

Article

Not peer-reviewed version

Epistemic Risk and the Transcendental Case Against Determinism

[Alessio Montagner](#)*

Posted Date: 25 December 2025

doi: 10.20944/preprints202510.2441.v3

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 and coherence is verified in Mace4, and the derivation of indeterminism 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 targets not our justification for believing determinism, but the conditions that make inquiry possible. 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. Why a Metric Framework?

Any argument based on risk requires a means of comparing possible worlds. Indeed, a belief is risky when there are worlds where it is false, and those worlds have a specific relation with the actual world. How can we represent such relations?

We do not want to rule out *a priori* the possibility of risk originating from inaccessible worlds. Furthermore, we need a notion of *equiproximity*. Consequently, we cannot express relations between worlds through a fuzzy accessibility relation—with higher accessibility corresponding to greater risk—or via a ranking function or partially ordered frame that instantiate a ternary relation $R(x, y, z)$ — reading as “the world x is closer to y than to z ”. Instead, a metric space provides a more perspicuous framework: distances can be represented numerically, thresholds can be fixed, and quantitative comparisons 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

Comparing different worlds is meaningful only if the facts that constitute them are comparable. For instance, it makes sense to compare two worlds by how far an arrow hits from the bullseye. By contrast, it is not meaningful to directly compare the fact that an arrow is shot with the fact that a stone fell into the sea, because there is no common dimension along which to quantify their difference.

Consequently, each world w in the space of possible worlds W is represented as a function from the set of parameters P to their values V :

$$w: P \rightarrow V$$

Each parameter defines a shared dimension along which worlds can be compared, identifying the set of comparable facts and ensuring that differences are measured on an equivalent scale. For every world, a value is assigned to each possible parameter, each value corresponding to a possible fact associated with that parameter — or to the possibility that no relevant fact occurs. For instance, the values of the “coin toss outcome” parameter are (*no toss, heads, tails*).

Parameters are temporalised: for instance, “coin toss outcome at t ” and “coin toss outcome at $t + 1$ ” are two distinct — yet diachronically connected — parameters. The set of parameters with the same temporal coordinate constitutes the state of that world at that time.

Parameters can be of two kinds:

1. **Ordered parameters**, which require quantification and are defined relative to an intrinsically ordered set (S, \leq) of values.
2. **Categorical parameters**, which require identification and are defined relative to an unordered set of values.

For instance, the “distance arrow – bullseye” parameter is ordered, the “coin toss outcome” parameter is categorical.

We further distinguish:

1. **Determining parameters**, whose values are not fixed by another parameter, and for which there exist possible worlds that differ only in the value of that parameter, such that the target-event occurs in one world but not the other (see §2.3 for an example).
2. **Non-determining parameters**, whose values are fixed by another parameter, or for which there are no possible worlds that differ only in the value of that parameter such that the target-event occurs in one world but not the other.

In other words, the determining parameters are those on which the event depends.

Finally, we distinguish:

1. **Random parameters**, whose possible values are equiprobable and independent of any other fact.
2. **Necessary parameters**, which take the same value in all accessible possible worlds.
3. **Contingent parameters**, namely, all parameters that are neither random nor necessary.

In general, a fact corresponds to a single parameter. However, if several facts are ontologically interconnected – that is, there is a causal, grounding, or other significant interdependence relation among them – they count as a single determining parameter. For instance, if an event depends on the outcome of a die roll, “die roll outcome” counts as a single parameter, and cannot be broken down into “position” parameters for each of its atoms.

2.3. The Modal and Normic Metrics

Two distinct metrics The standard metric axioms – identity, symmetry, and the triangle inequality – are not essential: the followings may be treated as quasi-metrics or pseudo-metrics without affecting our reasoning. are defined on the possible worlds space. The first, D , measures *modal distance* between worlds; the second, N , measure their normalcy.

