

Article

Not peer-reviewed version

---

# Epistemic Risk and the Transcendental Case Against Determinism

---

[Alessio Montagner](#)\*

Posted Date: 8 December 2025

doi: 10.20944/preprints202510.2441.v2

Keywords: free will; transcendental argument; epistemic risk; modal closeness; normalcy



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

# Epistemic Risk and the Transcendental Case Against Determinism

Alessio Montagner

Independent Researcher, Italy; alessio@hermesambiente.it

## Abstract

A familiar intuition holds that determinism creates an epistemically adverse context. This paper gives that intuition a formal shape by developing a new epistemic transcendental argument (ETA) grounded in the notion of epistemic risk. First, we formalise epistemic risk through a metric space  $W$  equipped with two metrics,  $D$  and  $N$ , corresponding to distinct theories of risk. Drawing on the notions of modal closeness and normalcy, we argue that these metrics better capture our intuitions about risk than traditional similarity-based accounts. Building on these insights, we articulate an argument based on five axioms. The axioms are philosophically motivated using the two metrics, their independence is verified in Mace4, and the derivation of the denial of determinism is formally carried out in Lean 4.

**Keywords:** free will; transcendental argument; epistemic risk; modal closeness; normalcy

---

## 1. Introduction

Epistemic transcendental arguments (ETAs) begin from the premise that knowledge is possible and infer that whatever would make knowledge impossible must be false. A notable class of ETAs targets determinism. The earliest instance appears in Epicurus (Honderich, 1988, p. 361), is echoed by Sextus Empiricus, and reemerges in Kant (Bernstein, 1988, p. 359) and several later thinkers See Wick (1964), Mascall (1965), Mabbott (1966), Malcolm (1968), Jordan (1969), Lucas (1970), Boyle, Grisez & Tollefsen (1972), Ripley (1972), Snyder (1972), Hasker (1973), Magno (1984), Walker (2001), Slagle (2016), Lockie (2018, pp. 182-183), Chevarie-Cossette (2019), and Steward (2020). For a pragmatic variant, see Latham (2019). Related arguments against naturalism include Balfour (1879, pp. 260-276), Haldane (1929, p. 209), Joad (1933, p. 99), Lewis (1947, §3), Taylor (1963, p. 110-111), Popper & Eccles (1984, p. 75), Moreland (1987, §3), and Plantinga (1993, pp. 229-237). The argument is often construed as a form of self-refutation: if determinism is false, one ought not to believe it; if true, one would believe it merely because one is determined to do so, thereby lacking justification in either case.

A different, largely neglected strategy begins from an epistemological rather than a doxastic concern. As Gisin observes,

If we did not have free will, we could never decide to test a scientific theory. We could live in a world where objects tend to fly up in the air but be programmed to look only when they are in the process of falling. (Gisin, 2014, p. 90)

The point is that, under determinism, an agent may be determined in epistemically undesirable ways — for instance, never to adopt a mostly true belief system. Since the agent may instead be determined in desirable ways, this merely introduces an *epistemic risk*: the risk of believing falsehoods. On epistemic risk as a refinement of anti-luck epistemology, see Pritchard (2016) and Navarro (2021). Pritchard addresses the puzzle of why luck arising from a high probability of false belief undermines knowledge, whereas luck arising from a high probability of no belief does not. Risk, unlike luck, denotes a negatively valued event, thereby explaining why only certain forms of luck are knowledge-undermining.

Of course, such risk is not exclusive to determinism: since a course of action could occur either because one is determined to carry it out or because one freely chooses to do so, epistemic risk can also arise under libertarian freedom. Yet the asymmetry is evident. In a libertarian setting, multiple possible futures remain open, and belief states can vary accordingly. In a deterministic setting, by contrast, there is a single possible future, and if the agent is determined to sustain false beliefs, nothing can alter that course. Hence, epistemic risk appears structurally higher under determinism.

Existing ETAs against determinism have overlooked this dimension. The intuition behind Gisin's remark, though familiar, has never been articulated as a formal argument. This paper develops that intuition into a new ETA: if knowledge requires low epistemic risk, and if determinism entails a systematically higher level of such risk, then determinism undermines the possibility of knowledge. The resulting argument diverges from the Epicurean lineage of anti-determinist reasoning and grounds a distinct transcendental challenge in epistemic risk aversion.

In §2 and §3, we formally state our conception of risk. In §4, we present the axioms of our ETA and, in §5, we provide their justification on the basis of our account of risk. In §6, we show that the adopted axioms are incompatible with determinism.

## 2. A metric Framework for Epistemic Risk

### 2.1. General Introduction

Any argument based on risk requires a means of comparing possible worlds. In fact, a belief is risky when there are worlds in which it is false, and those worlds have a specific relation with the actual world. To make this conception precise, we introduce a metric structure over the space of possible worlds.

Several formalisms could express relations between possible worlds. One could, for instance, define risk in terms of a fuzzy accessibility relation, with higher accessibility corresponding to greater risk. Yet accessibility and risk should remain conceptually distinct to avoid ruling out, *a priori*, the possibility of risk originating from inaccessible worlds.

