

Article

Not peer-reviewed version

---

# Epistemic Risk and the Transcendental Case Against Determinism

---

[Alessio Montagner](#)\*

Posted Date: 31 October 2025

doi: 10.20944/preprints202510.2441.v1

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

This paper advances a new epistemic transcendental argument (ETA) against determinism. I argue that determinism entails a higher degree of epistemic risk than libertarianism. To make this claim precise, I formalize risk through a metric space  $W$  and represent two theories of epistemic risk as alternative metrics,  $D$  and  $N$ . By exploring the concepts of modal closeness and normalcy, I show that these metrics are superior to traditional conceptions in accounting for intuitions about risk and ensuring objective assessment. Given some additional assumptions about knowledge and the past, I construct an ETA grounded in epistemic risk aversion: under determinism, all possible worlds count as modally close and normal, making many beliefs epistemically risky. This proposal departs from the Epicurean tradition and opens a new route against determinism.

**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<sup>1</sup>. 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.<sup>2</sup>

<sup>1</sup> 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).

<sup>2</sup> 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

Of course, such risk is not exclusive to determinism: since a course of action could occur either because I am determined to carry it out or because I freely choose 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 Gidin'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 what follows, I formalise the notion of epistemic risk and assess whether, under determinism, the conditions for knowledge fail for at least one class of propositions generally regarded as knowable.

## 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: risk and accessibility might not be proportional-

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. The formal system is a multi-sorted FOL with set-theoretic and arithmetic operators, whose general structure is summarised in *Table 1*.

Table 1.

Sorts	Variables/Symbols	Description
$W$	$w, v, u, \dots$	Possible worlds
$S(W)$	$R_m, R_n, F, \dots$	Sets of worlds
$B$	$B_1, B_2, B_3, \dots$	Beliefs
$F$	$a, b, c, \dots$	Facts

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.

$P$	$p_1, p_2, p_3, \dots$	Parameters
$\Phi$	$\varphi_w, \varphi_v, \varphi_w, \dots$	Possible pasts
$\mathbb{R}$	$1, 2, \dots, \delta, \varepsilon$	Real numbers

Predicates	Variables/Symbols	Description
Existence	$In(a, w)$	Membership of a fact in a domain
Knowledge	$K(B_i)$	Belief $B_i$ constitutes knowledge
Vulnerability	$V(B_i)$	Belief $B_i$ is vulnerable
Riskiness	$Risky_M(B_i), Risky_N(B_i)$	Belief $B_i$ is risky
Probability	$P(\varphi_w)$	Probability of occurrence of a past

Functions	Variables/Symbols	Description
Distance	$D: W \times W \rightarrow \mathbb{R}_{\geq 0}$	Metric for modal closeness
Normalcy	$N: W \rightarrow \mathbb{R}_{\geq 0}$	Metric for normalcy
Distance	$d_i: P \times P \rightarrow \mathbb{R}_{\geq 0}$	Calculates difference in parameter values
Explanation	$S_i: P \rightarrow \mathbb{R}_{[0,1]}$	How much worlds call for an explanation
Determination	$\alpha_i: P \rightarrow \mathbb{R}_{[0,1]}$	Whether a parameter is relevant
Time	$\pi_{td}: W \rightarrow W$	Regresses parameter values in time
Measure	$\mu: S(W) \times S(w) \rightarrow \mathbb{R}_{[0,1]}$	Calculates proportions between sets

Operations	Variables/Symbols	Description
Set-theoretic	$\{x \in A: \varphi(x)\}$	Standard operations to define sets
Arithmetic	$>, \geq, <, \leq, +, \cdot, -, \div, \dots$	Standard ordering and operations on $\mathbb{R}$
Logical	$\forall, \exists, \wedge, \vee, \rightarrow, \neg, \dots$	Standard FOL operators and quantifiers

Unless stated otherwise, standard FOL syntax, semantics, and inference rules apply.

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

$$w = \langle p_1(w), \dots, p_n(w) \rangle \quad (1)$$

Each parameter captures a dimension along which worlds may differ. Parameters can be of two kinds:

1. **Ordered parameters**, whose values belong to a naturally metrisable space  $(S_i, M_i)$ , representing continuous variations.
2. **Categorical parameters**, whose values belong to a discrete label space  $L_i = l_1, \dots, l_n$ .

