

We now turn to a discussion of Humean chance and then return to the question of how to justify $p(H) = 0.5$ in Section 4. The main difference between our and Keller's approach is that we make essential use of facts about the Humean mosaic (i.e., the totality of all actual or occurrent events – see section 3.3) and thereby avoid appeal to a limiting result.

3. Humean Objective Chance

In this section we introduce the concept of Humean Objective Chance (HOC), on which our reconciliation of determinism and chance is based.⁷ The views discussed here are an extension of those introduced in Hoefer (2007), but here presented in a way that pays particular attention to those features of the theory that bear on the issue of the compatibility of determinism and chance.

3.1 Defining Humean Objective Chance

The definition of chance that we present in this section differs from Lewis' canonical definition (1994, p.480). In part this is a matter of presentation; but in part it also results from correcting certain omissions and modifying a few central features. Three changes are particularly crucial. First, we correct the omission of any reference to the Principal Principle (PP) in Lewis' definition. In our view PP is essential for an understanding of objective chance and therefore has to appear in one way or another in its definition. Second, our definition is of chances or chance laws alone, and is not a definition of laws of nature more generally. And finally, of course, our definition will allow for there to be genuine chances in a world that is deterministic at bottom. We return to these points in due course.

Let e be an event, for instance a coin coming up heads or a die landing so that it shows three spots.⁸ We define chance as follows.

Definition 1 (HOC): The chance of event e , $ch(e)$, is a real number in the interval $[0, 1]$ such that:

- (1) the function ch satisfies the axioms of probability,

- (2) $ch(e)$ is the correct plug-in for X in the Principal Principle, and
- (3) the function ch supervenes on the Humean Mosaic in the right way.

Chances thus defined are Humean Objective Chances (HOC); for brevity we refer to them simply as ‘chances’. We use ‘THOC’ to refer to the entire theory of chance presented in this section. The elements of this definition are in need of explication, and providing the needed explications is the task for this section. Let us briefly indicate what this task involves.

The first clause is straightforward, but nevertheless not entirely trivial. Lewis thought it a major problem to prove that objective chances satisfy the axioms of probability, and he argued at length that chances indeed have this property.⁹ In our view there is nothing to prove here. THOC *defines* chance, and we are free to build into a definition whatever seems necessary. A function that does not satisfy the axioms of probability cannot be a chance function and so we simply require that ch satisfy the axioms of probability.

The second clause needs unpacking in two respects: we need to introduce PP, and we need to explicate what makes a plug-in for X a *correct* plug-in. Much hangs on this, and a careful exposition is imperative. For this reason we dedicate subsections to each point (Subsections 3.2 and 3.4).

The third clause is also problematic. We first have to introduce the Humean Mosaic, then say what we mean by a function supervening on the Humean Mosaic, and we then need to explicate the notion of supervening on the Humean Mosaic in the *right way*. The second clause of the definition enters here too, because an important part of what ‘the right way’ means here is: in such a way as to permit a solid argument justifying PP to be made. We turn to these issues in Subsection 3.3.

3.2 Introducing the Principal Principle

Chances, first and foremost, are guides to action. We look to chances when making decisions: if the chance for rain today is 0.95 I take my umbrella with me, but if it is 0.05 I do not. As Lewis insisted, the most central and important requirement on a theory of chance is that it make it possible to see how chances can play this action-guiding role. This aspect of chances

is enshrined in PP, which establishes a connection between chances and the credences a rational agent should assign to certain events, where by ‘credence’ we mean an agent’s subjective probability or degree of belief. The intuitive idea in PP is that a rational agent’s credence for an event e to occur should be set equal to the chance of e , as long as the actor has no ‘inadmissible’ knowledge relating to e ’s occurrence.

Definition 2 (Principal Principle): Let ‘ cr ’ stand for a rational agent’s credence. The Principal Principle (PP) is the rule that

$$cr(e|X\&K) = x, \tag{4}$$

where X is the proposition that the chance of e is x (i.e. $X = ‘ch(e) = x’$), and K is ‘admissible’ knowledge.

Before spelling out what we mean by admissible knowledge, let us add some clarifications about the purpose of K . At first sight it seems unclear why K should appear in Equation (4) at all, and more needs to be said the function that K is meant to be perform. The presence of K should not be interpreted as a request to gather a particular kind of knowledge before we can use PP. On the contrary, we always have knowledge about situations, and K simply stands for the sum of what we *de facto* happen to know. Depending on what kind of propositions K contains, we should or should not use Equation (4) to set our credences. The prescription is simple: if K contains no inadmissible knowledge then use Equation (4); if K does contain inadmissible knowledge then don’t. In the latter case PP is silent about how to set our credences.