A second option is to define a ternary relation  $R(x, y, z)$ , read as "the world  $x$  is closer to  $y$  than to  $z$ ". While this can be implemented by a ranking function or a partially ordered frame, it makes the notion of *equiproximity* — which we need — hard to express, and offers no straightforward way to impose numerical thresholds for risk.

For these reasons, a metric space provides a more perspicuous framework. Distances can be represented numerically, thresholds can be fixed, and quantitative comparisons between sets become possible.

In what follows, we formalise the concept of risk using a metric space. In §3, we show how these metrics were derived. We will then use them to justify the axioms we will use in our argument.

### 2.2. Possible Worlds and Parameters

Let  $W$  be the space of possible worlds. Each  $w \in W$  is represented as a tuple of parameters  $p$ :

$$w = \langle p_1(w), \dots, p_n(w) \rangle$$

Each parameter captures a dimension along which worlds may differ, identifying commensurable aspects between worlds and ensuring a canonical correspondence between facts. Parameters can be of two kinds:

1. **Ordered parameters**, whose values belong to a domain endowed with a total order or a topology that admits a natural metric — i.e., a pair  $(S_i, M_i)$  where  $M_i$  metrises  $S_i$ .
2. **Categorical parameters**, whose values belong to a discrete label space  $L_i = l_1, \dots, l_n$  lacking natural order.

For instance, the distance of a dart from a target's center is an ordered parameter, the outcome of a coin toss is a categorical one. Parameters whose values are purely random (i.e., equiprobable and independent) are treated as categorical, since their generating mechanism is invariant under any permutation of labels, and thus admits no intrinsic order or structure.

We further distinguish:

1. **Determining parameters**, which suffices for the occurrence of an event.
2. **Non-determining parameters**, which is not necessary for it.

In general, a fact corresponds to a single parameter. However, if several facts are ontologically interconnected — that is, there is a causal, grounding, or similar relation among them — they count as a single determining parameter.

We can thus outline a procedure for identifying and classifying relevant parameters. Given an event in a world  $w$ , consider the atomic facts whose occurrence is sufficient for the event's occurrence. If such facts are mutually dependent, or only jointly sufficient, they are grouped together. Each resulting group constitutes the value of a determining parameter. As a heuristic to classify these parameters, consider whether the labels of the parameter values can be permuted without altering other facts of the world. If this is possible, the parameter is likely categorical; otherwise, it is likely ordered.

### 2.3. The Modal and Normic Metrics

Two distinct metrics are defined on this space. The first,  $D$ , measures *modal distance* between worlds; the second,  $N$ , measure their normalcy:

$$\begin{aligned} D: W \times W &\rightarrow \mathbb{R}_{\geq 0} \\ N: W &\rightarrow \mathbb{R}_{\geq 0} \end{aligned}$$

The modal metric  $D$  is computed as:

$$D(w, v): \sum_{i=1}^n \alpha_i \cdot \left( \sum_{i=1}^n \alpha_i \cdot d_i(\pi_{td}(w), \pi_{td}(v)) \right) \quad (D)$$

Where:

- weight  $\alpha_i = 1$  for determining parameters and 0 otherwise.
- $d_i$  measures the difference between corresponding parameter values as  $|x - y|$  for ordered parameters,  $\varepsilon$  for categorical random parameters, and  $\epsilon$  for categorical non-random ones.
- The function  $\pi_{td}$  maps a world to a tuple by assigning each parameter the value that it had at the *time of determination*  $td$ . For instance, if a bomb is set to detonate at  $t$  and nothing can prevent its explosion at  $t + 99$ , the time of determination is  $t$ .

The normic metric  $N$  is defined as:

$$N(w) = \frac{\sum_{i=1}^n \alpha_i \cdot S_i(\pi_{td}(w))}{\sum_{i=1}^n \alpha_i} \quad (N)$$

Where:

- $S_i$  measures the extent to which a parameter value calls for an explanation. Parameters that are necessary or random receive the value 1, while the others are weighted by  $1 - g_i(p_i)$ , with  $0 < g_i(p_i) < 1$ .

For practical purposes, when discussing epistemic risk, it may be useful to transform the assessment of normalcy into a distance from the actual world  $w_a$ . In this case, the distance is inversely proportional to normalcy:

$$D(w_a, v)_N = \frac{1}{N(v)}$$

These metrics are defined relative to a target-event, which fixes which parameters count as determining. The determining parameters are those that are sufficient for the non-occurrence of the target-event. For example, to analyse the risk of a plane crashing,  $D$  measures the distance between the actual world, where the plane lands safely, and a counterfactual world in which it crashes. The determining parameters are those whose values, if varied, would have made the plane land safely instead. For instance, if both the absence of a storm and stricter departure checks would have been sufficient for a safe landing, and there is no connection between the two, we identify two distinct determining parameters.

In the case of epistemic risk assessment, the determining parameters are those that determine the belief's truth value and that makes the agent form a belief through a certain method at a certain time.

The standard metric axioms – identity, symmetry, and the triangle inequality – are not essential:  $D$  and  $N$  may be treated as quasi-metrics or pseudo-metrics without affecting our reasoning.

#### 2.4. Representing Risky Beliefs

Given a belief  $B_i$ , let  $R \subseteq W$  be the set of relevant worlds – those in which the same agent forms the same belief through the same method at the same time. We then define two subspaces:

- $R_m$ : the modally very close relevant worlds.
- $R_n$ : the most normal relevant worlds.