The former models continuous magnitudes — e.g., the distance of a dart from a target's center — while the latter models discrete outcomes — e.g., heads or tails. Parameters whose values are purely random (i.e., equiprobable and independent) are treated as categorical, since the numerical labels do not correspond to any underlying order or continuity in the generating process of the outcomes.

We further distinguish:

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

Determining parameters thus correspond to sufficient conditions in implications of the form  $x \rightarrow y$ . Consequently, a single parameter can be composed of many different facts.

## 2.2. 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:

$$D: W \times W \rightarrow \mathbb{R}_{\geq 0} \quad (2)$$

$$N: W \rightarrow \mathbb{R}_{\geq 0} \quad (3)$$

The modal metric  $D$  is computed as:

$$D(w, v): \sum_{i=1}^n \alpha_i \cdot (\sum_{i=1}^n \alpha_i \cdot d_i(\pi_{td}(w), \pi_{td}(v))) \quad (4)$$

Where:

- weight  $\alpha_i = 1$  for determining parameters and 0 otherwise.
- $d_i$  measures the difference between corresponding parameter values, either as  $|x - y|$  for ordered parameters or as  $\varepsilon$  for categorical 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:

$$N(w) = \frac{\sum_{i=1}^n \alpha_i \cdot S_i(p_i(w))}{\sum_{i=1}^n \alpha_i} \quad (5)$$

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  $1 - g_i(p_i)$ .

These metrics are defined relative to a target event, which fixes which parameters count as determining. For example, to analyse the occurrence of a plane crash,  $D$  measures the distance between the actual world, where the plane lands safe, and a counterfactual world in which it crashes. In this case, the determining parameters are those whose variations made the crash obtain in the counterfactual world.

Both  $D$  and  $N$  satisfy the standard axioms of a metric — identity, symmetry, and the triangle inequality — though the argument remains valid if they are construed instead as quasi- or pseudo-metrics.

### 2.3. Representing Risk

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 differs only by a single determining categorical parameter — i.e.,  $D(w, v) = \varepsilon$  — are intuitively very close; hence, we posit that  $\delta = \varepsilon$  for  $D$ . For  $N$ , we posit  $\delta = 1$ , which is the maximum possible value of  $N(w)$ .

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. Two worlds in which the same agent forms the same true belief can be more or less distant depending on the parameters by virtue of which this state of affairs obtains.

Let  $F(B)$  denote the set of relevant worlds in which the belief is false, and  $F_m(B)$ ,  $F_n(B)$  its modal and normic subsets:

$$F(B) := w \in R: \neg In(b, w) \quad (6)$$

$$F_m(B) := w \in R_m: \neg In(b, w) \quad (7)$$

$$F_n(B) := w \in R_n: \neg In(b, w) \quad (8)$$

The measure function  $\mu$  computes the relative portion of such worlds:  $\mu(F_m) = \frac{|F_m|}{|R_m|}$  and  $\mu(F_n) = \frac{|F_n|}{|R_n|}$ .<sup>3</sup> Belief  $B$  is then risky if a majority of relevant worlds falsify it:

$$Risky_M(B) = \begin{cases} 1, & \text{if } \mu(F_m) \geq 0,5 \\ 0, & \text{if } \mu(F_m) < 0,5 \end{cases} \quad (9)$$

$$Risky_N(B) = \begin{cases} 1, & \text{if } \mu(F_n) \geq 0,5 \\ 0, & \text{if } \mu(F_n) < 0,5 \end{cases} \quad (10)$$

### 2.4. Epistemological and Metaphysical Axioms

We assume the following epistemological constraints:

$$K(B) \rightarrow (Risky(B) = 0) \quad (AR)$$

$$\exists w(V(B) \wedge K(B)) \quad (PK)$$

where  $K(B)$  expresses that  $B$  constitutes knowledge and  $V(B)$  that  $B$  is *vulnerable*:

$$V(B) \leftrightarrow \mu(F) \geq 0,5 \quad (11)$$

Finally, the nature of the past is specified by one of two assumptions:

$$P(\varphi_a) = 1 \quad (NP)$$

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

where  $P(\varphi_a)$  denotes the probability of the actual past among all possible pasts  $\Phi$ .

To evaluate their independence and consistency, consider the following models:

<sup>3</sup> If  $R$  is uncountable,  $\mu$  can be extended as a normalized Borel measure over the metric space. Accordingly, cardinality ratios are reinterpreted as relative densities.

Table 2.