The question now is what counts as ‘admissible’ knowledge. Lewis’ original characterisation is:

“Admissible propositions are the sort of information whose impact on credence about outcomes comes entirely by way of credence about the chances of those outcomes.” (Lewis 1980, p.92)

This characterisation has given rise to controversy, and Lewis himself regarded it as too imprecise, and replaced it with a time-indexed version, in part in order to be able to say that all past events have chance 0 or 1. For a discussion of Lewis’ revised definition and the issue of time see Hoefer (2007, 553-5 and 558-60). We here build on this discussion and assume

that these corrections are not only unnecessary, but also wrong. Chances attach to circumstances (the ‘chance set-up’) and not to worlds-at-specific-times. The original definition of admissibility Lewis gave was essentially right. Chance is a guide to action *when better information is not available*. So the essence of the requirement of admissibility is to exclude the agent’s possession of other knowledge relevant to the occurrence of e , the kind of knowledge the possession of which might make it no longer sensible or desirable to set credence equal to objective chance. To use the usual (and silly) example: if you have a crystal ball that (you believe) reliably shows you future events, you may have inadmissible knowledge about a chance event such as the coin flip a week from now. If your crystal ball shows you the coin toss landing tails and you trust the ball’s revelations, you would *not* be reasonable to set your credence in tails to 0.5 for that flip; you have inadmissible knowledge. This example helps make the notion of admissibility intuitively clear, and also points toward a very important fact in our world: inadmissible evidence is not something we typically have – if we did, then chances would be rather useless to have. Still, it is possible to give a slightly more precise definition of admissibility (Hofer 2008, Ch. 2)

Definition 3 (Admissibility): A proposition P is admissible with respect to an outcome-specifying proposition E for chance set-up S (E says that event e occurs) iff P contains only the sort of information whose impact on reasonable credence about E , if any, comes entirely by way of impact on credence about the chances of those outcomes.

This definition makes clear that admissibility is relative to a chance set-up and its attendant possible outcome-events. It is also relative to the agent whose reasonable credence function is invoked in PP. The agent-relativity of admissibility may be more or less extreme, depending on how highly constrained a credence function must be in order to count as ‘reasonable’ or ‘rational’. For our purposes agent-relativity is not germane, and we will assume that all reasonable agents agree about whether a proposition P should or should not have an impact on credence in E , when P is added to a further stock of background knowledge K .

3.3 Humean Supervenience

The *Humean Mosaic* (HM) is the collection of everything that actually happens; that is, all occurrent facts at all times. There is a question about what credentials something must have to

be part of the mosaic. Nothing in what follows depends on how the details of this issue are resolved. What does matter is that irreducible modalities, powers, propensities, necessary connections and so forth are not part of HM. That is the ‘Humean’ in Humean supervenience.

The supervenience part requires that chances are entailed by the overall pattern of events and processes in HM; in other words, chances are entailed by what *actually* happens. Again we can make a comparison with actual frequentism, which satisfies Humean supervenience in a particularly simple way: the overall pattern of events uniquely determines the relative frequency of an event, and hence its probability. Actual frequentism has no frequency tolerance, and hence frequentist probabilities supervene on actual events. This contrasts with propensity theories, which have maximal frequency tolerance. THOC strikes a balance between these extremes by requiring that HOC’s supervene on HM, but not *simply*: THOC postulates that chances are the numbers assigned to events by probability rules that are part of a *Best System* of such rules, where ‘best’ means that the system offers as good a combination as the actual events will allow of *simplicity, strength* and *fit*.

The idea of a Humean Best System of chances can be illustrated with a thought experiment. To this end, we introduce a fictitious creature, Lewis’ Demon. In contrast to human beings who can only know a small part of the Humean mosaic, Lewis’ Demon knows the entire mosaic. The demon now formulates all possible systems of probability rules concerning events in HM, i.e. rules assigning probabilities to event-types such as getting heads when tossing a coin. In the mere formulation of such rules, no interpretation of probability is assumed. The rules in these systems assign numbers to events. These numbers have to satisfy the axioms of probability – this is why they are ‘probability rules’ – but nothing over and above this is required at this stage. Then the demon is asked to choose the best among these systems, where the best system is the one that strikes the best balance between simplicity, strength and fit. The probability rules of the system that comes out of this competition as the best system then, by definition, become ‘chance rules’, and the chance of an event *e* simply is the number that this chance rule assigns to it. That is, the chances for certain types of events to occur (given the instantiation of the setup conditions) simply are what probabilistic laws of the best system say they are.

Definition 4 (Humean *BS*-supervenience): A probability rule is Humean *BS*-supervenient on HM ('HBS-supervenes on HM', for short) iff it is part of a Best System, i.e. the system that strikes the best balance between simplicity, strength and fit given HM.

Clause (3) in Definition 1 can now be made precise: the function *ch* HBS-supervenes on HM.

Needless to say, much depends on how we understand simplicity, strength and fit. Before discussing these concepts in more detail, let us illustrate the leading idea of HBS-supervenience with an example. The question we have to ask is how certain aspects of event-patterns in HM may be captured by adding a chance rule about coin flips. Coins are fairly ubiquitous and we have the custom of flipping them to help make choices. So the event-type we call 'a good flip of a fair coin' is widespread in HM around here. Furthermore, it is a fact, first, that in HM the relative frequency of each discernible side-type landing upward is very close to 0.5 and, second, that there are no easily discerned patterns to the flip outcomes (it is not the case, for instance, that a long sequence of outcomes consist of alternating heads and tails). THOC now asks us to consider all possible probability rules for a given class of events and then choose the one that strikes the best balance between simplicity, strength and fit. There are of course infinitely many rules. One, for instance has it that $p(H)=0.1$ and $p(T)=0.9$; another rule postulates that $p(H)=p(T)=0.5$; and yet another says that $p(H)$ is the actual frequency of heads and $p(T)$ is the actual frequency of tails. Given that the frequency of heads and tails is roughly 0.5, the first rule has bad fit; at any rate its fit is worse than the fit of the other two. But how do we adjudicate between the second and the third rule?

