

# When Prediction Replaces Parts: A Criterion for What's Real

## Abstract

Mereological nihilism claims there are no hurricanes, only particles arranged hurricane-wise; no cells, only atoms arranged cell-wise. This paper argues that once we accept four uncontroversial premises (causality, locality, information, and thermodynamics), statistical boundaries become inevitable features of physical reality, not optional human conventions.

These boundaries create causal autonomy: systems become predictable and controllable through their boundaries alone, independent of external microstructure. We formalize this using the  $\epsilon$ -machine/ $\nu$ -machine framework from computational mechanics. When macro-variables predict the target class of future macro-states as effectively as complete micro-descriptions, relative to a fixed prediction task, the macro-level achieves computational sufficiency for that task. The whole genuinely causes its own future states.

A macro-description is non-leaky precisely when it is computationally self-contained: when fixing the macro state fixes the future macro dynamics, and no additional micro-information improves prediction or control. In such cases, the macro-level implements a closed computational process — one that can be simulated, intervened upon, and reasoned about independently of its micro-instantiation.

Disputes about whether wholes are ‘real’ often trade on an unacknowledged ambiguity in the term *real*. In practice, at least four distinct criteria are in play: **universality** (cross-agent convergence under shared constraints), **physicality** (spatiotemporal and energetic existence), **fundamentality** (inclusion in the lowest-level ontology), and **causal autonomy** (stable boundaries supporting reliable intervention). Our argument does not challenge fundamentality as a criterion but shows that causal autonomy and universality are objective properties of physical systems, even when fundamentality is denied.

This constrains the nihilist’s position to a largely terminological dispute about what counts as ‘existence,’ rather than a disagreement with scientific practice. If “facts about arrangements” perform all the causal work objects were supposed to do (supporting intervention, enabling prediction, structuring scientific practice), then denying their reality in the sense of causal autonomy becomes empty verbalism. The boundary does not label pre-existing molecules; it creates a new causal architecture. That architecture is the object.

We defend this view against retreats to “mere usefulness,” address challenges from consciousness, and demonstrate practical implications across medicine, ecology, and technology. The argument does not force concession but makes nihilism obsolete as a claim about causal autonomy. Where causal powers, counterfactual stability, and resistance to informational intrusion converge, objecthood is instantiated. Wholes are not fundamental, but they are metaphysically real as unified loci of causal powers. A pattern is universally discoverable when any sufficiently capable agent attempting to predict and intervene within the same constraints is forced to encode it, on pain of incurring

unsustainable algorithmic complexity. This universality is epistemic rather than metaphysical: it reflects convergence under shared constraints, not fundamentality. The hurricane is as real as its molecules in this sense.

## Introduction

*Is a hurricane “real,” or is it merely air molecules in motion?* While the question may initially appear trivial, we name hurricanes, track them, and flee from them. Mereological nihilism challenges this intuition. It claims there are no hurricanes, only molecules arranged hurricane-wise; no cells, only particles arranged cell-wise.

## Clarifying ‘Real’: Four Senses

Before proceeding, we must clarify what we mean by “real.” Disputes about whether wholes are ‘real’ often trade on an unacknowledged ambiguity in the term *real*. In practice, at least four distinct criteria are in play:

- **Universality:** Cross-agent convergence under shared constraints. Something is “real” if any sufficiently capable intelligence, interacting with the same constraints, would be forced to recognize it (e.g., hurricanes, cells,  $\pi$ ).
- **Physicality:** Spatiotemporal and energetic existence. Something is “real” if it occupies spacetime and exchanges energy (e.g., molecules, cells, hurricanes).
- **Fundamentality:** Inclusion in the lowest-level ontology required to explain everything else (e.g., quantum fields, spacetime structure, fundamental constants).
- **Causal Autonomy** (Leakiness): Stable boundaries supporting reliable intervention and prediction under perturbation (e.g., organisms, organs, hurricanes).

Our argument defends wholes as real in the senses of universality and causal autonomy, even if they lack fundamentality or (in some cases) physicality. Nihilism privileges fundamentality as the sole criterion, but this is a choice, not a revelation about the world.

In this paper, ‘real’ means: supporting a self-contained, non-leaky, computationally closed description that is sufficient for prediction, intervention, and control across a wide class of realizations.

The hurricane’s boundary is not a label; it is a statistical screen that makes the interior conditionally independent of the exterior. And boundaries like this are not optional features of reality; they are mathematical consequences of any dynamical system satisfying locality, finite information capacity, and energetic cost of computation.

The boundary is not optional. It emerges from the constraints governing any physical system. Let us make this precise. If we accept four premises that even most mereological nihilists grant, a fifth conclusion follows with mathematical necessity:

1. **Causality:** Events have causes and effects
2. **Locality:** Causes operate through local interactions
3. **Information:** Systems carry information about other systems
4. **Thermodynamics:** Processing information costs energy; complete tracking of all micro-information is physically expensive

We can *demonstrate* that certain arrangements of matter necessarily create statistical boundaries. By “demonstrate,” we refer not to a logical theorem, but to

an inevitability inherent in the physical constraints of the universe. These boundaries are neither optional nor projected; they are mathematical consequences. Once they emerge, they perform causal work that particles lacking a boundary-level description cannot.

This property requires a definition:

**Causal Autonomy:** A system is causally autonomous when its future states are fully predictable and controllable, to a specified tolerance, using only variables defined at its own boundary.