	$M_{AR}$	$M_{PK}$	$M_{NP/RP}$	$M_S$
$M \models$	$K(B)$ $V(B)$ $Risky(B) = 1$ NR or RP by choice	$K(B)$ $Risky(B) = 0$ NR or RP by choice	$K(B)$ $V(B)$ $Risky(B) = 0$	$K(B)$ $V(B)$ $Risky(B) = 0$ NR or RP by choice
Axioms satisfied	All except AR	All except PK	All except NP/RP	All

These axioms are therefore independent and consistent.

### 3. 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. Yet both situations seem to involve the same risk.

The modal metric  $D$  explains these divergences. Recall its definition:

$$D(w, v): \sum_{i=1}^n \alpha_i \cdot (\sum_{i=1}^n \alpha_i \cdot d_i(\pi_{td}(w), \pi_{td}(v))) \quad (4)$$

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

$$D(w_{fall}, v_{stable}) \propto |d_i(p_1(w_{fall}) - p_1(v_{stable}))| \quad (12)$$

When the die roll is introduced, it becomes the only determining parameter. Accordingly,  $\alpha_i = 0$  for displacement and 1 for the die outcome. Since this parameter is categorical, the modal distance between “falling” and “stable” worlds is constant:

$$D(w_{fall}, v_{stable}) = \varepsilon \quad (13)$$

**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.** Equation (4) 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 metric therefore succeeds where similarity-based accounts fail: it aligns modal closeness with the intuitive structure of risk.

#### 4. Philosophical Ground 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*.<sup>4</sup>

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

$$N(w) = \frac{\sum_{i=1}^n \alpha_i \cdot S_i(p_i(w))}{\sum_{i=1}^n \alpha_i} \quad (5)$$

<sup>4</sup> 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.

## 5. Philosophical Grounds of the Definition of Risky Belief

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:

$$Risky_M(B) = \begin{cases} 1, & \text{if } \mu(F_m) \geq 0,5 \\ 0, & \text{if } \mu(F_m) < 0,5 \end{cases} \quad (9)$$

$$Risky_N(B) = \begin{cases} 1, & \text{if } \mu(F_n) \geq 0,5 \\ 0, & \text{if } \mu(F_n) < 0,5 \end{cases} \quad (10)$$

## 6. Philosophical Grounds of the AR Axiom

The AR axiom expresses a commitment to an anti-risk epistemology:

$$K(B) \rightarrow (Risky(B) = 0) \quad (\text{AR})$$

That is, a belief qualifies as knowledge only if it is non-risky. 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 AR as a rational and principled axiom within the framework of our ETA.

## 7. Philosophical Grounds of the PK Axiom

The PK axiom asserts that some beliefs — which we call vulnerable — can still constitute knowledge, even though they are false in most relevant possible worlds:

$$\exists w(V(B) \wedge K(B)) \quad (\text{PK})$$

$$V(B) \leftrightarrow \mu(F) \geq 0,5 \quad (11)$$

Consider a clock so fragile that, in 90 percent of possible worlds, it is broken at the moment I consult 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, I have 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.

## 8. Philosophical Grounds of the NP and RP Axioms

The NP and RP axioms concern two alternative attitudes toward the nature of the past. Let *the past* be defined as the set of all facts obtaining prior to a certain instant *t*. Assume that multiple pasts are possible:

$$\exists w \exists v(\varphi_w \neq \varphi_v) \quad (14)$$

If this were not so, counterfactual claims as "Napoleon could have won at Waterloo" would be meaningless.

Once multiple pasts are admitted, a natural question arises concerning their probability distribution. The first view — 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:

$$P(\varphi_a) = 1 \quad (\text{NP})$$

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

$$\forall w \forall v(P(\varphi_w) = P(\varphi_v)) \quad (15)$$

Why not a skewed 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, allowing 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.

From this it follows that the probability of the actual past among all possible pasts is:

$$P(\text{past}_a) = \frac{1}{|\phi|} \quad (\text{RP})$$

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.

## 9. The ETA Against Determinism from Epistemic Risk

The argument proceeds by *reductio*. Let us assume, for the sake of argument, that determinism is true. Under determinism, every fact is fully determined by its past; equivalently, the only determining parameter of any event is the occurrence of a specific past:

$$\forall w \forall v (\varphi_w = \varphi_v) \rightarrow (w = v) \quad (16)$$

We now examine the consequences of this assumption under the two conceptions of the past: RP and NP. Let us first assume RP.