At this point considerations of strength come into play. In fact there may be an even better chance rule that could be part of the Best System, which would embrace coins and dice and tetrahedra and dodecahedra and other such symmetric, flippable/rollable solids. The rule would say that where such-and-such symmetry is to be found in a solid object of middling size with n possible faces that can land upward (or downward, thinking of tetrahedral), and when such objects are thrown/rolled, the chance of each distinct face being the one that lands up (or down) is exactly $1/n$. Given what we know about dice and tetrahedral and so forth, it is quite plausible that this rule belongs in the Best System; and it entails the coin-flip chances. So it enhances both simplicity and strength without much loss in fit, and hence on balance it is better than the system which sets chances equal to relative frequencies. Hence, the chance of

heads on a fair flip of a coin would seem certainly to exist, and be 0.5, in a Best System for our world.

How are we to understand simplicity, strength and fit? Let us begin with simplicity. This is a notoriously difficult notion to define precisely, yet we think that there is enough one can say about it to make THOC tick. As we understand it, simplicity has two aspects, *simplicity in formulation* and *simplicity in derivation*. The former is what is usually meant when simplicity arguments are put forward: a linear relation between two variables is simpler than a polynomial of order 325, a homogenous first order differential equation is simpler than a non-linear integro-differential equation, etc. It is not easy to pin down what general rule drives these judgments, but this does not represent a serious obstacle to us because nothing in what follows depends on simplicity judgments of this kind. Another component of simplicity in formulation is how many distinct probability rules a system contains. *Ceteris paribus*, the fewer rules a system has in it, the simpler it is. The second aspect of simplicity, *simplicity in derivation*, focuses on the computational costs incurred in deriving a desired result. The question is: how many deductive steps do we have to make in order to derive the desired conclusions? Some systems allow for shorter derivations than others. It is important not to confuse simplicity in this sense with a subjective notion of a derivation being ‘easy’ or ‘difficult’. The issue at stake here is the number of deductive steps needed to derive a conclusion, and this is a completely objective quantity, which could be quantified, for instance, by using a measure such as Kolmogorov’s computational complexity (roughly, the length of the shortest programme capable of deriving the result).

Simplicity could always be improved by cutting perfectly good chance rules out of the system. However, in general improving simplicity in this way is not a good strategy because it comes at too high a cost in terms of strength. The strength of the system is measured by its scope. The wider the scope of the system, the stronger it is. In other words, the larger the part of HM that the system is able to account for, the better it fares in terms of strength. The above example illustrates the point: a system that covers only coins is weaker than a system that also covers other chance setups such as roulette wheels, dice, etc.

The Best System should not only ascribe chances to lots of event types, and do so in as simple a way as possible; it should ascribe the *right* chances! But which are the right chances? Every system assigns probabilities to possible courses of history, among them the actual course.

With Lewis, we now postulate that the fit of the system is measured by the probability that it assigns to the *actual* course of history, i.e. by how likely it regards those things to happen that actually do happen. As an illustration, consider a Humean mosaic that consists of just ten outcomes of a coin flip: *HHTHTTHHTT*. It follows immediately that the first system above ($p(H)=0.1$ and $p(T)=0.9$) has worse fit than the second ($p(H)=p(T)=0.5$) since $0.1^5 0.9^5 < 0.5^{10}$. This example also shows that a system has better fit when it stays close to actual frequencies, as we would intuitively expect.¹⁰

So the ways in which we evaluate systems is objective and no appeal to ‘pragmatic’ or specifically ‘human’ values or limitations has been made. Nevertheless, we accept two (not very controversial) assumptions that assure that the Best System, whatever its concrete form, shares at least some essential characteristics with science as we, humans, know it. The first assumption is *ontological pluralism*, which denies that only basic/micro entities exist. Some hard-headed reductionists deny that anything except the basic micro entities exist. Thus, chairs, rivers, cats, trees, etc. are said not to exist. We deny this. That coins consist of atoms does not make coins unreal. Coins exist, no matter what micro physics tells us about their ultimate constitution, and so do rivers, chairs, and cats. Hence, even in a classical world, HM consists of much more than elementary particles and their trajectories.

The second assumption is *linguistic pluralism*, the posit that the language in which the Demon formulates the systems that subsequently enter into the simplicity-strength-fit competition contains terms for macroscopic kinds. That is, the language has not only the vocabulary of microphysics, but also contains terms like ‘coin’ and ‘river’. So we not only believe that macro objects exist, we also equip the demon with a language in which he can talk about these as *sui generis* entities.¹¹

3.4 Justifying the Principal Principle

There is controversy not only over the correct formulation of PP, but also over its status. Strevens (1999) argues that it is no more possible to offer a sound argument justifying PP than it is to justify induction, and that we therefore have to accept it as something like a first principle. But not everybody shares this pessimism. In fact, we believe that the unique features of HOC permit a demonstration that it is irrational not to apply PP when its

conditions are fulfilled. Space precludes a full discussion here, so we will simply present a brief version of the argument; for an in-depth discussion see (Hoefer 2008, Ch. 4).