The modal metric D is computed as:

$$D(w, v) : \sum_{p \in P_{det}} d(\pi_w(p), \pi_v(p)) \quad (D)$$

Where:

- P_{det} is the set of determining parameters, a subset of P .
- The sum $\sum_{p \in P_{det}}$ is understood as the supremum over all finite partial sums. If the number of determining parameters with different values in the two worlds is finite, D reduces to a finite sum over determining parameters. If it is countably infinite, D equals the sum of the corresponding series (finite if convergent, ∞ if divergent). If it is uncountably infinite, D takes the value ∞ .
- d measures the difference between corresponding parameter values as $|x - y|$ for ordered parameters, and ε for categorical parameters. Indeed, the values of a categorical parameter, lacking an order, are incommensurable.
- The function π maps a world w to the parameter values that it had at the *time of determination* – that is, it decides which is the relevant temporal state of the world. For instance, if a bomb is set to detonate at $t - 99$ and nothing can prevent its explosion at t , the time of determination is $t - 99$.

In other words, D adds the differences between determining parameters at the time of determination.

The normic metric N is defined as:

$$N(w) = \int_{P_{det}} S(\pi_w(p)), d\mu(p) \quad (N)$$

Where:

- μ is a measure on P_{det} such that $\mu(P_{det}) = 1$. Uniform weighting of parameters is assumed. In the finite case, this is the discrete uniform measure which assigns $1/|P_{det}|$ to each determining parameter. In the countably infinite case, uniform weighting may be implemented via equal infinitesimal weights, or limiting procedures such as Cesàro means. In the uncountably infinite case, it may be taken as the normalised Lebesgue measure.
- S measures the extent to which a parameter value calls for an explanation. Values of necessary or random parameters receive the value 1, while the others are weighted by $1 - g(p)$, with $0 < g(p) < 1$.

In other words, N averages the normalcy degrees of the determining parameters at the time of determination.

To allow a comparison between the two metrics, the degree of normalcy of a world can be converted into a distance from the actual world w_a . In this case, the distance is inversely proportional to normalcy:

$$D_N(w_a, v) = \frac{1}{N(v)} - 1$$

These metrics are defined relative to a target-event, which fixes which parameters count as determining. For example, we may apply D to measure the distance between the actual world, where a plane lands safely, and a counterfactual world where it crashes. The determining parameters are those whose variation would be sufficient to switch between these outcomes in at least some pair of possible worlds. Let us consider three non-interconnected parameters with two possible values, and eight worlds distinguished from each other only by these parameters:

Table 1.

	p_1	p_2	p_3	Outcome
w_1	1	1	1	Lands
w_2	1	1	0	Lands
w_3	1	0	1	Lands
w_4	0	1	1	Lands
w_5	1	0	0	Crashes
w_6	0	1	0	Crashes
w_7	0	0	1	Crashes
w_8	0	0	0	Crashes

The plane does not crash if at least two parameters have value 1. For instance, a change in p_1 is sufficient to switch from w_2 , where the plane lands, to w_6 , where it crashes. Consequently, they are

all determining parameters. This also applies to cases where the three parameters are diachronically connected (i.e., different temporal states of the same parameter) and the target-event requires a minimum duration.

In the case of epistemic risk assessment, the determining parameters are those that, if varied, would not have led the agent to form a false belief at a certain time through a certain method.

2.4. Representing Risky Beliefs

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

- R_D : the modally very close relevant worlds.
- R_N : the most normal relevant worlds.

These subspaces are delimited by a threshold δ : $R_D(\delta) = \{v \in R : D(w_a, v) \leq \delta\}$, $R_N(\delta) = \{w \in R : N(w) \geq \delta\}$.

For N , we posit $\delta = 1$, which is the maximum possible value of $N(w)$.

For D , we posit $\delta = \varepsilon$ — that is, worlds that differ by a single determining categorial parameter are very close. This choice is methodologically conservative: it yields an extremely permissive safety condition, and makes it difficult to challenge determinism on epistemic risk grounds.