This creates a dilemma for the nihilist: either accept that “facts about arrangements” do everything objects were supposed to do (making their denial of objects empty verbalism), or explain why causal efficacy is not sufficient for existence. They have no principled alternative criterion. The universe is not flat; it is layered. And those layers are as causally autonomous as the bedrock.

Before proceeding, we must clarify what we mean by “causal work.” This matters because the nihilist might grant that boundaries are useful while denying they are real.

Here we adopt the interventionist theory of causation :  $X$  causes  $Y$  if intervening on  $X$  changes  $Y$ , holding other factors fixed. On this view, causation is not about “little things banging into other things” but about which variables make a difference under intervention. On the interventionist account of causation, successful macro-level interventions establish that macro-variables are causally sufficient for prediction and control. This establishes explanatory and interventional autonomy, not ontological fundamentality.

Predictive equivalence matters because it reveals which level of description captures the intervention-relevant structure. If macro-variables predict as well as micro-variables, then the macro-level *is* the causal level for that phenomenon. Adding micro-detail changes nothing about what interventions will succeed.

The Ptolemaic objection (that epicycles predicted planetary motion without being real) actually supports our case. Epicycles failed precisely because they did not support intervention: you could not steer a planet by manipulating its epicycle. The macro-objects we defend (cells, hurricanes, economies) pass the interventionist test that epicycles fail.

Crucially, this is not a denial of microphysical completeness, nor an argument for “spooky” emergence. It is an account of why causal units appear at multiple scales in a universe constrained by locality, information flow, and thermodynamics.

The argument proceeds in four stages. First, we demonstrate that statistical boundaries are inevitable consequences of these four premises (Parts 1-2). Second, we show these boundaries create causal autonomy, allowing systems to be predicted and controlled through their boundaries alone (Part 3). Third, we formalize a rigorous test for when boundaries constitute genuine objects (Part 5). Finally, we address objections and show why this matters practically (Parts 6-10). At each stage, we ask: can the nihilist consistently accept the step while maintaining their position?

# What the Nihilist Actually Says (And Why Their Foundation Is Shaky)

Before showing why boundaries are inevitable, let us confront the challenge. Alex O'Connor articulates the strongest nihilist case: *“When did this glass begin? No new matter was created, just rearranged soup. The only distinction between objects is mental labeling. All things emerge from a quantum soup of foundational matter; the soup is all that is real.”*

This position appears scientific but conflates material substrate with causal structure. To be fair, this reductionist view has been spectacularly successful for its intended domain: understanding fundamental forces by smashing atoms. But most of science, and indeed all of life, is not about smashing atoms. The four premises do not merely *allow* boundaries; they *require* them in certain configurations. A cup’s boundary is not an abstract essence of “cup-ness”; it is the measurable fact that its walls screen off the liquid’s temperature from external atoms.

Here is the *tu quoque*: The nihilist assumes particles are self-subsistent objects in the soup. But quantum mechanics (entanglement, indistinguishability) shows particles lack inherent self-identity. You cannot tag an electron and track it like a billiard ball. They are nodes in a relational structure. This means they have no ontological privilege over “wholes.” Both are Real Patterns in the formalism. If nothing is an inherently “separate” object at the base level, the nihilist’s attempt to deny higher-order objects based on their composition loses its anchor. The “soup” is not made of things; it is made of constraints.

## Part 1: The Necessity of Statistical Boundaries

Given causality, locality, and information, we can analyze the consequences of local particle interactions. These interactions form causal networks where, inevitably, certain nodes screen off others from the rest of the graph.

Imagine a set of 100 particles bouncing in a perfectly insulated box. External particles can only affect the interior through wall collisions. Once you know *exactly* what is happening at the walls (momentum, angle, timing), knowing about external particles adds zero predictive power.

This represents a Markov blanket, a boundary where knowing the state of the barrier renders the exterior irrelevant for predicting the interior. Think of a castle with thick stone walls. To predict the temperature inside the great hall, you only need to know the temperature of the stones on the inner surface of the wall. The weather five miles away is irrelevant except insofar as it has already affected those stones. The wall screens off the rest of the universe.

Formally:

$$P(\text{inside} \mid \text{walls}, \text{outside}) = P(\text{inside} \mid \text{walls})$$

When this equality holds, the “outside” variable becomes mathematically redundant. The nihilist objects: “It is particles arranged box-wise.” But arrangement is not neutral. This property, a screen between regions, did not exist before the box. The boundary is a new causal constraint on information flow. (Part 5 formalizes this as computational closure: when macro-variables predict the future as well as micro-variables, the micro-details become redundant.)

Think of the landscape of possible boundaries like a topographical map: certain “valleys” in configuration space guarantee boundary formation. Self-reinforcing feedback loops are one valley. Lipid bilayers are another. Evolution searches this landscape and finds boundaries because they work.

Not every arrangement achieves this stability; we will see examples of failure when we formalize the test.

Physical walls are straightforward cases. But the profound question is: can statistical boundaries alone create causal autonomy?

## Part 2: When Boundaries Become Causal Firewalls

More profound than physical walls are statistical boundaries that emerge from interaction patterns alone.

Consider an *E. coli* bacterium. Its membrane proteins: - Sense external glucose - Pump glucose inside - Regulate internal metabolism

Over time, the internal state becomes predictable from boundary states alone. External chemistry becomes irrelevant for predicting internals, *given* what is crossing the membrane.