As we have seen in the last subsection, it is a result of a careful analysis of what it means for the function ch to supervene on the Humean Mosaic in the right way that whenever there is a large number of instances of a chance setup, the chance of a certain outcome is close to the relative frequency of that outcome. For this reason, THOC can be understood as a (major) sophistication of finite frequentism, and understanding why PP is justified for HOC begins by recalling this affinity.¹²

There are two ways of justifying PP based on this affinity, an ‘*a priori*’ and a ‘consequentialist’ one. The former is similar to the justification of PP Howson and Urbach (1993) give for von Mises frequentism. A subjective degree of belief corresponds (by definition) to the odds at which an agent feels a bet on either side of a question (E versus not- E) would be fair. An agent who has no inadmissible information pertinent to the outcomes of a *long series* of instances of chance setup S should have the same degree of belief in the E -outcome in *each* trial – having a *reason* to assign a higher or lower degree of belief to E on a specific trial automatically and by definition amounts to possessing inadmissible information. Hence, if an agent assigns degree of belief p to outcome E in a single trial of chance setup S , he should assign the same credence to an E -outcome in each instance of a(n indefinitely) *long series* of trials of S ; not to do so is to take oneself to have information relevant to an E -outcome that does not come from E ’s chance (which by stipulation is the same in each trial of S), and hence to have *inadmissible* information. Having inadmissible information makes PP inapplicable, so we may assume for the rest of the argument that the agent does not vary his credence in E -outcomes from trial to trial.

So the agent takes betting on E in each trial of an indefinitely long series at odds $p:(1-p)$ to be fair. Assume that he bets on the same side in all trials in the sequence, i.e. either on E in all trials or not- E in all trials. Because the agent thinks the bet is fair, he must think that there is no advantage to betting on E rather than not- E (or vice versa); that is, he must be indifferent towards which side of the bet he takes. By assumption there is a chance for getting E on a trial, $ch(E)=q$. From the account of THOC above we know that the relative frequency of E ’s in an indefinitely long sequence of trials must be equal (or at least very close to) the chance of E . It is a simple result of probability calculus that if agents don’t bet in accordance with

relative frequencies, then one side of the bet is doing better. This cannot be if the agent believes the bet to be fair. It is then a simple arithmetic fact that if q differs non-trivially from p , and the agent bets on E at $p:(1-p)$ odds throughout the long series, then the agent will certainly lose (or win) in the long run. The agent, understanding THOC and that $ch(E) = q$, must know all this too; but he cannot believe this and yet believe the long series of bets to be fair. So if $p \neq q$, the agent holds contradictory beliefs, which is irrational. So the only rational assignment of probabilities is $p = q$, as PP prescribes.

The consequentialist argument is more straightforward. It asks us, in the spirit of Humeanism, to look at HM, which not only contains outcomes of trials but also agents placing bets. If we look at all agents placing bets across the entire mosaic and check on how they are doing, we will see that those agents who set their credences equal to the chances obtain – at most places and times, at least – better results than those who adopt credences significantly different from the chances. In other words, in the wider domain just as in Las Vegas, if one has to gamble on chancy events, one does best if one knows the objective probabilities. For this reason it is rational to set one's credences to objective chances, as PP requires.¹³

3.5 The Epistemology of HOC's

Let us close this section with a brief remark about the epistemology of HOC's. At first sight, an approach to probability whose central concepts are defined in terms of everything that actually happens at any point in time and at any spatial location – the HM – and an omniscient creature – Lewis' Demon – may strike some as rather disconnected from actual human endeavours. This impression is mistaken. Needless to say, the appeal to HM and Lewis' demon are idealisations, and ones that take us rather far away from our actual epistemic situation. But we deny that this turns THOC into an epistemic pipe-dream. First, the limitations of actual human experience are a factor that every epistemology has to cope with. In particular, also those positions who believe in metaphysically 'thick' laws and probabilities (universals, causal powers, etc.) have to base their views on the nature and character of these on actual experience and there is the possibility that future events may prove them wrong. Any view about probabilistic laws – Humean or not – has to base claims about these on our limited actual experience, and this involves an inductive leap. How handle this leap is of course a time-honoured philosophical puzzle on which much ink has been spilled, and there is

no royal road to success. The point to stress here is that the Humean is not alone with this problem. Second, the requirement that only occurrent properties be part of HM is in harmony with scientific practice as we know it, since occurrent properties are what science can observe. In this respect THOC is even closer to actual science than approaches that postulate modal entities that science can never observe. Third, the rules that are given to the Demon have an obvious ‘human flavour’: the omniscient Demon himself would probably not care about simplicity and strength since he knows everything anyway. These requirements are metatheoretical virtues humans value in science and hence what the Demon is asked to do is in the end ‘human style’ science as best as it can be done. Hence the Demon’s activity is not different in kind from the endeavours of human scientists; the difference is that he can perform with perfection what we can do only inadequately.

4. Coin Flipping for Humeans

Let us now return to flipping coins. A striking feature of the discussion so far is the almost complete mismatch between how probabilities for the coin flip were treated in Sections 2 and 3 respectively. The treatment in Section 2 started with a deterministic mechanical model and sought to retrieve the 50/50 chance rule from mechanical laws plus a probability distribution over initial conditions. The approach taken in Section 3 did not mention mechanics at all and instead focussed the pattern of outcomes in HM. At least on the face of it these approaches have little in common and so the question arises whether they are compatible at all, and if so how.

In this section we argue that they are compatible, and, what is more, that they are actually complementary. In order to reach this conclusion we need to analyse the two accounts and their status *vis-a-vis* each other in greater detail. To facilitate the discussion, we set up a temporary debate between two viewpoints: ‘mechanicism’ *versus* ‘macro-statistics’, their proponents being ‘mechanicists’ and ‘macro-statisticians’.