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

$$F(B) := w \in R : \neg I_n(b, w)$$

$$F_D(B) := w \in R_D : \neg I_n(b, w)$$

$$F_N(B) := w \in R_N : \neg I_n(b, w)$$

$\mu(F)$ represents the measure of relevant worlds in which the belief is false (i.e., if R is finite, $\frac{|F|}{|R|}$). Similarly, $\mu(F_D)$ and $\mu(F_N)$ represents the measure of its two subsets (i.e., $\frac{|F_D|}{|R_D|}$ and $\frac{|F_N|}{|R_N|}$). Belief B is then risky if the proportion of close or normal relevant worlds falsifying it is at least ψ :

$$\text{Risky}_D(B) = \begin{cases} 1, & \text{if } \mu(F_D) \geq \psi \\ 0, & \text{if } \mu(F_D) < \psi \end{cases}$$

$$\text{Risky}_N(B) = \begin{cases} 1, & \text{if } \mu(F_N) \geq \psi \\ 0, & \text{if } \mu(F_N) < \psi \end{cases}$$

To what is ψ equivalent? The literature includes various proposals, Williamson (2000) and Pritchard (2007, p. 292) maintain that knowledge requires truth in all close worlds. Coffman (2007, p. 390) proposes considering a belief safe if it is true in at least half of close worlds. Even more permissive thresholds are conceivable. and this paper does not take a position on them.

3. Philosophical Grounds of the Metric Framework

3.1. Philosophical Grounds of the Modal Metric

Many theorists maintain that risk correlates, not with mere probability, but with modal distance (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 is not only a matter of how many worlds make the target-event occur, but also of how close those worlds are: given two target-events which occur in an equal number of worlds, the target-event associated with the smaller modal distance from the actual world represents the greater risk.

Yet, several examples indicate that similarity and risk do not always coincide. Consider the following three examples, assuming that the number of worlds where the target-event occurs is always the same:

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 (although in both cases there are infinitely many possible displacements that produce a fall). 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 just enough to make it fall. 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 (a parameter we assume to be categorical). 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* departs from a naive similarity measure to explain such divergences:

$$D(w, v): \sum_{p \in P_{det}} d(\pi_w(p), \pi_v(p)) \quad (D)$$

Ex.1. Without the die roll, displacement is the sole determining parameter, so modal distance is equal to its variation:

$$D(w_{fall}, v_{stable}) = |p_w - p_v|$$

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 π 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 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.

Thus, *D* succeeds where naive similarity-based accounts fail: specifying which differences are relevant, it aligns modal closeness with the intuitive structure of risk.

3.2. Philosophical Grounds of the Normic Metric

Alongside the modal conception of risk, a second notion has recently gained relevance: risk correlates with the *normalcy* of the worlds where the target-event occurs.

According to Ebert, Smith and Durbach (2020, §5), 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)

Intuitively, *y* calls for an explanation if it lacks one, yet we have reason to believe that this explanation *x* might exist.

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 the values of necessary and random parameters:

$$N(w) = \int_{P_{det}} S(\pi_w(p)), d\mu(p) \quad (\text{N})$$

As with the D metric, we evaluate N at the time of determination and on determining parameters only. What matters is the normalcy of the occurrence of the target-event: the normalcy of parameters independent of the target event, or arising after its determination, is irrelevant for this assessment.

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:

Table 2.

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 Φ	$\varphi_w, \varphi_v, \varphi_u, \dots$	Possible initial states
K	$K(vB)$	vB constitutes knowledge
Nec	$Nec(\varphi)$	φ is the value of a necessary (or random) categorical parameter
$DepOnly$	$DepOnly(w, v, \varphi)$	$D(w, v)$ is calculated relatively to a target-event for which φ is the only determinant parameter
$Close$	$Close(w, v)$	w and v are very close
$Risky$	$Risky(vB)$	vB is risky

Here, *vulnerable beliefs* are those beliefs that are false in at least ψ of the relevant possible worlds – i.e., those satisfying $\mu(F) \geq \psi$.