These subspaces are delimited by a fixed threshold  $\delta$ :  $R_m(\delta) = \{w : D(w, v) \leq \delta\}$ ,  $R_n(\delta) = \{w : N(w) \geq \delta\}$ . Worlds that depend on a single determining random parameter – i.e.,  $D(w, v) = \varepsilon$  – are intuitively very close. In fact, if the occurrence of an event is random, such occurrence does not require any prior difference and therefore is ontologically simple. Hence, we posit that  $\delta = \varepsilon$  for  $D$ . For  $N$ , we posit  $\delta = 1$ , which is the maximum possible value of  $N(w)$ .

Alternatively, we may set  $\delta = \epsilon$ , with  $\varepsilon < \epsilon$ , for  $D$ , extending the reasoning to non-random categorical parameters. Treating only worlds that differ by only a single parameter as very close remains a methodologically conservative choice: it yields an extremely permissive safety condition, making it difficult to challenge determinism on epistemic risk grounds.

Let  $F(B)$  denote the set of relevant worlds in which the belief is false – i.e., those in which the fact  $b$  that would make it true does not exist – and  $F_m(B)$ ,  $F_n(B)$  its modal and normic subsets:

$$\begin{aligned} F(B) &:= w \in R : \neg In(b, w) \\ F_m(B) &:= w \in R_m : \neg In(b, w) \\ F_n(B) &:= w \in R_n : \neg In(b, w) \end{aligned}$$

The measure function  $\mu$  computes the relative portion of such worlds:  $\mu(F_m) = \frac{|F_m|}{|R_m|}$  and  $\mu(F_n) = \frac{|F_n|}{|R_n|}$ . If  $R$  is uncountable,  $\mu$  can be extended as a normalised Borel measure over the metric space. Accordingly, cardinality ratios are reinterpreted as relative densities. Belief  $B$  is then risky if a majority of relevant worlds falsify it:

$$Risky_M(B) = \begin{cases} 1, & \text{if } \mu(F_m) \geq 0,5 \\ 0, & \text{if } \mu(F_m) < 0,5 \end{cases}$$

$$Risky_N(B) = \begin{cases} 1, & \text{if } \mu(F_n) \geq 0,5 \\ 0, & \text{if } \mu(F_n) < 0,5 \end{cases}$$

### 3. Philosophical Grounds of the Metric Framework

#### 3.1. Philosophical Grounds of the Modal Metric

Many theorists maintain that risk correlates, not with probability, but with modal distance. The guiding definition is this:

**Modal risk** — A belief  $B$  formed by an agent  $A$  through the method  $M$  at time  $t$  is risky if the worlds where  $B$  is false are a large enough portion of the very close possible worlds where  $A$  forms  $B$  through  $M$  at  $t$ . (cf. Pritchard, 2016)

The notion of modal closeness originates with Lewis (1973, §2.4), who employed it as a measure of the similarity to analyse counterfactuals; Williamson (2000, §5.3) later adapted it to capture epistemic safety and risk. On this view, risk increases with similarity: the closer the worlds in which a negative event occurs, the greater the risk. Yet, several examples indicate that similarity and risk do not always coincide. Our intuition about what counts as “close” sometimes diverges from our intuitions about what counts as “risky”.

**Ex. 1** Imagine a ball resting on the tip of a cone and another on the flat top of a frustum. The risk of falling is evidently greater in the first case, since a smaller displacement suffices to make the ball fall. The “falling” worlds are thus closer to the “stable” ones. Now suppose the ball’s position is determined by a die roll: an even number keeps it balanced (on either surfaces), an odd number displaces it so that it falls. Although the cone requires a smaller displacement, intuition holds that both cases now involve equal risk.

**Ex. 2** Consider two balls,  $A$  and  $B$ . For  $A$ , the ball falls immediately if an odd number is rolled at  $t$ . For  $B$ , it falls at  $t$  if an odd number was rolled at  $t-99$ . Thus, the “falling” and “stable” worlds of  $A$  coincide until  $t$ , whereas those of  $B$  coincide until  $t-99$  and are very different at  $t$ . Nonetheless, the intuitive degree of risk at  $t$  is the same for both.

**Ex. 3** Suppose ball  $A$  falls if a coin lands heads; ball  $B$  falls if, at the origin of the universe, a physical law  $x$  rather than  $y$  is actualised, where which one is actualised is random. Altering a fundamental law constitutes a far greater difference between worlds than altering a coin’s outcome: a law may comprise many different facts. Yet both situations seem to involve the same risk.

The modal metric  $D$  explains these divergences:

$$D(w, v) = \sum_{i=1}^n \alpha_i \cdot \left( \sum_{i=1}^n \alpha_i \cdot d_i(\pi_{td}(w), \pi_{td}(v)) \right) \quad (D)$$

**Ex.1.** Without the die roll, displacement is the sole determining parameter, so risk correlates with its continuous variation:

$$D(w_{fall}, v_{stable}) \propto |d_1(p_1(w_{fall}) - p_1(v_{stable}))|$$

When the die roll is introduced, it becomes the only determining parameter. Since this parameter is categorical, the modal distance between “falling” and “stable” worlds is constant:

$$D(w_{fall}, v_{stable}) = \varepsilon$$

**Ex. 2.** The function  $\pi_{td}$  regresses the parameters values at the time of determination. For ball  $A$ , that time is  $t$ ; for  $B$ , it is  $t-99$ . Hence, the modal distance at  $t$  is measured at  $t$  for  $A$  and at  $t-99$  for  $B$ , and subsequent divergence is irrelevant.

