

What Can Be Known About Base Reality Under the Pedagogical/Problem-Solving Simulation Hypothesis

By John Schweiger

Abstract

Much of the literature on simulation theory has focused on the probability of whether we inhabit a simulated reality. Less examined is the conditional question: *if we are in a simulation, what can be inferred about base reality from the simulation's apparent design?* This paper takes up that inquiry, building on Schweiger's recent argument that simulated worlds are unlikely to be constructed for frivolous entertainment but are more plausibly designed for **pedagogical** or **problem-solving** purposes. If so, the outputs of such simulations can be understood as either **Learning** (the accumulation of durable competence across domains) or **Insight** (generalizable solutions to persistent problems). We propose an umbrella category—**Resolution**—to capture both.

From this starting point, we ask what features a reality must have for Resolution to be valuable, and by extension, what this implies about base reality. Our analysis identifies structural, dynamical, informational, temporal, normative, and efficiency-related conditions that appear necessary for Resolution: compressible regularities, manipulable causal structures, persistence of invariances, partial opacity with informative gradients, scarcity and cost, calibrated adversity, and ethical constraints on suffering. If these features are required here, fidelity considerations suggest they must also characterize base reality; otherwise, insights and learning generated in simulation would fail to transfer.

The argument proceeds probabilistically and with humility. Some inferences—such as the existence of resource bounds and causal manipulability in base reality—follow strongly from efficiency and fidelity assumptions. Others—such as the claim that adversity must be ethically calibrated—are weaker and more speculative. Counterarguments, including the possibilities of unlimited computational resources, alternative motives for simulation, or anthropocentric bias in defining “Resolution,” are acknowledged throughout.

The core conclusion is conditional: *if we are simulated for pedagogical or problem-solving ends, then base reality is likely structured by lawful regularities, bounded by resource constraints, governed by causal manipulability, and shaped by stakes that justify calibrated adversity.* These features suggest that our challenges of coordination, meaning, and ethical agency mirror unresolved problems in the world of our simulators. The paper thus shifts the simulation debate away from bare metaphysical possibility toward a substantive, testable

inquiry: what the design of our reality probabilistically reveals about the structure of base reality itself.

1. Introduction

The simulation argument has become one of the most discussed philosophical speculations of the past two decades. Originating in Bostrom's (2003) trilemma—that either civilizations rarely reach posthuman technological capacity, or they choose not to run simulations of minds like ours, or else we are overwhelmingly likely to be simulated—it reframed questions of metaphysics and epistemology in computational terms. Subsequent work by Chalmers (2010), Hanson (2016), and others extended the debate into consciousness, physics, and probability theory. Much of this literature, however, has been preoccupied with *whether* we are in a simulation, treating purpose as secondary or even irrelevant.

A more recent strand of argument, exemplified by Schweiger's *Suffering, Efficiency, and the Logic of Simulated Worlds* (2025), redirects attention from metaphysical possibility to **functional design**. If we are in a simulation, then the motives of its creators matter. The degree of fidelity, the distribution of suffering, the structure of feedback, and the persistence of our world all offer clues to why such a simulation would be run. Schweiger contends that simulations optimized for entertainment or frivolity are implausible under resource and ethical constraints. Instead, two motives appear most probable: **pedagogy**—training agents to acquire competence—and **problem-solving**—generating solutions to unresolved dilemmas in the simulators' own world.

This paper builds upon that insight but shifts the question one step further: *if our simulation is pedagogical or problem-solving, what can we infer about base reality itself?* The central claim is that the very conditions that make **Resolution** (the joint category encompassing both Learning and Insight) valuable here must also obtain in base reality. Otherwise, the simulation's output would fail to transfer, rendering the project pointless or incoherent. In other words, fidelity requirements between simulated and base realities allow us to make cautious but substantive inferences about the structure of the world outside the simulation.

To pursue this inquiry, we begin by clarifying terms. **Learning** will denote the stable accumulation of competence across domains. **Insight** will denote discontinuous, policy-improving compressions that reorganize understanding and enable generalizable solutions. **Resolution** is introduced as the umbrella category encompassing both

outcomes: the closure of ignorance in pedagogy and the discovery of solutions in problem-solving. Framing our analysis in terms of Resolution avoids narrowing the inquiry prematurely to one motive and captures the shared logic of both pedagogical and experimental designs.