The nihilist might object: “But this is simply lipids obeying chemistry!”

Response: This is true, but the specific arrangement achieves the Markov condition we saw in Part 1. The membrane creates conditional independence, a measurable property that raw chemical “soup” lacks. We do not *choose* where to draw the boundary; we *discover* it by measuring whether this screening-off effect holds.

However, which boundaries count? Not all mathematical boundaries are equal. The cell membrane represents a “joint” in nature because it corresponds to a self-maintaining feedback loop that evolution has independently discovered multiple times. By contrast, failed boundaries are gerrymandered attempts that never achieve stable closure. The cell membrane achieves robust closure. That is the difference between a natural boundary and a gerrymandered surface.

Nature brims with natural boundaries: cell membranes, organ systems, ecosystem edges. They are not labels we impose; they are discoverable features of causal networks. Evolution acts as a blind computational engine, scanning the landscape of configuration space and discovering these boundaries because they work to ensure the stability of the system against thermodynamic decay.

It is crucial to clarify that this process requires no observer. Evolution acts as a blind selection mechanism against thermodynamic decay. Configurations that fail to establish effective Markov blankets are unstable; they dissipate because they cannot regulate their internal entropy against environmental perturbations. The boundaries we observe are the survivors of this selection process. These boundaries are not arbitrary conceptual carvings; they are the specific subset of arrangements that possess the physical property of self-maintenance. They constitute natural kinds defined by their thermodynamic stability.

However, these boundaries are not binary walls; they are statistical gradients. Boundary closure is a matter of degree because leakiness can be quantified as the rate at which algorithmic complexity must be imported from the environment to sustain accurate prediction. Boundaries with low complexity influx support stable compression and intervention; those with high influx collapse into additive descriptions. A granite boulder has extremely tight closure; very little external

information (short of a sledgehammer) disrupts its macro-state prediction. A social category like class or race is leakier; its predictive power fluctuates based on context. The boulder is “more real” only in the sense that its closure is more robust across a wider range of perturbations.

It is tempting to view dimensions like “Law” or “Culinary Tradition” and call them observer-relative because they depend on human minds. But this confuses origin with ontology. A dimension exists wherever a specific set of constraints creates a stable landscape for prediction. Whether a “Checkmate” exists is not a matter of opinion; it is a mathematical fact derived from the constraints of Chess. Whether a “Corporation” exists is not a matter of feeling; it is a fact derived from the constraints of Law. To say a dimension is “observer-relative” implies we can wish its boundaries away. But try ignoring the legal dimension (e.g., stop paying taxes) and you will collide with its reality just as hard as you collide with a physical wall. The boundaries we encounter are forced upon us by the constraints of that system. Leakiness does not equal non-existence; so long as the boundary screens off enough noise to allow intervention, the entity is causally autonomous.

Statistical boundaries exist and achieve screening-off. But what do they actually *do*?

### Part 3: The Birth of Causal Autonomy

When a system achieves conditional independence, it gains causal autonomy: it becomes predictable and controllable through its boundary alone.

**Scenario A: Particle Soup** 100 particles bounce randomly. To predict particle #57’s path, one must track all 99 others. No shortcuts.

**Scenario B: Proto-cell** The same 100 particles organize: 20 form a boundary, 80 cluster inside. Now: - Predicting internal particle #57 requires only boundary states - External particles become irrelevant

This is not computational convenience. It is a new causal architecture. You can *intervene* on the whole by tweaking boundary conditions (shining light on photoreceptors, adjusting membrane permeability) rather than manipulating individual particles.

This shift from particle soup to proto-cell illustrates a specific mathematical relationship: dimensional reduction. The boundary formation reduces a high-dimensional state space (requiring tracking of  $10^{23}$  particle coordinates) into a low-dimensional state space (requiring only boundary temperature, pressure, or permeability) without significant loss of predictive fidelity. This reduction is objective; it relies on the mathematical fact that internal degrees of freedom have become statistically decorrelated from external variables. The Markov condition is not merely a description but a measurable dimensional collapse in the state space.

When does composition occur? It occurs when Markov conditions are satisfied to the degree that intervention on the boundary reliably controls the interior. Biological boundaries are leaky, but they achieve robust closure across relevant timescales and perturbations.

Critics might argue this makes reality interest-relative, since closure depends on an error tolerance. Who decides how much “leakage” is acceptable?

The answer is thermodynamics and survival. The tolerance is not arbitrary; it is the threshold of disintegration. If an organism sets its tolerance too loose

(ignoring dangerous external variables), it dies. If it sets it too tight (trying to track every atom), it starves from the computational cost of prediction. The “correct” tolerance is an objective survival parameter discovered by evolution, not chosen by observers. The boundaries we observe are not the ones we like; they are the ones that successfully balanced predictive accuracy against metabolic cost over billions of years.

Nature itself enforces which compressions work. We do not decide that the cell membrane works as a boundary; experiments converge on the same boundaries regardless of investigator or theoretical commitment. Projectibility (support for counterfactual predictions) is objective. Approximate closure does not undermine reality; it is precisely what allows robustness across noise, perturbation, and scale.

This explains why doctors treat organs, not cells. Heart failure is not fixed by adjusting cardiomyocyte ion channels. Cardiologists alter organ-level dynamics. The heart’s boundary creates causal autonomy.

Evolution selects for boundaries, not particles. The nihilist’s flat ontology cannot explain why life builds hierarchies.