Mechanicists are likely to argue that their point of view is privileged since their account is based on fundamental laws: by assumption we live in a classical universe and so HM consists of trajectories of objects, among them the trajectories of coins, and classical mechanics is the fundamental theory of this universe.¹⁴ Chance rules, if there are any at all, have to be

formulated in terms of the fundamental entities of HM and in the language of the fundamental theory describing them. Equation (3) fits the bill: it is a rule that assigns probabilities to getting heads when tossing a coin, and it does so solely in terms of basic mechanic properties. The rule is simple, has good fit, and since Keller (1986, pp.194-6) shows that it can easily be generalised to other chance setups such as roulette wheels it also has strength. So we have good reasons to believe that within the class of all probability rules Equation (3) is the one that strikes the best balance between simplicity, strength and fit, and hence the probabilities it defines are chances in the sense of THOC.

The macro-statistician disagrees with this point of view for two reasons. The first objection is conceptual, the second technical. The conceptual objection takes issue with the mechanist's reductionist outlook. Even if the world is classical at bottom and classical mechanics is the fundamental theory of the universe, it does not follow that everything that can be said about HM has to be said in the language of the fundamental theory. More specifically, the macro-statistician adopts a *methodological pluralism* (MP), the position that probability rules can be formulated in a macro language pertaining to a certain level of discourse, and that probabilities thus introduced are *sui generis* HOC's if the probability rules in question strike the best balance between of simplicity, strength and fit relative to all other Best Systems. To do this, they need not prove logical independence from micro-level chance rules; they need only win out in competition with alternate rules *formulated in the same language*, that of the macro-level. On this view, then, the $1/n$ rule for gambling devices is a *sui generis* chance rule because it strikes a better balance between the three basic metatheoretical virtues than any other probability rule formulated in the language of coins, wheels, throws, etc. (We come back to this principle at length below.)

The macro-statistician's technical objection to mechanicism turns on the status and mathematical form of the distribution $\rho(v,\omega)$ in Equation (3). At a general level the worry is that the mechanist is 'cheating'. No probabilities can ever come out of a purely deterministic approach ('no probabilities in, no probabilities out'), and the mechanist just puts them in by hand when he introduces $\rho(v,\omega)$, which is not warranted by (or even related to) any posit of mechanics. Therefore the introduction of $\rho(v,\omega)$ is an *ad hoc* manoeuvre, unmotivated from the point of view of mechanics. And, as is often the case with such manoeuvres, it may well raise more question than it answers. The first problem with $\rho(v,\omega)$ is that it is not clear what it is a distribution *for*. The most basic question we have to ask about every probability

distribution is: what are these probabilities probabilities for? And it is not clear what the answer in the case of the coin is. We might take it to be giving the probability of a coin flip's having initial conditions within a given range $v + dv$, $\omega + d\omega$. But nothing in mechanics can ground such a distribution.

The problem with the mathematical form of $\rho(v,\omega)$ is the following. Keller's limiting argument shows that the mathematical form of $\rho(v,\omega)$ is immaterial, and hence the question of what $\rho(v,\omega)$ to chose becomes obsolete. However, this limiting argument is of no relevance to *actual* coin tosses. Diaconis has shown in experiments that for typical coin tosses the initial upwards velocity v is about 5 mph and the frequency ω lies between 35 and 40 revolutions per second (1998, p.802). This is very far away from infinity! The problem is that once as we revoke the infinite limit, it is no longer irrelevant what $\rho(v,\omega)$ one chooses. So which $\rho(v,\omega)$ is the right one to plug into Equation (3)? Intuitively one would choose a uniform distribution. For one it is simple; for another, it would give (roughly) a 0.5 probability for heads since, as becomes obvious from Figure 1, the black and the white stripes occupy approximately the same area. But nothing in the mechanical approach justifies this assumption.

Let us now step back, evaluate the arguments on either side, and explain how the two views eventually come together. Take the mechanist's insistence on fundamental laws first. He will object to MP on the grounds that the chance for heads is not independent of the micro physics of the world. Surely, so the argument goes, there must be some dependence there! If the physics of our world was vastly different from what it is, then the chance for heads should be different too!

There is a grain of truth in this, but we must not be misled. The physics of our world might be vastly different, and yet (for *whatever* reason) the pattern of heads- and tails-outcomes in HM might be exactly the same; in which case, the chances would be the same. But in most reasonably imaginable counterfactual scenarios, the physics will matter much less than the actual pattern of outcomes in HM. Macro level events depend *ontologically* on micro-level facts – in the obvious compositional sense – but in our world do not depend on them *constitutively* (i.e., the macro chance facts are not entailed directly or indirectly by fundamental physics; they depend on the pattern of macro events no matter what the micro physics is.).

Once this is realised, the other problems have elegant solutions too. We can chop the Gordian knot in four cuts.

First, with the macro-statistician we affirm MP, from which it follows that chances for macro events like coin flips depend on the outcome pattern, not on the details of the underlying physics. (We justify MP below).

Second, with the mechanist we take ontological dependence seriously. The question is how to take this into account, and this is where a new element enters. We share with the mechanists the view that Equation (3) – and similar equations – matter, but interpret them differently. This equation does not *give* us the chance for heads. We don't need to be given anything – we have the chance, and the chance is (constitutively) independent of the micro physics. Rather, we see Equation (3) both as a 'consistency check' and an explanation. Let us take these in turn. The different parts of a Best System have to be consistent with each other (which is not to say that one has to be derivable from the other). For this reason, whenever the setup conditions of a macro-level chance rule and a micro-level chance rule are the same (extensionally equivalent), then the chances they ascribe must agree or be very nearly in agreement. This, of course, does not rule out the possibility of minor adjustment. For example, assume we adopted the 50/50 rule for heads and tails. Now we know for sure that we get reductive relations right and we have the correct micro theory, and based on these we find 49/51. This is no real conflict because there is some flexibility about the macro chances and if there are very good overall reasons for making adjustments, then the Humean can make these. But there is a breaking point: if the micro theory predicts 80/20, we have to go back to the drawing board. The second element is explanation. We don't want to place too much emphasis on this, but there is the pervasive intuition that if a macro result can be derived from a more fundamental theory, there is explanation. Those who share this intuition – among them us – can see Equation (3) as providing an explanation. Those who don't can renounce explanatory goals and rest content with the consistency requirement.