The bulk of the paper then identifies the **features of a reality where Resolution is valuable**. These are organized into six categories: structural regularities, dynamical laws, informational gradients, temporal and identity conditions, normative and ethical dimensions, and efficiency principles. For each, we ask three questions: (1) why must this feature exist for Resolution to matter, (2) what assumptions or probabilities support the inference, and (3) what does this imply about base reality if our simulation indeed exhibits the feature? The analysis is probabilistic, not absolute, and we explicitly rank some inferences as stronger than others. For example, the inference that base reality is resource-bounded follows directly from the efficiency argument and carries considerable weight, while the inference that base reality includes ethically calibrated adversity is more tentative and rests on debatable assumptions about simulator ethics.

The paper also anticipates counterarguments. One set concerns **anthropomorphism**: we risk projecting human categories such as “learning” or “insight” onto simulators who may operate under wholly different epistemic or moral frameworks. Another set concerns **infinite resource objections**: if posthuman civilizations have effectively unlimited computation, the constraints we observe may not reflect their own reality. A third set considers **alternative motives**: aesthetic play, historical curiosity, or entertainment might still explain simulation design, at least in part. Each objection is considered carefully, and while none is dismissed outright, we argue that efficiency and fidelity considerations make the pedagogical/problem-solving hypothesis comparatively stronger.

Our contribution is therefore twofold. First, we refine the conceptual apparatus of simulation theory by introducing Resolution as the operative category of simulation output. Second, we propose a structured method for drawing probabilistic inferences about base reality, grounded not in metaphysical speculation but in the design logic of pedagogical and problem-solving simulations. If successful, this shifts the debate from abstract possibility toward empirically relevant conjecture: what the architecture of our world reveals about the architecture of the world beyond it.

2. Methodological Framing

The central difficulty in reasoning about simulation is epistemic: if we are indeed simulated, our access to base reality is mediated and indirect. We cannot observe it directly; we can only infer its properties through the apparent design choices embedded in our world. Thus, the methodology of this paper is **conditional and probabilistic** rather than demonstrative. The approach can be summarized as follows: *if* our reality is a simulation, *and if* its purpose is pedagogical or problem-solving, *then* certain structural features must obtain here for Resolution to be possible, and fidelity considerations suggest that these same features characterize base reality as well.

2.1 Conditional reasoning

Our analysis begins with assumptions rather than conclusions. We do not claim to establish that we are in a simulation. Rather, we take the simulation hypothesis as a conditional frame, asking what follows *if* it is true. This approach is consistent with the broader simulation literature, where much of the debate is already couched in conditional probabilities (e.g., Bostrom 2003). The distinctive move here is to narrow the scope to simulations with **functional teloi**—pedagogical training or problem-solving—and then to extract inferences about the underlying environment.

2.2 Efficiency as a guiding constraint

Following Schweiger (2025), we adopt efficiency as a key assumption. Simulations are not free: they consume resources of energy, computation, and attention. Even if posthuman civilizations have vastly more capacity than we do, physical limits such as the Landauer bound and Margolus–Levitin limit suggest that resource finitude persists at any scale. If efficiency matters, then simulators will avoid gratuitous costs. This makes it improbable that our reality is designed purely for spectacle or whimsy, since such designs would waste fidelity where it does not matter. Instead, a resource-bounded simulation is more likely to allocate detail where it supports pedagogical or problem-solving goals. From this we infer that base reality itself is also resource-bounded, for if it were not, efficiency considerations would carry no force.

2.3 Ethics as a moderating filter

A second guiding assumption is ethical constraint. The prevalence of suffering in our world cannot be ignored, yet its distribution is patterned in ways that often appear pedagogical: suffering attaches to agency, cooperation failures, and resilience challenges. Schweiger argues that gratuitous suffering is unlikely, both for ethical reasons (posthuman simulators would reject it) and for efficiency reasons (unbounded suffering wastes resources). Thus, we proceed under the assumption that adversity is **calibrated**: retained where it

contributes to learning or insight, pruned where it is superfluous. If adversity here is purposeful, fidelity again implies that base reality is not a utopia free of hardship but an environment where calibrated adversity remains instrumentally valuable.