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

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

- (1) $\forall \varphi \text{ Nec}(\varphi)$
- (2) $\forall w \forall v \forall \varphi (\text{DepOnly}(w, v, \varphi) \wedge \text{Nec}(\varphi)) \rightarrow \text{Close}(w, v)$
- (3) $\forall v B (\forall w \forall v \text{Close}(w, v)) \rightarrow \text{Risky}(vB)$
- (4) $\forall v B \text{Risky}(vB) \rightarrow \neg K(vB)$
- (5) $\exists v B K(vB)$

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 Axiom 1

Axiom 1 states that the “initial state” parameter is necessary:

$$\forall \varphi \text{ Nec}(\varphi) \tag{1}$$

In other words, the value of this parameter (i.e., which initial state of the universe among the various possible ones actualised) holds in all accessible worlds. Although other possible initial states abstractly exist, the worlds where they hold are inaccessible.

However, for the purposes of the argument, this view is interchangeable with the one according to which the “initial state” parameter is random:

$$\text{Nec}(\varphi) \leftrightarrow \text{Random}(\varphi)$$

Under this second view, each possible initial state is independent of other facts and equiprobable (since nothing precedes its actualisation and hence nothing can influence it). So all worlds with different initial states are accessible, and each could occur with equal probability.

Although counterintuitive, this position is defensible. It corresponds to Norton’s (2021, §6) principle of mediocrity in infinite lottery logic, which addresses the measure problem in multiverse models and permits uniform probability over infinite sets of possible universes. Moreover, on a propensity view of probability, probability asymmetries reflect physical asymmetries (Strevens 1998); yet the initial state, encompassing all facts that hold at time 0, has no antecedent condition that could constitute an asymmetry.

This axiom can therefore be interpreted in both ways, and the argument remains unchanged.

5.2. Philosophical Grounds of Axiom 2

Under determinism, every fact is fully determined by the initial state of the universe. If the universe has an infinite past, any of its temporal states can be considered its “initial state”. Indeed, under determinism, each state is determined by the preceding one and determines all the subsequent ones; hence, any state suffices to determine the entire temporal sequence, both past and future. In accordance with contemporary physics, each fact depends not on any particular initial state factor, but on the initial state as a whole. Indeed, the initial state is represented by a global Hilbert-space state vector that cannot be factorised into independent local states. This global quantum state exhibits holistic correlations: varying any purportedly local component would require changing the entire

state structure. Consequently, following the criteria expressed in §2.2, the initial state constitutes a single parameter.

Thus, we may define *determinism* as the thesis that the only determining parameter of any fact or event is the occurrence of a specific initial state. Although the determinists might propose a different conception of their view — denying that the initial state constitutes a single parameter — this move would amount to abandoning a scientifically robust notion of determinism.

Axiom 2 reads as follows. Consider two worlds and a target-event. If that event depends solely on which initial state actualised, and this initial state constitutes the value of a single necessary (or random) parameter, then those worlds are very close:

$$\forall w \forall v \forall \varphi (DepOnly(w, v, \varphi) \wedge Nec(\varphi)) \rightarrow Close(w, v) \quad (2)$$

This follows automatically from our conception of risk:

1. Applying the N metric, if the occurrence of the event depends solely on a necessary or random parameter, then the world where it occurs is maximally normal, and hence very close to the actual world;
2. Applying the D metric, if two worlds are assessed relatively to an event that depends solely on a categorical parameter, they are very close. Since the “initial state” parameter requires identification rather than quantification, it lacks an intrinsic natural ordering. Although the Hilbert space is endowed with the Fubini-Study metric, applying it directly to possible initial states would amount to assigning an order to arbitrary labels — much like ordering the faces of a die. The Fubini-Study concerns our ability to distinguish quantum states experimentally, not a natural ordering of facts: it is primarily an epistemic metric, not a purely metaphysical one: events depend on which initial state obtains, and not on what quantity it represents on a scale. Hence, it qualifies as a categorical parameter. Therefore, the two worlds are very close.