**Ex. 3.**  $D$  assigns no intrinsic weight to the metaphysical significance of parameters: it values only their number and whether they are determining or not, ordered or not. Both the coin toss and the initial-law event are categorical, and thus entail equal modal distance and equal risk.

The  $D$  metric emerges from the idea that risk depends on the number of determining parameters and their value at the time of determination. Thus, it succeeds where similarity-based accounts fail: it aligns modal closeness with the intuitive structure of risk. Rather than evaluating similarity abstractly,  $D$  specifies which differences are relevant.

### 3.2. Philosophical Grounds of the Normic Metric

Alongside the modal conception of risk, a second notion has recently gained prominence:

(NR) **Normic risk** — A belief  $B$  formed by an agent  $A$  through the method  $M$  at time  $t$  is risky if the worlds where  $B$  is false are a large enough portion of the most normal possible worlds where  $A$  forms  $B$  through  $M$  at  $t$ . (cf. Ebert, Smith & Durbach, 2020, §5)

According to Ebert, Smith and Durbach, the more an event calls for an explanation, the less normal it is. But what does it mean for something to call for an explanation?

Philosophers of science have proposed three major models of explanation:  $x$  explains  $y$  only if  $y$  is derived from  $x$  through...

**Deductive-Nomological** — ...a law-like generalisation. (Hempel, 1965, p. 365)

**Causal-Mechanical** — ...causal processes and interactions. (Salmon, 1984, p. 9)

**Unificationist** — ...an argumentative scheme used to derive many other facts. (Kitcher & Salmon, 1989)

Although identifying all the possible facts that do not call for an explanation exceeds the scope of this paper, there is a broad agreement on two limiting classes of facts that do not call for an explanation and therefore possess maximal normalcy.

First, random facts do not call for an explanation: they are by definition equiprobable and independent, and hence there are no facts from which they can be derived through a law-like generalisations, causal mechanisms, or unificatory schemes.

Second, necessary facts also fail to call for an explanation. Because they cannot fail to obtain, their existence is self-explanatory. Likewise, if  $y$  is necessitated by  $x$ , no further explanation of  $y$  is required beyond that of  $x$ . Although Baras (2020) concludes that whether a necessary fact can call for an explanation depends on how one understands this expression, he also recognises that almost all available accounts imply that a necessary fact cannot call for an explanation.

In other words, random facts do not call for an explanation because they cannot have one, necessary facts because they cannot lack one. The normic metric operationalises these insights by assigning the maximum normalcy value to determining parameters corresponding to random or necessary facts:

$$N(w) = \frac{\sum_{i=1}^n \alpha_i \cdot S_i(\pi_{td}(w))}{\sum_{i=1}^n \alpha_i} \quad (\text{N})$$

$N$  assigns to each world the average of the degree of normalcy of the values of its parameters. As a consequence, as the number of relevant parameters increases, the distance  $D$  increases, while normalcy  $N$  decreases proportionally: the more ways there are to escape a negative event, the lower its risk; the more coincidences required for the event to occur, the lower its normalcy.

### 3.3. Philosophical Grounds of the Representation of Risky Beliefs

Both definitions of risk — the modal and the normic — state that a belief is risky when the portion of very close or most normal worlds in which it is false is “large enough.” Yet what counts as “large enough”?

Williamson (2000) and Pritchard (2007, p. 292) maintain that knowledge requires safety: a belief constitutes knowledge only if it could not easily have been false. In their formulations, safety demands truth in *all* close worlds. Under this strict interpretation, a belief becomes risky — and hence fails to amount to knowledge — if it is false in even a single very close world.

Such an absolute standard captures an important intuition about especially important beliefs (cf. Broncano-Berrocal, 2013, 28), but it may be too restrictive for epistemological analysis in general. Coffman (2007, 390) proposes a more permissive and arguably tractable criterion: a belief is safe if it is true in at least half close worlds. Because this condition sets a higher evidential bar for those who wish to argue that determinism undermines knowledge, it is methodologically conservative. Adopting it ensures that the subsequent argument does not rely on an artificially narrow notion of risk:

$$\begin{aligned} Risky_M(B) &= \begin{cases} 1, & \text{if } \mu(F_m) \geq 0,5 \\ 0, & \text{if } \mu(F_m) < 0,5 \end{cases} \\ Risky_N(B) &= \begin{cases} 1, & \text{if } \mu(F_n) \geq 0,5 \\ 0, & \text{if } \mu(F_n) < 0,5 \end{cases} \end{aligned}$$

#### 4. Axioms of the ETA

The conception of risk presented in §2 provides the framework within which the axioms of our ETA can be philosophically and formally justified. The ETA is formulated in a three-sorted FOL in which several notions relevant to epistemic risk are treated as predicates:

Category	Variables/Symbols	Description
Sort $VB$	$vB_1, vB_2, vB_3, \dots$	Vulnerable beliefs
Sort $W$	$w, v, u, t, \dots$	Possible worlds
Sort $\Phi$	$\varphi_w, \varphi_v, \varphi_u, \dots$	Possible pasts
$K$	$K(vB)$	The vulnerable belief $vB$ constitutes knowledge
$RanNec$	$RanNec(\varphi)$	The past $\varphi$ is a random (or necessary) parameter
$DepOnly$	$DepOnly(w, v, \varphi)$	Whether $w$ or $v$ actualised depends only on the past $\varphi$
$Close$	$Close(w, v)$	The worlds $w$ and $v$ are very close
$Risky$	$Risky(vB)$	The vulnerable belief $b$ is risky

Here, *vulnerable beliefs* are those beliefs that are false in at least half of the relevant possible worlds – i.e., those satisfying  $\mu(F) \geq 0,5$  – and the *past* of a world is the set of all facts obtaining prior to a certain instant  $t$ .

Standard FOL quantifiers, connectives, syntax, semantics, and inference rules apply.

Within this framework, we introduce five axioms for our ETA:

$$\forall vB (\forall w \forall v \text{Close}(w, v)) \rightarrow \text{Risky}(vB) \quad (1)$$

$$\exists vB K(vB) \quad (2)$$

$$\forall \varphi \text{RanNec}(\varphi) \quad (3)$$

$$\forall w \forall v \forall \varphi (\text{DepOnly}(w, v, \varphi) \wedge \text{RanNec}(\varphi)) \rightarrow \text{Close}(w, v) \quad (4)$$

$$\forall vB \text{ Risky}(vB) \rightarrow \neg K(vB) \quad (5)$$

We used Mace4 (ver. LADR-2009-11A) to show their consistency and their mutual independence. The source code is archived at (Montagner, 2025).

## 5. Philosophical Grounds of the Axioms

### 5.1. Philosophical Grounds of the Axiom 1

A belief is defined as *vulnerable* if it is false in at least half of the relevant possible worlds, and as *risky* if it is false in at least half of the very close relevant possible worlds. The first axiom states that, if all worlds were very close, then any vulnerable belief would also be risky:

$$\forall vB (\forall w \forall v \text{ Close}(w, v)) \rightarrow \text{Risky}(vB) \quad (1)$$

This axiom follows immediately from the definitions and requires no additional justification. Note that the predicate *Close* subsumes the  $N$  metric via the equivalence described in §2.3: maximally normal worlds are treated as maximally close to the actual world when the  $N$  metric is applied.

### 5.2. Philosophical Grounds of the Axiom 2

Given a possibilist reading of the quantifiers, the second axiom states that vulnerable beliefs can still constitute knowledge, even though they are false in most relevant possible worlds:

$$\exists vB K(vB) \quad (2)$$

Consider a clock so fragile that, in 90 percent of possible worlds, it is broken at the moment one consults it. A belief about the time formed through this clock is thus vulnerable: it is false in the majority of relevant worlds. Yet if, in the actual world, one has good reasons to believe that the clock is functioning, there is no obstacle to that belief counting as knowledge. The explanation is straightforward: in this context, the worlds in which the clock is broken are not close or normal. A majority of falsifying worlds is not enough to undermine knowledge without a specific modal topology.

### 5.3. Philosophical Grounds of the Axiom 3

Assume that multiple pasts are possible. In fact, if this were not so, counterfactual claims as “Napoleon could have won at Waterloo” would be meaningless. Axiom 3 states that the *RanNec* predicate holds for every possible past:

$$\forall \varphi \text{ RanNec}(\varphi) \quad (3)$$

This predicate can be interpreted in two different ways, depending on the way one distributes probabilities across possible pasts.

The first interpretation — naturally congenial to the determinist — holds that the past is necessary — i.e., common to all accessible worlds. Other worlds, with different equally necessary pasts, are inaccessible from ours. Accordingly, the probability of the actual past among all possible pasts equals unity:

$$P(\varphi_a) = 1$$

The alternative interpretation, instead, allows to assume, by a *principle of indifference*, that all possible pasts are equiprobable:

$$P(\varphi_a) = \frac{1}{|\Phi|}$$

In that case, since, as mentioned, no fact precedes the actualisation of a past, nothing can influence its actualisation. Therefore, the possible pasts are independent, and we can consider the occurrence of the actual past not just equiprobable, but random.

Why not a non-uniform probability distribution instead? One motivation comes from cosmology: Norton's (2021, §6) principle of mediocrity in infinite lottery logic offers a similar resolution to the measure problem in multiverse models and allows for uniform probability distribution even in uncountable sets.

Another justification emerges from a propensity interpretation of probability. As Strevens (1998) argues, asymmetries in probability reflect physical asymmetries. But the past includes all prior temporal and even atemporal facts: the latter, precisely because they are atemporal, already obtain before any  $t$ , as is evident if we assume that time emerges from them. Hence, the past admits no antecedent condition: nothing precedes its actualisation that could constitute an asymmetry.

As we will see below, these two conceptions of the past (randomness and necessity) are merged into a single predicate since the conclusion follows from both in an equivalent way.