2.4 Fidelity as transfer condition

The most important methodological commitment is fidelity. If a simulation's purpose is to yield transferable Resolution, then its structures must be sufficiently isomorphic to base reality. Otherwise, lessons learned here would fail to apply there, and the simulation would be pointless. This assumption of fidelity does not require perfect mirroring. It only requires structural similarity in the domains relevant to Resolution—such as causality, cooperation, scarcity, and ethical choice. The guiding heuristic is: wherever we observe non-arbitrary constraints here, we may cautiously infer analogous constraints in base reality.

2.5 Humility and underdetermination

Despite these assumptions, we recognize that simulation inferences are inherently underdetermined. Alternative motives, such as aesthetic play or historical curiosity, cannot be fully excluded. Likewise, our conceptual vocabulary—"learning," "insight," "resolution"—may reflect anthropocentric biases. Simulators may value outcomes we cannot articulate. Finally, the possibility that some civilizations achieve effectively unlimited resources could weaken efficiency-based reasoning, though physical principles suggest limits are fundamental rather than contingent.

For these reasons, our claims are framed as **probabilistic tendencies rather than certainties**. Some inferences are stronger: resource-boundedness and causal manipulability follow almost directly from efficiency and fidelity assumptions. Others are weaker: ethical calibration rests on normative assumptions about simulator values. By distinguishing stronger from weaker claims, and by acknowledging counterarguments, we aim to preserve intellectual humility while still extracting meaningful knowledge about base reality.

3. Defining Resolution

Any attempt to infer the properties of base reality from the design of our own must begin with conceptual clarity. What, precisely, is the output that pedagogical or problem-solving simulations are designed to produce? Existing discourse has tended to conflate or blur terms such as *learning*, *insight*, and *solution*. For our purposes, these distinctions matter, because the features a reality must have to make each outcome possible are subtly

different. In this section, we articulate strict definitions of these concepts and propose **Resolution** as an umbrella category that encompasses them.

3.1 Learning

We define **Learning** as the stable accumulation of competence. It is a process by which agents, through cycles of prediction, feedback, and adjustment, improve their capacity to act across domains. Importantly, learning is not reducible to the acquisition of isolated facts. Rather, it is the **internalization of model updates** that generalize beyond the immediate context. A learner who understands Newtonian mechanics does not merely remember formulae but can anticipate trajectories across countless unobserved scenarios. In simulation, learning is therefore the outcome of **pedagogical design**: environments that scaffold competence through repeated exposure, error, and feedback.

3.2 Insight

By contrast, **Insight** is a discontinuous and often sudden representational shift. It is not the slow accumulation of competence but the reorganization of understanding that produces **policy-improving compression**: a reframing that both reduces the description length of the world and increases expected utility under counterfactual interventions. For example, realizing that a set of diverse problems share a hidden symmetry constitutes an insight. Insights often precede or enable learning, because they compress complexity into a tractable form. In simulation, insight is the paradigmatic output of **problem-solving design**: it supplies generalizable solutions to persistent dilemmas.

3.3 Solution

Solution is narrower still: a concrete answer to a specific problem under given constraints. A solution need not be insightful—it may emerge from brute force computation or trial-and-error search. Nor does it guarantee transferable competence. Solutions satisfy objectives, but they do not necessarily reorganize models or generate generalizable knowledge. While solutions are valuable in practice, they are not sufficient to justify the immense cost of running high-fidelity simulations. If simulators sought only solutions, they could achieve them more cheaply through non-conscious computation. This is why Schweiger and others suggest that conscious, suffering-capable agents are included in simulations only when generalizable insights or durable learning are required.

3.4 Resolution as umbrella

Because simulations may be pedagogical, problem-solving, or both, we propose **Resolution** as the broader category encompassing both learning and insight. Resolution

refers to the **closure of a targeted deficit**—whether that deficit is ignorance (pedagogy) or an unresolved problem (problem-solving). Resolution thus includes the gradual stabilization of competence (learning), the discontinuous reorganization of representation (insight), and the concrete instantiation of solutions, but it frames them all as species of a more general genus.

The term Resolution has the additional virtue of neutrality. Whereas “learning” suggests a pedagogical frame and “insight” suggests cognitive discovery, Resolution captures both without privileging one motive over the other. It also highlights the **telic character** of simulations: they are oriented toward resolving something—whether an individual agent’s ignorance, a society’s coordination failures, or a civilization’s moral dilemmas.