5.3. Philosophical Grounds of Axiom 3

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

$$\forall v B (\forall w \forall v Close(w, v)) \rightarrow Risky(vB) \quad (3)$$

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.

5.4. Philosophical Grounds of Axiom 4

Axiom 4 expresses a commitment to an anti-risk epistemology: a (vulnerable) belief qualifies as knowledge only if it is not risky.

$$\forall v B Risky(vB) \rightarrow \neg K(vB) \quad (4)$$

A risk-tolerant epistemology seeks to maximise the number of true beliefs formed, even at the cost of forming some false ones. An anti-risk epistemology seeks to minimise the number of false beliefs formed, even at the cost of forgoing some true ones. For anti-risk epistemology, a rational agent would not form beliefs with too high a risk of being false, so they are not justified, and therefore cannot constitute knowledge.

Many ingenious defenses of the anti-risk approach are found in Rawls (1974). Although Rawls defended it in political philosophy, his reasoning transposes naturally into epistemology. Thus, the idea that a rational agent should be epistemically risk-averse 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 than the alternatives.
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. Risk-neutral frameworks do not imply such strategies, but they tolerate them nonetheless.

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 4 as rational and principled within the framework of our ETA.

5.5. Philosophical Grounds of Axiom 5

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

$$\exists vB K(vB) \tag{5}$$

Consider a clock so fragile that it is broken in 99 percent of relevant worlds. One uses that clock to form a belief about the time. 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: it is a true belief for which, in the actual world, one has a valid justification.

Indeed, if one has good reasons to believe that the clock is functioning, the worlds where it is broken are not close or normal, hence the belief is not risky. Risk requires falsity in a proportion ψ of the *very close* or *most normal* relevant worlds. However high that proportion may be, falsity in a proportion ψ of relevant worlds *simpliciter* is not enough to undermine knowledge without a specific modal topology.

This is the axiom that defines our argument as transcendental: the assumption of the possibility of knowledge. It merely embodies the intuitive idea that not all falsifying worlds are necessarily very close or normal.

6. The ETA Against Determinism from Epistemic Risk

We have defined *determinism* as the thesis that every fact or event depends solely on the initial state of the universe. The axioms were translated into Prover9 (ver. LADR-2009-11A) to verify that the negation of determinism — that is, the existence of an event that does not depend on any initial state — follows. Then, we developed in Lean (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 initial state actualised is the only relevant parameter:

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

By axiom 1, every initial state is the value of a necessary (or random) parameter. Thus, given (6), every fact depends solely on a necessary (or random) parameter:

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

(7), given axiom 2, leads to a modal collapse: every world is very close to the actual one. Indeed, for any target-event and any world pair considered, the “initial state” parameter is the only determining parameter:

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

Given axiom 3, from (8) follows that every vulnerable belief is risky:

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

Therefore, by axiom 4, no vulnerable belief can constitute knowledge:

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

But (10) contradicts axiom 5. Therefore, our assumption must be discarded, and determinism is transcendently rejected. Indeed, the axioms imply the existence of events that do not have an initial state as their only determining parameter:

$$\exists w \exists v \forall \varphi \neg \text{DepOnly}(w, v, \varphi) \quad (11)$$

Libertarianism, instead, avoids the modal collapse. Indeed, under libertarianism worlds may also differ in the agent’s free deliberative choices — which constitute additional determining parameters — thus allowing for different modal distances. Moreover, certain deliberative choices may be impossible with certain other parameter values in falsifying worlds, excluding those worlds from the possibility space and thus reducing the proportion of falsifying worlds among the very close ones.

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 5), that risk undermines knowledge (axiom 4), and define risk in terms of modal closeness or normalcy (axioms 2 and 3).

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 initial state of the universe — is itself necessary or random (axiom 1).

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.

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>.
- 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.
- 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.18047016>.
- 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.
- 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.