#### 5.4. Philosophical Grounds of the Axiom 4

Under determinism, every fact is fully determined by its past. In accordance with contemporary physics, present facts do not depend on any particular past facts, but on the past as a whole. Specifically, all present facts depend on the Hilbert space state vector at the immediately preceding instant, which describes that instant as a non-decomposable whole composed of non-factorisable, entangled states. In other words, the global state cannot be reduced to local states. Consequently, the past constitutes a single parameter.

Thus, we may define *determinism* as the thesis that the only determining parameter of any fact is the occurrence of a specific past. Although the determinists might propose a different conception of their view — denying that past facts constitute a whole — this move amounts to abandoning a scientifically robust notion of determinism.

Axiom 4 states that if the occurrence of one world rather than another depends solely on which past actualised, and this past constitutes a random or necessary parameter, then those worlds are very close:

$$\forall w \forall v \forall \varphi (DepOnly(w, v, \varphi) \wedge RanNec(\varphi)) \rightarrow Close(w, v) \quad (4)$$

Like axiom 1, this one follows automatically from our conception of risk:

1. Applying the  $N$  metric, if the occurrence of a word depends solely on a random or necessary parameter, then that world is maximally normal, and hence very close to the actual world;
2. Applying the  $D$  metric, if two worlds depend solely on a random parameter, their distance is  $\varepsilon$ , and given the threshold  $\delta = \varepsilon$ , they are very close;
3. Instead, if two worlds depend solely on a necessary categorical parameter, their distance is  $\varepsilon$ . Since each possible past constitutes a complete and discrete configuration of the world, the values of the "actual past" parameter lack a natural ordering: their labels are mutually permutable without altering present facts. Hence, it qualifies a categorical parameter. Hence, if  $\delta = \varepsilon$ , the two worlds are very close even if the past is interpreted as necessary.

To refute axiom 4, the determinist would need to assume only the metric  $D$ , the necessity of the past, and a risk threshold  $\delta < \varepsilon$ . However, this would constitute an untenable permissive epistemic standard: a divergence in a single categorical parameter would suffice to render falsifying worlds distant, thereby making one's belief safe. Such a conception of safety contrasts sharply with the restrictive interpretations advanced by Williamson and Pritchard.

#### 5.5. Philosophical Grounds of the Axiom 5

Axiom 5 expresses a commitment to an anti-risk epistemology: a (vulnerable) belief qualifies as knowledge only if it is non-risky.

$$\forall vB \text{ Risky}(vB) \rightarrow \neg K(vB) \quad (5)$$

The assumption reflects a long tradition that prioritises the avoidance of error over the acquisition of truth. William James (1979, §7) famously identified two epistemic precepts:

1. Believe truth!
2. Shun error!

An epistemology is anti-risk when it accords priority to the second. It treats the avoidance of falsehoods as an epistemic good even when this comes at the cost of forgoing some truths.

This attitude has deep historical roots. Clifford (1999) gave it moral form in his maxim: “It is wrong always, everywhere, and for anyone, to believe anything upon insufficient evidence”. Wald (1945) formulated the maximin rule: one should select the option whose worst possible outcome is the best available. Rawls (1974) later defended risk-avoidance, proposing the minimax as a rational principle in political philosophy.

Rawls’s reasoning transposes naturally into epistemology. Thus, epistemic risk aversion can be defended on at least four grounds:

1. **Robustness.** Risk-taking lowers the reliability of belief-forming processes. If baseline reliability is low, a risk-tolerant epistemology could generate a mostly false belief system; by contrast, risk-avoidance increases reliability across all baselines.
2. **Asymmetry of loss.** Adding a falsehood to one’s belief system worsens it, whereas failing to add a truth merely leaves it unchanged. Epistemic error carries a strictly greater disvalue than epistemic omission.
3. **Practical tractability.** Given falsificationism (Popper, 2002, ch. 4) and underdetermination (Kyle, 2023), falsehoods are typically easier to identify and eliminate than truth to confirm. A strategy that privileges error-avoidance is thus more implementable in practice.
4. **Intuitive coherence.** As Pettigrew (2015, pp. 257-261) observes, risk-tolerant frameworks license counterintuitive strategies — such as believing contradictory propositions simultaneously or assigning greater credence to one of several equiprobable alternatives.

A related Rawlsian defence of epistemic minimax has been advanced by Alfano et al. (forthcoming), who argue that risk-averse strategies promote certain epistemic goods in some social groups.

A full defense of anti-risk epistemology would require volumes. Yet for present purposes, the considerations above suffice to justify axiom 5 as rational and principled within the framework of our ETA.

## 6. The ETA Against Determinism from Epistemic Risk

We have defined *determinism* as the thesis that every world depends solely on a past. The axioms were translated into Prover9 (ver. LADR-2009-11A) to verify that the negation of determinism — that is, the existence of worlds that do not depend on any past — follows. Then, we developed in Lean 4 (v4.26.0-rc2) a proof of the derivation. All source code is archived (Montagner, 2025). We present here an intuitive sketch proof.