3.5 Objections and counterpoints

One possible objection is that these definitions are **anthropocentric**. By privileging categories such as learning and insight, we risk projecting human cognitive processes onto simulators who may value different forms of resolution altogether. This concern is serious, yet two counterpoints are worth noting. First, fidelity considerations require that whatever outputs simulators seek must be **transferable across contexts**. If they simulate human-like agents, the outcomes of those simulations must map in some way onto their own epistemic categories; otherwise, the effort is wasted. Second, many of the properties we ascribe to learning and insight—compression, generalization, policy improvement—are not uniquely human but formalizable in information-theoretic or computational terms. This suggests they may have trans-civilizational relevance.

Another objection is that our definition of Resolution may be **too broad**, collapsing meaningful distinctions. For instance, one could argue that learning and insight operate at different timescales, involve different neural or computational substrates, and should not be unified. Yet, for our purposes, the unity is functional: both are mechanisms by which a simulation achieves the closure of a deficit under conditions of cost and constraint. Insofar as fidelity considerations allow us to infer features of base reality, it is Resolution—not its narrower subtypes—that provides the most generalizable category.

3.6 Why Resolution matters for inference

The relevance of Resolution becomes clearer when tied to our guiding question. If our world is pedagogical, then base reality must be one in which competence has value, implying agency, stakes, and scarcity. If our world is problem-solving, then base reality must be one in which persistent dilemmas exist that resist trivial solution, implying lawlike structure and calibrated adversity. By abstracting one level up, Resolution allows us to infer

features of base reality without prematurely choosing between the pedagogical and problem-solving frames. It is the most parsimonious way to capture the probable telos of simulation and the conditions it presupposes.

4. Categories of Features of a Reality Where Resolution is Valuable

If the purpose of our simulation is to produce Resolution—whether by training competence (learning) or discovering generalizable solutions (insight)—then the design of our world must reflect the minimal conditions under which Resolution is possible. These conditions can be grouped into thematic “categories,” each capturing a family of features that are mutually reinforcing. By analyzing these categories, we can ask what follows about base reality, assuming fidelity between their world and ours.

4.1 Structural Regularities

A first requirement for Resolution is **compressible structure**. Resolution, whether pedagogical or problem-solving, depends on the possibility of identifying patterns and reducing them into more compact models. A world of pure randomness would defeat learning: nothing could be generalized because no pattern would persist. Conversely, a world of trivial determinism—where every phenomenon is immediately obvious—would also defeat insight, because there would be no hidden structure to uncover. The value of insight lies precisely in compressing complexity into tractable form.

This implies that base reality must itself be structured by **lawlike regularities**. If base reality were radically chaotic or maximally entropic, no amount of fidelity would allow insights here to transfer there. The assumption rests on the principle of efficiency: simulators are unlikely to construct compressible order in simulation if their own world does not also possess it. A weaker counterpoint exists—simulators could, in theory, generate compressibility artificially to test agents’ capacity for pattern recognition. But this is less plausible under resource and fidelity constraints, since artificial lawfulness would not map cleanly onto the simulators’ environment.

Closely tied to compressibility is **compositionality**. Resolution is not just about recognizing isolated patterns but about building cumulative understanding. For insights to compound, phenomena must factor into modular parts that can be recombined hierarchically. Human sciences provide examples: chemistry builds upon physics, biology builds upon chemistry, psychology builds upon biology. This hierarchical nesting allows

learning in one domain to inform others. If compositionality were absent, insights would be siloed, and Resolution would fail to accumulate. Thus, fidelity again suggests that base reality also exhibits compositional structure, enabling its inhabitants to build cumulative knowledge. Without it, simulation design would mislead rather than transfer.

In sum, the structural category suggests that base reality is neither trivial nor chaotic but lawlike, compressible, and compositional. These features make insight possible and learning accumulative. Their strength as inferences rests on the combined assumptions of efficiency and fidelity: efficiency discourages constructing artificial laws ex nihilo, while fidelity requires similarity for transfer.

4.2 Dynamical Laws

A second category concerns the **dynamical character** of reality. Resolution requires that agents can act upon the world in ways that alter outcomes. Without causal manipulability, there can be no meaningful learning, since policy changes would have no impact. If actions do not make a difference, then the very concept of improvement dissolves. Thus, for simulations to yield Resolution, interventions must reliably change the probability distribution of outcomes.