## Part 4: The Dimensionality of Wholes

But hierarchy is not the only structure. Reality is also dimensional. We must distinguish between scale and dimension. A corporation has no causal closure in physics. It is just people and paper. But in the dimension of law, it has a tight, impenetrable boundary. It can sue, be sued, and hold assets.

Consider the hotdog dilemma: Is a hotdog a sandwich? In the culinary dimension, the boundary is drawn by texture and expectation, perhaps excluding the hotdog. In the regulatory dimension (tax law), the boundary may explicitly include it. Both boundaries are causally autonomous because both allow for perfect prediction within their respective domains (predicting a tax bill versus predicting a culinary experience). Reality is not just layered vertically; it is dimensional. Within a given domain of constraints (physical, biological, legal), entities that achieve stable predictive and interventional roles function as domain-relative objects.

Crucially, these dimensions are not merely observer-relative; they are constraint-relative. It is tempting to view the legal or culinary dimensions as subjective projections, but this confuses origin with ontology. A dimension exists wherever a specific set of constraints, whether thermodynamic, legal, or game-theoretic, creates a stable landscape for prediction. Once these constraints are established, the closure they generate is objective. Whether a move is “Checkmate” is not a feeling; it is a mathematical fact derived from the constraints of Chess. Whether a corporation is “Liable” is a fact derived from the constraints of Law. We may choose which dimension to engage, but once we enter the domain, the boundaries we encounter are forced upon us by the system’s constraints. These constraint-relative entities function as objective invariants within their dimensional logic, imposing non-negotiable consequences on the agent’s future possibilities just as rigidly as physical walls impose spatial constraints.

But showing that boundaries *can* exist is not enough. We need a test for when they *do* exist.

## Part 5: The Information-Theoretic Test

The nihilist may object: “You have not proven these arrangements are *objects*, only that they are causally useful.” To address this, we must differentiate between genuine wholes and gerrymandered collections.

Consider a thought experiment: A trumpet is connected to a lightbulb. We derive a mathematical formula that predicts exactly how bright the light will be based on the note played. We possess a description, but do we possess an *explanation*? A skeptic might say no; we have not described the wires or the electrons. We have a rule, not a mechanism.

But the Rosas et al. (2024) framework offers a rebuttal: If the system is causally closed, the ‘Trumpet Formula’ is not merely a description; it constitutes the *maximal explanation* possible. Two types of prediction machines capture this distinction:

- $v$ -machine (Upsilon): The complete micro-story, the best possible prediction of the macro-future using every molecule’s exact position and velocity.
- $\varepsilon$ -machine (Epsilon): The compressed macro-story, the best prediction of the macro-future using only macro-data like temperature, membrane state, or the note played.

When both predict equally well, the macro-level has achieved computational closure. This is the condition where macro-variables fully determine future macro-states without needing micro-details. Causal closure here does not mean the macro-level is independent of microphysics, but rather that it is sufficient for prediction and intervention over the target variables without needing to look “down.” The system becomes informationally self-contained.

This test serves as the necessary firewall against ontological triviality. A critic might ask: “What prevents us from defining a trivial dimension where the Moon and my left shoe constitute a single object?” The answer lies in information compression.

Consider an intuitive test: can you describe the whole more efficiently than enumerating its parts?

To describe a “Shoe-Moon” object, you must first describe every property of the shoe (material composition, shape, lacing structure, wear patterns), then describe every property of the moon (orbital parameters, surface topology, gravitational effects, geological composition). Grouping them under a single label saves no descriptive effort and compresses no information. The description is purely additive.

By contrast, describing a hurricane is qualitatively different. Instead of enumerating the position, velocity, and thermodynamic state of  $10^{23}$  independent air molecules, you describe a single structural constraint: a rotating atmospheric vortex with specific pressure gradients and angular momentum. This single pattern specification captures the collective behavior of trillions of particles. The compression ratio is astronomical.

This intuition can be formalized via Kolmogorov complexity: Compression is not epistemic convenience but ontological diagnosis. A system that admits a stable, low-complexity description of its causal evolution thereby constitutes a unified causal power. Where such compression fails, objecthood fails with it. The “Shoe-Moon” fails this test because it is algorithmically incompressible relative to its components. The hurricane succeeds because the vortex pattern achieves massive compression.



This compression is not mere notational convenience. Information is physical ; processing information costs energy. A system that can predict its environment by tracking a single boundary state (the hurricane’s eye wall pressure and rotation) rather than tracking  $10^{23}$  particle trajectories has an enormous thermodynamic advantage. Natural selection operates on this energetic reality. An organism that models its environment using the “Shoe-Moon” additive method will starve from computational overhead. Evolution selects for compressive boundaries because they minimize the metabolic cost of prediction.

Trivial dimensions are brittle; they require constant ad-hoc parameters to maintain prediction. Genuine dimensions are low-maintenance because they compress the state space. Real entities earn their ontological status by reducing the computational cost of prediction. If a boundary does not achieve compression, it does not exist.

Adding substrate information yields zero predictive improvement. The whole is not a convenient fiction; it genuinely causes its own future states. If knowing the position of every electron ( $v$ ) adds *zero* predictive power over simply knowing the note ( $\varepsilon$ ), then the mechanism is informationally redundant. We must distinguish between the signal (the note) and the noise (the specific electron paths). The bulb does not care about the vibrational micro-state of the brass; it cares only about the frequency profile. The macro-variable screens off the micro-noise. In a causal sense, the signal is what matters.