The third cut is that the mechanist has to admit that the introduction of $\rho(v,\omega)$ is a step beyond mechanics and as such $\rho(v,\omega)$ has to be justified elsewhere. But far from being a problem, this actually is an advantage. When thinking about $\rho(v,\omega)$ in the 'THOC way', we immediately have a natural interpretation of $\rho(v,\omega)$: it is the relative frequency of certain

initial conditions. Of course all actual initial conditions are a collection of points in the $v-\omega$ plane, and not a continuous distribution. But arguably a continuous distribution is much simpler (in the sense of simplicity in formulation) than a huge collection of points, and so the Humean can argue convincingly that fitting a suitable continuous distribution through the points makes the system simpler and stronger. This distribution then is just an elegant summary of the actual initial conditions of all coin flips in HM.

Fourth, the common intuition that there is something epistemic about the chance of getting Heads on this flip – after all it has one and only one initial condition and given this initial condition it is determined whether it comes up head or tails – is addressed by paying close attention to THOC's prescription about when to use chances to guide our credence. Information about the precise initial condition of a given coin flip is certainly inadmissible: such information logically implies the coin toss outcome and hence provides knowledge about the outcome of a toss that does not come by way of information about chances. The crucial point is that in typical situations in which we toss a coin, we just don't have inadmissible information, and that is why we use chances and PP to set our degrees of believe. So we use chances when we lack better knowledge.

Let us illustrate the admissibility point in some more detail. Consider the scenario described in Section 2, and an agent A who has only the usual sort of knowledge in his background K and who needs to decide how to bet on the coin flip. A should apply PP, clearly, and set his credences for heads and tail outcomes to 0.5. But now consider agent L , a Laplace-demon-in-training, who also must decide how to bet. L knows all that A knows, but – crucially – L also knows the exact micro-state of the world (or a big enough local region of it) just prior to the flip, and knows the laws of Newtonian mechanics. Should L set her credences for heads and tails outcomes equal to 0.5? Evidently not! She can calculate, on the basis of her background K , precisely what will happen. Let's assume she calculates that the coin will in fact land heads. L has inadmissible knowledge.¹⁵ She has information relevant to whether the coin will land heads (*maximally* relevant!), and the information is *not* relevant by way of telling her about the objective chances. So L should not apply PP; and this is intuitively the right verdict. The conclusion is not that the objective chance of heads is 1. It is that (given the past state and the laws), the coin *will* land heads; and anyone who is aware of these facts should set their credence in heads to 1 (as the rules of subjective probability require), and not to 0.5.¹⁶ The truth of deterministic laws entails that, given a complete-enough state of affairs at a moment

of time (and perhaps boundary conditions), future events are fully determined. And this entails that *if you can get* such Laplace's-demon style information, and if you can actually calculate anything with it, then you may have better information with which to guide your credences about future events than the information HOCs give you. What is entailed, in other words, is not that objective chances do not exist, but rather that certain godlike beings may not have any use for them. We humans, alas, never have had nor will have either such information about initial conditions, or such demonic calculational abilities. For us, it is a good thing that objective chances exist, and that we can come to know (and use) them.

With these points in mind, now we can see how determinism and non-trivial objective chances are compatible, and we also see that the admissibility clause in PP plays a crucial role in that.

We now turn to a defence of MP, which we merely stated above. Why should we subscribe to this principle? Why would a best system contain anything like chance-rules about coins and other macro objects? Let us distinguish two cases, a world in which physicalist reductionism about chance is true, and one in which it is false. Physicalist reductionism about chance is the claim that all chance-facts arise out of the laws of physics. Physicalist reductionism quite generally (not merely about chance) is popular in particular with elementary particle physicists; see for instance Weinberg (1994).

If reductionism of this kind is false, then it is obvious that the best system would contain rules about macro objects: these rules don't follow from basic laws of physics and therefore putting them into a system will greatly increase its strength. The more difficult case is if chance reductionism is true. If the rules about coins and wheels are but complicated applications of the laws of physics, why would we have such rules in our best system?¹⁷ This seems to make the system less simple without adding strength. The reason to put them in nevertheless is what we above called simplicity in derivation: it is hugely costly to start from first principles every time you want to make a prediction about the behaviour of a roulette wheel. So the system becomes simpler in that sense if we write in rules about macro objects.

There is also a more intuitive argument why this independence of chances from micro physics is correct. It is the basic posit of Humeanism that the chance of a certain event HBS-supervene on the pattern of occurrence of events of the same kind in HM, and as such this

chance is independent of how these events relate to other features of HM. In our concrete example this means that the chance of heads only depends on the pattern of heads in HM, or perhaps the pattern of outcomes in rolls/flips of n-sided solids with the appropriate symmetries and not on the relation that ‘obtaining heads’ bears to other parts of HM, in particular the basic mechanical properties of matter. As noted above, these sorts of patterns may obtain even in worlds with radically different micro-laws. Imagine a universe in which in which matter is a continuum and obeys something like the laws of Cartesian physics; imagine that coins exist in this universe and are tossed repeatedly. Despite the basic physics being very different, suppose it turns out that the overall pattern of outcomes of rolls/tosses of such n-sided objects in the continuum universe’s HM is very similar to the pattern in our universe. What would the chance of heads be in the continuum universe? Clearly it would be given by the $1/n$ rule, since this is the best rule relative to the HM, irrespective of the micro-constitution of matter.