Let us assume, for *reductio*, that determinism is true — i.e., that, for every fact, which past actualised is the only relevant parameter:

$$\forall w \forall v \exists \varphi \text{ DepOnly}(w, v, \varphi) \quad (6)$$

By axiom 3, every past is a random (or necessary) parameter. Thus, given (6), every world depends solely on a random (or necessary) parameter:

$$\forall w \forall v \exists \varphi \text{ DepOnly}(w, v, \varphi) \wedge \text{RanNec}(\varphi) \quad (7)$$

(7), given axiom 4, causes a modal collapse: every world is very close to the actual one:

$$\forall w \forall v \text{ Close}(w, v) \quad (8)$$

From (8), given axiom 1, follows that every vulnerable belief is risky:

$$\forall v B \text{ Risky}(vB) \quad (9)$$

Therefore, by axiom 5, no vulnerable belief can constitute knowledge:

$$\forall v B \neg K(vB) \quad (10)$$

But (10) contradicts axiom 2. Therefore, our assumption must be discarded, and determinism is transcendentally rejected.

Libertarianism, instead, avoids this *reductio*. A libertarian agent is free to believe or disbelieve, act or procrastinate, apply one method or another. Each of these undetermined but non-random choices constitutes a new determining parameter, as they are independent and each would have been sufficient to avoid a relevant world in which one believes a falsehood through a specific method in a specific time. Consequently, it is never the case that the target-events depend on a single categorical parameter.

## 7. Concluding Remarks

It is difficult to imagine an alternative formalisation of the intuition captured by Gisin's remark.

Gisin's intuition does not claim that determinism necessarily entails an epistemically undesirable context, but only that it might. It therefore calls for the notion of epistemic risk: it must assume that knowledge is possible at least in libertarianism (axiom 2), that risk undermines knowledge (axiom 5), and define risk in terms of modal closeness or normalcy (axioms 1 and 4).

Yet if modal closeness is understood merely as a measure of similarity, the argument cannot proceed: under either determinism or libertarianism, the worlds remain equally similar. A different metric is therefore required. However, any adequate one must rest on an unbiased analysis of the concepts of modal closeness and normalcy. Given the intuitive and widespread view that necessary and random facts are maximally normal, any adequate metric should account for this, as we did in §2.

Consequently, the argument can succeed only by showing that the determining fact in determinism — namely, the past — is itself necessary or random (axiom 3).

Thus, all steps of the argument are forced and necessary.

The argument does not purport to demonstrate that libertarianism is the only alternative to determinism, nor does it exclude all conceivable forms of determinism. However, our formalisation suffices to establish that a scientifically grounded form of determinism entails a knowledge-undermining risk for certain classes of intuitively knowable propositions.

## References

- Alfano, M., Ferreira, M, Reimann, R., Cheong, M., & Klein, C. (Forthcoming). Epistemic Minimax and Related Principles in the Contemporary Epistemic Environment. In M. Popa-Wyatt (ed.), *Misinformation and Other Epistemic Pathologies*. Cambridge University Press.
- Balfour, A. (1879). *A Defence of Philosophical Doubt: Being an Essay on the Foundations of Belief*. Longmans, Green & Co.
- Baras, D. (2020). How Can Necessary Facts Call for Explanation. *Synthese*, 198(12): 11607-11624. <https://doi.org/10.1007/s11229-020-02817-7>.
- Bernstein, M. (1988). Justification and Determinism: an Exchange. *Monist*. 71(3): 358-364. <https://doi.org/10.5840/monist198871322>.
- Boyle Jr., M.J., Grisez, G., & Tollefsen, O. (1972). Determinism, Freedom and Self-Referential Arguments. *Review of Metaphysics*, XXVI(1). <https://www.jstor.org/stable/20126164>.
- Broncano-Berrocal, F. (2013). *Luck and the Control Theory of Knowledge*. <https://api.semanticscholar.org/CorpusID:170453651>.