Crucially, this test applies across dimensions. In the dimension of tax law, an  $\varepsilon$ -machine tracking ‘corporate assets’ predicts legal outcomes perfectly, while a  $v$ -machine tracking the physical location of every employee adds nothing. The corporate entity achieves computational closure in legal space, satisfying the test just as robustly as the trumpet does in acoustic space.

Consider phlogiston theory in eighteenth-century chemistry. It failed this test decisively because every new experiment revealed information leakage. When metals gained weight upon burning (instead of losing phlogiston as predicted), theorists invented “negative phlogiston.” When different substances showed different weight changes, more ad-hoc parameters appeared. This constant maintenance signaled that no causally autonomous screening-off was occurring. Phlogiston was not a natural boundary but a failed attempt to carve one. This is the operational difference between natural joints and gerrymandered surfaces: causally autonomous patterns are low-maintenance; fake ones require constant propping up and ad-hoc patches.

Think of water flowing through a pipe. An  $v$ -machine would try to predict the flow by tracking the vector and velocity of all  $10^{23}$  water molecules. An  $\varepsilon$ -machine uses the Navier-Stokes equations, treating the water as a continuous fluid in terms of pressure and viscosity. When the fluid equations predict the flow as accurately as the molecular simulation, the “fluid” is causally real. It’s not just a summary; it’s the level where the causal constraints operate.

The nihilist might concede the test is rigorous but deny its ontological significance.

These considerations reveal a common structure underlying causal autonomy, universality, and dimensional overlap. A boundary that exhibits low leakiness permits sustained compression of the system’s dynamics. Wholes recur across domains because the same computational structures arise as stable, non-leaky coarse-grainings under widely different micro-dynamics. Universality here reflects the reappearance of the same effective computation, not a privileged ontological

status. Cross-dimensional overlap is therefore not an additional metaphysical ingredient, but a signal that the boundary resists complexity leakage under many forms of intervention. Entities that persist across dimensions are those whose boundaries minimize algorithmic complexity inflow across a wide range of constraints.

## Part 6: The Nihilist's Retreat / The Implementation Layer

A sophisticated nihilist might respond:

“I agree that particles arranged membrane-wise satisfy Markov conditions, and that this is discoverable. However, *satisfying a condition* does not create a new entity. There are particles, and there are *facts about* particles. Facts are not objects.”

If facts about arrangements perform causal work, denying their objecthood becomes an exercise in empty language. The nihilist has reached a position that is formally consistent but doing no explanatory or predictive work. What remains is a purely verbal distinction regarding whether “real” denotes “fundamental particle” or “constraint-determined pattern.”

The nihilist makes a category error by conflating fundamentality with reality.

Fundamentality refers to the abstraction hierarchy and the implementation substrate (e.g., silicon is fundamental to software).

Reality refers to the tightness of causal closure (e.g., the software bug causes the crash, not the silicon).

Physical fundamentality has a special status not because it is more real, but because it is the implementation layer for everything else. Legal reality runs on paper, servers, and human brains; software reality runs on silicon and electricity. While dimensions are orthogonal in their causal logic; one cannot explain a tax code using voltage; they are tethered by existential dependency. If the physical implementation layer is destroyed, the higher-order dimensions vanish as well.

A sophisticated objection must be addressed here. The nihilist might grant that macro-entities achieve causal closure while insisting they remain ontologically derivative because they supervene on the micro-substrate. We must distinguish between Existential Dependence and Causal Independence:

1. Existential Dependence: The macro-entity requires the micro-substrate to exist. If the atoms vanish, the cell vanishes.
2. Causal Independence: The causal trajectory of the macro-entity is determined by macro-variables, with micro-fluctuations screened off.

The nihilist error is to assume that (1) implies the negation of (2). However, in systems achieving computational closure, the specific micro-trajectory becomes chemically irrelevant to the macro-outcome, provided it remains within the manifold of states compatible with the macro-description. The implementation layer provides the *capacity* for the system to exist, but the macro-structural constraints determine the *direction* of its future states. Supervenience establishes a dependency relation, not an identity relation. The software bug supervenes on silicon states but is not reducible to them; the causal explanation operates at the software level.

Consider the phenomenon of Multiple Realizability. You can run the exact same spreadsheet calculation on silicon transistors, on vacuum tubes, or with a room full of people using abacuses. If the spreadsheet crashes due to a logic error (division by zero), the crash is caused by the algorithmic structure, not by the voltage of the silicon, the filaments in the tubes, or the wood of the abacuses. Because the crash can occur on three totally different physical substrates, the *cause* cannot be physical; the cause must be at the algorithmic level.

The physical layer provides the capacity for existence, but the structural layer dictates the outcome. The nihilist who insists on reducing everything to the substrate is like a hardware engineer trying to find a spelling error by inspecting the motherboard with a microscope. They are examining the reality of the implementation while missing the reality of the dynamic.

This even resolves the “China Brain” puzzle . If a population of a billion people used walkie-talkies to simulate the firing of neurons, would a collective mind exist? Our framework suggests that if the macro-pattern achieves computational closure, enabling predictions and interventions that tracking the individual citizens cannot, then that “mind” is a valid causal entity. It is software running on a substrate of people, just as Excel runs on a substrate of transistors.