Consider two arbitrary worlds  $w$  and  $v$  differing in some determining parameter  $p_i$ . By the contrapositive of (16), this difference depends solely on the occurrence of distinct pasts. But under RP, which past occurs is random. Hence, applying the metric  $D$ :

$$\forall w \forall v D(w, v) = \varepsilon \quad (17)$$

Similarly, applying the metric  $N$ :

$$\forall w (N(w) = 1) \quad (18)$$

Intuitively, because all differences depends on a single random parameter, every world is equally close to every other, and all are maximally normal.

Now consider a vulnerable belief  $B$ , made true by a fact  $b$  in the actual world. By (16)-(18), we have:

$$\forall w \forall v (In(b, w) \wedge \neg In(b, v) \rightarrow D(w, v) = \varepsilon) \quad (19)$$

$$\forall w \forall v (In(b, w) \wedge \neg In(b, v) \rightarrow N(v) = 1) \quad (20)$$

Since  $B$  is vulnerable, it is false in at least half of the relevant worlds. Given (19) and (20), these falsifying worlds are all both very close and maximally normal. Consequently:

$$\forall B (V(B) \rightarrow (Riskym(B) = 1)) \quad (21)$$

$$\forall B (V(B) \rightarrow (Riskyn(B) = 1)) \quad (22)$$

By AR:

$$\forall B (V(B) \rightarrow \neg K(B)) \quad (23)$$

But this contradicts PK. Therefore, the initial assumption must be rejected:

$$\exists w \exists v (\varphi_w = \varphi_v \wedge w \neq v) \quad (24)$$

Let us now consider the alternative assumption, NP. Since the  $N$  metric assigns maximal normalcy to both random and necessary parameters, the consequences parallel those under RP: equations (18), (20), and (22)-(24) still hold.

Libertarianism, instead, avoids this *reductio*. A libertarian agent is free to believe or disbelieve, act or procrastinate, apply one method or another. These choices, which constitute new determining parameters, are neither random nor necessitated by the past. Consequently, it is never the case that the only determining parameter is a random one; rather, the model ensures that  $D(w, v) > \varepsilon$  and  $N(w) < 1$ .

The argument does not imply that libertarianism is the only theory capable of satisfying (24); nevertheless, its goal is achieved: we have shown that the possibility of knowledge is undermined under determinism, and not under libertarianism.

## 10. Concluding Remarks

The reasoning behind our ETA can now be briefly summarised.

The starting intuition is simple: if determinism is true, one could be determined to hold false beliefs and lack any capacity for correction. Such a condition seems to undermine the very possibility of justified belief.

Yet this intuition is inconclusive. Determinism might equally determine one to believe only truths. What it truly expresses is an epistemic risk. This risk is generally analysed through the modal and normic structure of possible worlds. Determinism entails greater risk only if those worlds are, in some systematic way, closer or more normal than under libertarianism.

The traditional conception of modal closeness — as mere similarity — cannot support this conclusion, since the set of possible worlds remains identical under determinism and libertarianism alike. Hence, a reconception of closeness and normalcy is required. Our metric framework provided that reconception, and proved to be more effective than the traditional one in accounting for our intuitions about risk.

The ensuing analysis revealed that whenever the actualisation of one world rather than another depends solely on a random or necessary event, those worlds appear close and normal. Determinism, by definition, makes every fact depend on the past. But the past itself, under a principle of indifference echoing themes in multiverse cosmology, must be conceived as random. And if it is instead conceived as necessary, it follows the same under a normic interpretation.

From this point, the ETA follows straightforwardly. If determinism holds, every vulnerable belief is risky; by anti-risk epistemology, risky beliefs cannot amount to knowledge; yet it is intuitive that some vulnerable beliefs can be knowledge. Hence the assumption of determinism must be rejected.

Each step of this reasoning is forced, making it hard to see any alternative way to formalise the initial intuition.

If the argument succeeds, it reframes the traditional debate on freedom. The issue is no longer whether determinism threatens moral responsibility or assertability, but whether it undermines the modal structure that makes knowledge possible.

\*\*\*

I am an independent researcher working on the implications of recent epistemological trends for classical metaphysics. Feedback on this preprint is highly welcome:

Alessio Montagner

alessio@hermesambiente.it

\*\*\*

## 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.
- 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.