- Chevarie-Cossette, S.P. (2019). Is Free Will Scepticism Self-Defeating? *European Journal of Analytic Philosophy*, 15(2), 55-78. <https://doi.org/10.31820/ejap.15.2.3>.
- Clifford, W.K. (1999). The Ethics of Belief. In T. Madigan (ed.), *The Ethics of Belief and Other Essays* (pp. 70-96). Prometheus.
- Coffman, E.J. (2007). Thinking About Luck. *Synthese*, 158, 385-398. <https://doi.org/10.1007/s11229-006-9046-8>.
- Ebert, P.A., Smith, M., & Durbach, I. (2020). Varieties of Risk. *Philosophy and Phenomenological Research*, 101(2), 432-455. <https://doi.org/10.1111/phpr.12598>.
- Gisin, N. (2014). *Quantum Chance: Nonlocality, Teleportation and Other Quantum Marvels*. Springer International Publishing.
- Haldane, J. B. (1929). *Possible Worlds, and Other Essays*. Chatto and Windus.
- Hasker, W. (1973). The Transcendental Refutation of Determinism. *Southern Journal of Philosophy*, 11(3), 175-183. <https://doi.org/10.1111/j.2041-6962.1973.tb01133.x>.
- Hempel, C. G. (1965). *Aspects of Scientific Explanation and other Essays in the Philosophy of Science*. Free Press.
- Honderich, T. (1988). *A Theory of Determinism: the Mind, Neuroscience, and Life-Hopes*. Clarendon Press.
- James, W. (1979). The Will to Believe. In F. Burkhardt, F. Bowers, and I. K. Skrupskelis (eds.), *The Will to Believe and Other Essays in Popular Philosophy*. Harvard University Press.
- Joad, C.E. (1933). *A Guide to Modern Thought*. Faber & Faber.
- Jordan, J.N. (1969). Determinism's Dilemma. *Review of Metaphysics*, 23(1), 48-66.
- Kitcher, P., & Salmon, W. 1987. *Scientific Explanation*. Minneapolis, University of Minnesota Press.
- Kyle, S. (2023). Underdetermination of Scientific Theory. In E. N. Zalta and U. Nodelman (eds.), *The Stanford Encyclopedia of Philosophy* (Summer 2023 Edition). <https://plato.stanford.edu/archives/sum2023/entries/scientific-underdetermination/>.
- Latham, A.J. (2019). The Conceptual Impossibility of Free Will Error Theory. *European Journal of Analytic Philosophy*, 15(2), 99-120. <https://doi.org/10.31820/ejap.15.2.5>.
- Lewis, C.S. (1947). *Miracles*. Collins.
- Lewis, D.K. (1973). *Counterfactuals*. Blackwell Publishers.
- Lockie, R. (2018). *Free Will and Epistemology: A Defence of the Transcendental Argument for Freedom*. Bloomsbury.
- Lucas, J.R. (1970). *The Freedom of the Will*. Clarendon Press.
- Mabbott, J.D. (1966). *An Introduction to Ethics*. Hutchinson.
- Magno, J.A. (1984). Beyond the Self-Referential Critique of Determinism. *Thomist: A Speculative Quarterly Review*, 48(1), 74-78. <https://doi.org/10.1353/tho.1984.0057>.
- Malcolm, N. (1968). The Conceivability of Mechanism. *The Philosophical Review*, 77(2), 45-72. <https://doi.org/10.2307/2183182>.
- Mascall, E.L. (1965). *Christian Theology and Natural Science: Some Questions in Their Relations*. Archon Books.
- Montagner, A. (2025). Proofs for 'Epistemic Risk and the Transcendental Case Against Determinism'. <https://doi.org/10.5281/zenodo.17842159>.
- Moreland, J.P. (1987). *Scaling the Secular City: a Defense of Christianity*. Baker Publishing Group.
- Navarro, J. (2021). Epistemic Luck and Epistemic Risk. *Erkenntnis*, 88(3), 1-22. <https://doi.org/10.1007/s10670-021-00387-9>.
- Norton, D. J. (2021). Eternal Inflation: When Probabilities Fail. *Synthese*, 198, 3853-3875. <https://doi.org/10.1007/s11229-018-1734-7>.
- Pettigrew, R. (2015). Jamesian Epistemology Formalised: An Explication of 'The Will to Believe.' *Episteme*, 13(3), 253-268. <https://doi.org/10.1017/epi.2015.44>.
- Plantinga, A. (1993). *Warrant and Proper Function*. Oxford University Press.
- Popper, K.R. (2002). *The Logic of Scientific Discovery*. Routledge.
- Popper, K.R., & J. C. Eccles. (1984). *The Self and Its Brain*. Routledge.
- Pritchard, D. (2007). Anti-Luck Epistemology. *Synthese*, 158, 277-298. <https://doi.org/10.1007/s11229-006-9039-7>.
- Pritchard, D. (2015). Risk. *Metaphilosophy*, 46(3), 436-461. <https://doi.org/10.1111/meta.12142>
- Pritchard, D. (2016). Epistemic Risk. *Journal of Philosophy*, 113(11), 550-571. <https://doi.org/10.5840/jphil20161131137>.

- Rawls, J. (1974). Some Reasons for the Maximin Criterion. *Papers and Proceedings of the Eighty-sixth Annual Meeting of the American Economic Association*, 64(2), 141-146.
- Ripley, C. (1972). Why Determinism Cannot Be True. *Dialogue*, 11(1), 59-68. <https://doi.org/10.1017/s0012217300028389>.
- Salmon, W. (1984). *Scientific Explanation and the Causal Structure of the World*. Princeton University Press
- Slagle, J. (2016). *The Epistemological Skyhook: Determinism, Naturalism, and Self-defeat*. Routledge.
- Snyder, A.A. (1972). The Paradox of Determinism. *American Philosophical Quarterly*, 9(4), 353-356. <http://www.jstor.org/stable/20009464>.
- Steward, H. (2020). Free Will and External Reality: Two Scepticisms Compared. *Proceedings of the Aristotelian Society*, 120(1), 1-20. <https://doi.org/10.1093/arisoc/aoaa001>.
- Strevens, M. (1998). Inferring probabilities from symmetries. *Noûs*, 32(2), 231-246. <https://www.jstor.org/stable/2671966>
- Taylor, R. (1963). *Metaphysics*. Prentice Hall.
- Wald, A. (1945). Statistical Decision Functions which Minimize the Maximum Risk. *Annals of Mathematics*, 46(2), 265-280. <https://www.jstor.org/stable/1969022>.
- Walker, M.T. (2001). Against One Form of Judgment-Determinism. *International Journal of Philosophical Studies*, 9(2), 199 – 227. <https://doi.org/10.1080/09672550110035907>.
- Wick, W. (1964). Truth's Debt to Freedom. *Mind*, LXXIII(292), 527-537. <https://www.jstor.org/stable/2252100>.
- Williamson, T. (2000). *Knowledge and Its Limits*. Oxford University Press.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.