This distinction clarifies how orthogonal dimensions interact without collapsing into incoherent pluralism. How can a legal concept (a warrant) cause a physical event (an arrest)? They interact through the implementation layer. Dimensions are distinct in their causal logic (since one cannot explain a tax statute using voltage) but remain tethered by existential dependency. Interaction occurs when a causal chain in one dimension triggers a state change in the implementation layer, which then propagates upward into another dimension. For example: a judge signs a warrant (Legal Dimension). This moves physical ink on paper (Implementation Layer). This physical token is observed by a police officer, changing their cognitive state (Psychological Dimension), leading to a physical arrest (Physical Dimension). We avoid the “ghost in the machine” problem because we are not positing distinct substances passing through one another; we are describing software processes running on the same hardware.

This logic applies recursively across every level of the universe. The rock is more fundamental than the corporation, but the corporation is just as causally autonomous within its causal domain.

The boundary changes what would happen if you intervened. If the boundary is real and does causal work, what could it mean to deny the objecthood of what it bounds? In a meaningful sense, the boundary defines the object.

Consider what it means for a boundary to “define” an object. The hurricane’s boundary is not a label we apply to existing molecules; it is a causal constraint that changes what those molecules collectively can do. Before the boundary forms, particle #57’s trajectory depends on all other particles in the region. After the boundary forms, particle #57’s trajectory depends only on boundary conditions. The boundary has created a new causal architecture. That architecture is the object.

If the nihilist retreats to facts, we must ask: what kind of thing is a fact that does causal work?

This raises a deeper question about the relationship between physical structure and higher-order properties. Nowhere is this more pressing than in understanding consciousness.

## Part 7: Where Is the Triangle?

A significant challenge arises from consciousness. A critic might ask: “*When I visualize a triangle, where is it? If I open the brain, I see only neurons firing. The triangle I’m experiencing isn’t identical to those neurons.*”

This question contains two distinct puzzles, and intellectual honesty requires separating them:

1. The location problem: What kind of thing is mental content, and how does it relate to neural activity?
2. The phenomenal problem: Why is there *something it is like* to visualize a triangle?

Our framework addresses the first problem directly. The paper by Rosas et al. provides formal support for the analogy of the mind as software: a system runs software when its macroscopic state determines its future state irrespective of hardware details. When you visualize a triangle, that representational pattern is substrate-independent (a valid causal entity because it is the effective theory of that cognitive process). The “triangle” (as functional content) is the  $\epsilon$ -machine; the firing neurons are the hardware implementation. The reason you cannot find the triangle by dissecting neurons is the same reason you cannot find Microsoft Word by examining silicon: you are looking at the wrong level of description. Mental content is best characterized at the level where cognitive patterns achieve explanatory and interventional sufficiency, independent of their neural implementation.

The “screen” objection (O’Connor argues computer analogies fail because there is no screen in the brain) misses this point: the  $\epsilon$ -machine *is* the screen. It is the level at which the pattern achieves explanatory and interventional sufficiency. No external observer is needed because the system itself operates at that level. The triangle exists at the  $\epsilon$ -level, not the  $\nu$ -level. It is not *in* the neurons any more than Python code is *in* silicon. It is a constraint-determined pattern that runs on the substrate.

Consider a child drawing a dragon. Biologically, the dragon does not exist (it has no DNA). But in representational space, the dragon is a compression of predator, a high-fidelity symbol that predicts fear responses, narrative structures, and cultural transmission.

If we ask “Is the dragon causally autonomous?”, we must ask “In what dimension?” In biology: No. In the causal architecture of human psychology: Yes. It achieves closure because intervening on the symbol (showing a picture of a dragon) reliably produces a predictable macro-state (fear/awe) without the subject needing to process the micro-details of actual lions or snakes. The dragon is a causally autonomous psychological interface for a dangerous world.

But what about phenomenal experience (the redness of red, the felt quality of visualizing that triangle)?

Here we must acknowledge our framework’s current limitations. The explanatory gap between “this system processes information in a self-modeling way” and “there is something it is like to be this system” remains open. Perhaps phenomenal consciousness is functional organization (Dennett), perhaps computational closure is necessary but not sufficient, or perhaps proto-phenomenal properties are fundamental. We do not adjudicate between these options, but we recognize the obligation our framework creates: if minds are real

patterns achieving causal closure, we must eventually explain why such patterns are accompanied by experience.

Setting aside the phenomenal problem's unresolved status, these considerations point toward a specific metaphysical picture.

## Part 8: Rainforest Realism

The debate becomes semantic. Ladyman and Ross propose Rainforest Realism: reality (in multiple senses) is lush with entities at every scale where patterns achieve projectibility (support counterfactuals) and information compression. A hard-nosed metaphysician might insist this is an account of explanation, not existence. We accept the charge but deny the distinction. There is no principled gap between “what exists” and “what plays an indispensable causal role in the best explanation of reality.” Fundamentality without an autonomous causal role is an honorific, not an explanatory category.

This position has three primary implications for how we understand reality.

First, scale relativity. At scale A (micro), the cup does not exist; at scale B (macro), the cup *does* exist. Both scales are equally causally autonomous; physics does not grant the micro-scale ontological privilege. Consider: a human being exists simultaneously as a quantum field, a cellular colony, and a voting citizen. None of these descriptions is “more real” than the others. Each level achieves closure at its own scale. The voter is a valid object because you can predict election outcomes using “voters” without knowing anything about their mitochondria. This predictive independence demonstrates causal autonomy at the civic scale. The nihilist's attempt to reduce the cell to atoms is a category error: it confuses one scale of description with the only reality.