5. Envoy

As we have indicated in the introduction, this paper is about more than coins. In fact exactly the same considerations can be used to explain chance in statistical mechanics (SM). A full exposition of this theory is beyond the scope of this paper, but we would like to bring our discussion to a close by very briefly indicating how the insights gained with the example of the coin carry over to SM.¹⁸ Consider a typical SM system, for instance a gas in container. The gas consists of about 10^{23} molecules. These molecules bounce around under the influence of the forces exerted onto them when they crash into the walls of the vessel and when they collide with each other. The motion of each molecule under these forces is governed by the laws of mechanics. Hence the gas is a large mechanical system: its state is fully specified by a point in its (6×10^{23} -dimensional) phase space – in this context referred to as its ‘micro state’ – and its evolution over time is fully determined by the laws of mechanics.

At the same time the system is always in a certain macro-state, which is characterised by the values of macroscopic variables, in the case of a gas pressure, temperature, and volume. It is one of the fundamental posits of (Boltzmannian) SM that a system’s macro-state supervenes on its micro-state, meaning that a change in the macro-state must be accompanied by a change in the micro-state. For instance, it is not possible to change the pressure of a system and at the

same time keep its micro-state constant. Hence, to every given micro-state there corresponds exactly one macro-state. This determination relation, however, is not one-to-one. In fact many different micro-states can correspond to the same macro-state. We now group together all micro-states corresponding to the same macro-state, which yields a partitioning of the phase space into non-overlapping regions. We can then define an entropy function (the so-called Boltzmann entropy) that assigns a particular entropy value to every macro-state.

Systems characteristically start off in a low entropy state and then evolve into equilibrium, the macro-state with maximum entropy. The Second Law of thermodynamics tells us that this is what invariably must happen. One of the central aims of SM is to show that the Second Law – which is a purely macroscopic law – actually is a consequence of the mechanical motion of the molecules of the gas, and it does so by showing that the approach to equilibrium is overwhelmingly likely. And this is where we make contact with the coin example. In order to judge something as likely, trivially, we must introduce probabilities. SM does this by putting a uniform probability measure over the region of phase space which corresponds to the system's initial low entropy state, and then aims to show that micro conditions that lie on trajectories which eventually move towards equilibrium are overwhelmingly likely. The logic of this is like in the case of the coin, the only difference being that we sort initial conditions into ones that behave as the Second Law requires and ones that don't, rather than into ones that yield heads and one that yield tails. Let us then mark the ones that behave as we expect white and the other ones black. We then put a measure over these all initial conditions of the same kind as ρ above. The difference just lies in the values: we now don't expect a 50/50 division between white and black, but rather something like 99.9999/0.0001 (omitting many 9s and 0s here for brevity). But the basic idea is the same: put a distribution over initial conditions and show that the outcome probabilities entailed fit well with the patterns in actual events. And indeed they do, not only the (essentially) exceptionless pattern of Second Law behaviour for macroscopic fluids, but also non-trivial probabilities for smaller collections of particles.

So what we have learned from the coin also solves the problem of interpreting probabilities in SM! They can be elegantly accommodated in a Humean theory of objective chance.

Notes

¹ The authors are listed alphabetically; the paper is fully collaborative. To contact the authors write to r.p.frigg@lse.ac.uk and carl.hoefer@uab.es.

² Or if one insists that at bottom the world is quantum mechanical, then the problem is that probabilities like the ones we attach to coin flips don't reduce to the micro probabilities since quantum mechanics assigns values close to either 1 or 0 rather than 1/2 to events like getting heads when tossing a coin.

³ For a discussion of determinism see Earman (1986, Ch. 2).

⁴ This point is often made in the context of statistical mechanics; see for instance Albert (2000, p.64), Loewer (2001, p.611) and Goldstein (2001, p.48); see also Hoefer (2007, p.557, pp.563-4) and Maudlin (2007, pp.281-2).

⁵ Loewer (2001; 2004) has presented a reconciliation of determinism and chance from a Humean perspective. However, we believe this reconciliation to be problematic for the reasons discussed in Frigg (2008b).

⁶ Diaconis *et al.* (2007) generalise this result by relaxing some of the above modelling assumption and thereby taking into account the precession of the coin. This adds interesting features to the model, but since the main features remain the same we keep using the simple model discussed in this section.

⁷ One might argue that 'objective chance' is a pleonasm since chances are objective by definition. True enough, but the phrase 'objective chance' has become customary in the literature and so we stick to it here.

⁸ Two disclaimers are in order. First, nothing in what follows depends on a more precise characterisation of events. Second, we attribute chances to events because this looks most natural in the cases we discuss. But nothing hangs on that; we could take propositions instead. In fact, as will become clear from the context, in certain formulae below letters such as e and X will stand for propositions describing events rather than directly for events. This is inconsequential for our views on chance.

⁹ For a discussion of Lewis' arguments see Hoefer (2007, pp.560-62).