Second, the **Five Axes of Realness**. We can distinguish between different “degrees” of reality by measuring entities along five independent axes: 1. **Causal Closure** (Leakiness): The effectiveness of the boundary. How completely does the boundary screen off the interior from external noise, supporting reliable intervention? 2. **Fundamentality**: The level in the implementation hierarchy. Quarks are fundamental (low abstraction); corporations are highly abstract. 3. **Constraint Tightness**: How negotiable the boundaries are. Thermodynamics is inflexible; culinary categories are socially negotiable. 4. **Sharedness** (Universality): The degree of agent convergence. The value of  $\pi$  is a universal invariant; specialist expertise is a narrow one. 5. **Cross-Dimensional Overlap**: The multiplicity of independent constraint systems in which the entity appears without reinterpretation. An entity is “more real” the more distinct domains (e.g., physics, chemistry, biology, law) converge on the same boundary.

An entity scores differently across these dimensions. A granite boulder scores high on all five: tight closure, low abstraction, inflexible constraints, universal sharedness, and overlap across physics, engineering, and geology. A corporation is abstract and socially negotiable, yet it achieves tight causal closure in legal space, with high sharedness among lawyers and low overlap beyond law. Hurricanes score highly on universality, causal autonomy, and cross-dimensional overlap (appearing in physics, meteorology, economics, and public safety). By distinguishing these axes, we avoid conflating “fundamental” with “real” and can respect the “lushness” of the rainforest without descending into relativism. No single notion of ‘real’ plays all explanatory roles; insisting that only one does is itself a metaphysical choice, not a scientific discovery.

Third, structural invariance. These boundaries are not arbitrary labels but stable solutions to the constraints of survival and coordination. Consider an analogy from mathematics:  $\pi$  is not a physical object, nor is it a mere invention. It is a constraint-determined invariant. Given the axioms of geometry, any intelligence working with circles must eventually discover  $\pi$ ; it is forced into existence by the constraints of the problem. Biological and social boundaries are similar. To exist, in any meaningful sense, is to be an invariant solution to a constraint problem.

A potential objection must be addressed: is compression observer-relative? Perhaps hurricanes seem “simple” only because we describe them in human language rather than in some alternative representational system. This concern confuses the representation with what is being represented.

The compressibility of a pattern is independent of the particular symbolic system used to encode it. A hurricane achieves massive compression whether described in English prose, differential equations, binary machine code, or a hypothetical alien notation. What matters is not the absolute length of any particular encoding, but the ratio: how much shorter is the pattern description compared to the full enumeration of component states? This compression ratio remains invariant across representational frameworks, provided those frameworks are computationally universal (capable of expressing the same range of computable functions).

The invariance is structural, not accidental. The hurricane’s rotating vortex is a stable attractor state in the dynamical system governing atmospheric physics. The constraint that creates this stability (conservation of angular momentum coupled with thermodynamic gradients) exists in the physical system itself, not in our linguistic or mathematical conventions. Any sufficiently capable intelligence attempting to predict atmospheric behavior must eventually recognize and encode this pattern. Failing to do so means expending vastly more computational resources tracking irrelevant microscopic trajectories that average out at the macro-scale. The pattern is forced upon any predictor by the thermodynamic structure of the domain.

This connects directly to the  $\pi$  analogy. Just as any intelligence working with Euclidean circles must discover  $\pi$  (because it is forced by geometric constraints), any intelligence modeling hurricane dynamics must discover the vortex compression (because it is forced by thermodynamic constraints). The boundaries we identify are objective in precisely this sense: they are constraint-determined invariants that emerge necessarily from the physics, independent of the observer’s representational choices.

This explains why the nihilist’s alternative; the belief that “causally autonomous” causation is limited to “microbangings” of particles; fails. Fundamental physics does not describe particles banging into each other; it describes mathematical constraints. Causation is an interventionist concept that applies to special sciences like biology and medicine, not to fundamental physics.

These considerations have implications for long-standing debates in physics.

## **Part 9: Free Will**

Physicists often mock Superdeterminism as requiring “cosmic conspiracy,” implying atoms must magically coordinate to trick us. But the computational perspective we have developed clarifies what is at stake.

If the universe is a deterministic system that is compressible into stable macro-states without loss of predictive power, then high-level agents like us naturally emerge as computationally closed entities. This compressibility is not an *a priori* assumption but an empirical discovery. The evidence is straightforward: the special sciences work. Chemistry reduces to physics. Biology reduces to chemistry. The reductions succeed. This success is evidence that the universe has the hierarchical structure we describe.

Does this vindicate superdeterminism? Not entirely. Here is what it does and does not establish:

*What it establishes:* If superdeterminism is true, it need not involve “conspiracy.” The correlation between measurement choices and particle histories would be mediated by the same computational closure that makes agents possible in the first place. The physicist’s choice is determined by their macroscopic “software,” a computationally closed high-level boundary, which is shielded from arbitrary micro-details by Markov boundaries. The correlation is not spooky; it is structural, the same boundary-based shielding we have described throughout.

*What it does not establish:* That this structural shielding is *sufficient* for scientific inference. The critic may press: if hidden correlations exist between settings and outcomes, how do we know our statistical methods track real regularities rather than artifacts of the correlation? This is a genuine open problem. Our framework suggests where to look for an answer (in the formal conditions under which coarse-graining preserves causal structure), but we do not claim to have provided one. The point is narrower: superdeterminism, if true, is compatible with the emergence of genuine agents. Whether it is compatible with the epistemology of science remains contested.

Beyond physics, these ideas matter for practical intervention.

## Part 10: Real World Implications

These are not mere philosophical niceties. Mistaking the relevant causal level leads to predictable failures.

**Medicine:** Atrial fibrillation is an organ-level dynamical breakdown. In this condition, the micro-ontology (cells, proteins, ion channels) remains intact, yet the system fails because the electrical propagation pattern has lost coherence. Effective treatment involves ablating tissue to restore boundary integrity. Cardiologists do not manipulate individual cardiomyocyte ion channels; they alter the topological structure of electrical connectivity. This demonstrates that the causal variable responsible for cardiac function is not the sum of cellular parts but the topology of their connectivity. When intervention targets the topology to restore function, we empirically demonstrate that the object of medical science is the relational structure, not the aggregate of cells. A causal theory recognizing only cells cannot explain why topological constraints dictate the outcome. A nihilist ontology would literally not know where to intervene.

The pattern repeats across domains.

**Ecology:** When Pacific salmon populations collapsed in the Columbia River system, the initial response focused on the species level: hatcheries, breeding programs, fishing restrictions. But salmon are not isolated agents. They are nodes in an ecosystem with river flows, sediment patterns, predator-prey balances, and nutrient cycling. Interventions focused on individual fish failed because they treated salmon as independent units. The system has emergent stability

properties invisible at the species level. Restoration only succeeded when it targeted the ecosystem boundary: restoring flow dynamics of the watershed, reconnecting floodplains, reestablishing predator balances. When you intervene at the right level (treating the ecosystem as a causally autonomous unit), salmon populations recover. Optimize for individual fish while ignoring ecosystem dynamics, and you get expensive failure.

**Technology:** The internet has traffic patterns and robustness routers lack. Optimize routers while ignoring network properties, and you get cascading failures.

Each case demonstrates what we have formalized: intervention succeeds when targeting computationally closed levels.

**Quantum Challenge:** “Does entanglement not violate locality?” Response: Locality holds at the macroscopic scales where wholes emerge. Even quantum mechanics preserves statistical locality thermodynamically, as no information can be transmitted faster than light. Our argument applies where objects exist: above the quantum decoherence threshold.

These practical examples point toward a broader philosophical conclusion.

## Conclusion: The Hierarchy of Truth

We began with a hurricane and asked whether it was real. From four premises about causality, locality, information, and thermodynamics, we derived the inevitability of statistical boundaries. We showed these boundaries create causal autonomy, formalized the test for genuine wholes via computational closure, and demonstrated their practical indispensability. What picture of reality emerges?

Reality is not a flat soup of particles, nor an unordered jungle of patterns. It is a hierarchy, a lattice of computational machines stratified by scale. Science is the process of discovering which machine (which level of the lattice) effectively predicts the phenomenon at hand. When we find a level that is causally closed, we have hit bedrock. We have found a Whole that is as causally autonomous as its Parts.

What this argument establishes is that ontology ignoring causal-informational structure is incomplete. Once you accept that science tracks real patterns (in the sense of universality and causal autonomy), that patterns support intervention, and that interventions would fail without treating wholes as units, then whether we *call* them objects becomes secondary.

The nihilist position remains formally consistent but explanatorily barren. To deny the hurricane’s reality in the sense of causal autonomy after accepting its causal efficacy is to privilege etymological tradition over scientific practice. This position is not strictly inconsistent, but it is scientifically idle and explanatorily parasitic. Nihilism survives only by privileging fundamentality as the sole criterion of reality—a choice, not a discovery.

The hurricane we named, tracked, and fled was not a linguistic convenience. It was a pattern this universe’s constraints forced into existence.

The analysis suggests that disputes over the ‘existence’ of wholes often obscure a deeper point: scientific explanation and intervention require treating certain patterns as units, regardless of their metaphysical status. The relevant question is not “Do wholes exist?” but “What earns causal autonomy?” This is the question that matters for science and survival. Where causal powers, counterfactual



stability, and resistance to informational intrusion converge, objecthood is instantiated. Wholes are not fundamental, but they are metaphysically real as unified loci of causal powers. Any agent, human or otherwise, interacting with the atmosphere's constraints must eventually adopt a concept like "hurricane" to predict efficiently. A pattern is universally discoverable when any sufficiently capable agent attempting to predict and intervene within the same constraints is forced to encode it, on pain of incurring unsustainable algorithmic complexity. This universality is epistemic rather than metaphysical: it reflects convergence under shared constraints, not fundamentality, whether that pattern is the physical wall of a hurricane, the legal boundary of a corporation, or the psychological reality of a dragon. Each is a discoverable invariant in its own domain.

Note: Arguments addressed here are drawn from O'Connor's subscriber Q&As (1.25M and 1.75M). The  $\epsilon$ -machine formalization is from ; the Rainforest Realism and microbangings critique from .

## References

Rosas, Fernando E., et al. 2024. "Disentangling High-Order Mechanisms and High-Order Behaviours in Complex Systems." *Nature Physics* 20: 1095–1104. <https://doi.org/10.1038/s41567-024-02477-4>.