¹⁰ Elga (2004) argues that this definition of fit runs into problems if there are infinitely many chancy events, and suggests a solution to based on the notion of a typical sequence. This concern is orthogonal to the problems we discuss in this paper and hence we will not pursue the matter further.

¹¹ For further discussion of the issue of the language used in formulating laws see Lewis (1983).

¹² Since clause 2 of our definition of HOC above stipulated that $ch(e)$ is the correct plug-in for the Principal Principle, one might expect a quick and easy demonstration that HOCs satisfy PP: it is true by definition! Clearly, this is a bit *too* easy. The two justifications of PP for HOC offered in this section are substantial, departing from connections between HOC and frequencies of events, and are entirely non-circular.

¹³ Caveats and details of the consequentialist argument are discussed in Hoefer (2007, sec. 5) and (2008, Ch. 4).

¹⁴ Those who also uphold micro-reductionism – the view that matter consist of atoms and the behaviour of macroscopic objects like coins eventually has to be explained in terms of the behaviour of its micro constituents – can replace the trajectory of a coin by the bundle of trajectories pertaining to the atoms making up the coin. Mutatis mutandis the arguments remain the same.

¹⁵ According to Lewis' official definition of admissibility, information about laws of nature and about past states of the world are fully admissible, hence L does not have inadmissible information. This adjudication makes it

impossible for Lewis to retain non-trivial chances if the true laws of nature are deterministic. For a discussion of this point see Hoefer (2007, pp.553-555).

¹⁶ Formally, $Cr(H|XK) = 1$ is required by the probability axioms, since $K \supset H$. We emphasize, it is *not correct* by contrast to say that $K \supset [Ch(H) = 1.0]$.

¹⁷ It may be hard to see how probability-facts could follow from fundamental physical laws, or laws plus initial conditions even, if the laws are fully deterministic. We do believe that 'no probabilities in, no probabilities out' holds here. But one might posit a fundamental-physics probability law as a supplement to the deterministic laws, precisely in order to allow derivation of probabilities for a variety of physical event types, including perhaps macro events. Loewer's version of Best System Humeanism does precisely this; see Loewer (2001).

¹⁸ For a detailed discussion of statistical mechanics see Uffink (2006) and Frigg (2008a).

References

David Albert, *Time and Chance*. Cambridge/MA and London: Harvard University Press 2000.

Persi Diaconis, "A Place for Philosophy? The Rise of Modeling in Statistical Science", in: *Quarterly of Applied Mathematics* 56, 4, 1998, pp.797-805.

Persi Diaconis, Susan Holmes, and Richard Montgomery, "Dynamical Bias in the Coin Toss", in: *SIAM Review* 49, 2, 2007, pp.211-235.

John Earman, *A Primer on Determinism*. Dordrecht: Reidel 1986.

Adam Elga, "Infinitesimal Chances and the Laws of Nature", in: *Australasian Journal of Philosophy* 82, 2004, pp.67-76.

Roman Frigg, "A Field Guide to Recent Work on the Foundations of Statistical Mechanics", in: Dean Rickles (Ed.), *The Ashgate Companion to Contemporary Philosophy of Physics*. London: Ashgate 2008, pp.99-196.

——— "Humean Chance in Boltzmannian Statistical Mechanics", *forthcoming in Philosophy of Science* 2008.

Sheldon Goldstein, "Boltzmann's Approach to Statistical Mechanics", in: Jean Bricmont, Detlef Dürr, Maria Carla Galavotti, Gian Carlo Ghirardi, Federico Petruccione and Nino Zanghì (Eds.), *Chance in Physics: Foundations and Perspectives*. Berlin and New York: Springer 2001, pp.39-54.

Carl Hoefer, "The Third Way on Objective Probability: A Sceptic's Guide to Objective Chance", in: *Mind* 116, 463, 2007, pp.549-596.

——— *Chance in the World*: Book Manuscript 2008.

Colin Howson and Peter Urbach, *Scientific Reasoning: The Bayesian Approach*. 2nd ed. Chicago: Open Court 1993.

Joseph B. Keller, "The Probability of Heads", in: *American Mathematical Monthly* 93, 3, 1986, pp.191-197.

David Lewis, "A Subjectivist's Guide to Objective Chance", in: Richard C. Jeffrey (Ed.), *Studies in Inductive Logic and Probability. Vol. 2*. Berkeley: University of California Press 1980, reprinted in Lewis 1986, 83-132, with postscripts added.

——— "New Work for a Theory of Universals", in: *Australasian Journal of Philosophy* 61, 4, 1983, pp.343-377.

——— "Humean Supervenience Debugged", in: *Mind* 103, 1994, pp.473-490.

- Barry Loewer, "Determinism and Chance", in: *Studies in History and Philosophy of Modern Physics* 32, 2001, pp. 609-629.
- "David Lewis' Humean Theory of Objective Chance", in: *Philosophy of Science* 71, 2004, pp.1115-1125.
- Tim Maudlin, "What Could Be Objective About Probabilities?" in: *Studies in History and Philosophy of Modern Physics* 38, 2007, pp.275-291.
- Michael Strevens, "Objective Probability as a Guide to the World", in: *Philosophical Studies* 95, 1999, pp.243-275.
- Jos Uffink, "Compendium of the Foundations of Classical Statistical Physics", in: Jeremy Butterfield and John Earman (Eds.), *Philosophy of Physics*. Amsterdam: North Holland 2006, pp.923-1047.
- Steven Weinberg, *Dreams of a Final Theory: The Scientist's Search for the Ultimate Laws of Nature*. New York: Vintage 1994.