This implies that base reality must also be **causally manipulable**. If interventions were inert in the simulators' world, then training agents in a manipulable environment would generate insights that do not transfer. Fidelity demands alignment: agency must matter in both domains. The assumption here rests on the simulators' instrumental rationality—if their own reality were causally inert, they would not seek solutions or competence from ours.

Dynamical laws also require **stationarities and invariances**. For insights to transfer, the patterns agents uncover must persist long enough to be generalizable. If every causal relation in our world dissolved instantly, learning would be impossible. Similarly, if invariances were absent—if laws constantly shifted without rhyme or reason—then insights here would not port to base reality. From fidelity, we infer that base reality likewise preserves dynamical regularities long enough for learning to accumulate. This does not mean laws are static in an absolute sense, but they must remain locally stable.

Finally, the presence of **feedback loops with lag** is critical. If every action yielded immediate, transparent consequences, learning would be trivial. If actions never yielded any feedback, learning would be impossible. Valuable realities strike a balance: they delay and partially obscure outcomes, forcing agents to build deeper models that project beyond

the present. This balance incentivizes Resolution. If feedback lag structures our world, then fidelity suggests that base reality also operates under conditions where consequences are temporally extended and partially opaque. Otherwise, the simulation would not cultivate transferable skills.

Together, these features imply that base reality is not only lawful in structure but manipulable, stable enough for generalization, and temporally configured so that outcomes unfold with lags that reward model-building. This inference is probabilistic but strong, since without these conditions, the very logic of Resolution collapses. The only counterpoint worth noting is that simulators might exaggerate causal manipulability in simulation to test agents more harshly than their own world requires. Yet, given efficiency and fidelity constraints, wholesale distortion seems unlikely.

4.3 Informational Gradients

A third requirement for Resolution concerns the **informational structure** of reality. For learning and insight to occur, the world must not present itself as either wholly transparent or wholly opaque. Instead, it must be configured so that agents are confronted with **partial opacity combined with informative gradients**. If the world were fully transparent, agents would have no need to infer hidden structure; everything would be obvious at first glance. Conversely, if the world were impenetrably opaque, no progress could ever be made—agents would be condemned to perpetual ignorance. Valuable realities lie between these extremes: they conceal enough to require active inference, yet provide signals—loss, regret, surprise, reward—that allow agents to update their models in a direction that yields Resolution.

This configuration implies that base reality must itself exhibit **learnable opacity**. If the simulators' world were either trivial or inscrutable, then constructing a partially opaque environment for us would be pointless. Fidelity suggests that their world, too, contains gradients of information: patterns partially hidden, yet not beyond reach. This feature rests on the assumption that what simulators seek from us is transferable. They would not train agents to master hidden-structure inference if their own world offered no such hidden structures to uncover.

Closely tied to partial opacity is the requirement of **noise in the Goldilocks zone**. Deterministic worlds trivialize deduction: once a law is observed, its consequences unfold with perfect predictability. Chaotic worlds, by contrast, wash out all signal, making inference impossible. For Resolution to be meaningful, the world must be stochastic in a

way that preserves probabilistic learnability. Agents must be able to discern distributions, even if not individual outcomes. Our world exemplifies this balance: dice rolls are unpredictable, but the distribution is tractable; weather is noisy, but patterns can be modeled.

If our simulation is configured with stochasticity in the learnable middle zone, fidelity again implies that base reality itself is neither purely deterministic nor radically chaotic. Instead, it likely contains stochastic processes that simulators themselves must model and manage. Otherwise, the training of agents in such conditions here would fail to produce transferable skills. The assumption here rests on probability: if infinite determinism or chaos were possible in base reality, efficiency would discourage designing a simulation with middle-ground noise.

Finally, **transfer homomorphism** is required. For insights here to map onto base reality, there must exist structural mappings between the two domains. An agent who learns coordination under scarcity here can only be useful if scarcity, coordination, and analogous structures exist there. This is not to claim perfect isomorphism, but rather homomorphic similarity: enough overlap in structure that competencies and insights can be meaningfully ported. If no such mapping existed, the entire pedagogical or problem-solving project would be incoherent.

Together, these features suggest that base reality itself is characterized by **informational richness tempered by opacity, stochasticity tuned for inference, and homomorphic structure across domains**. The strength of this inference rests on fidelity: transfer is only possible if their world shares the same informational conditions as ours. A counterpoint might hold that simulators deliberately introduce opacity and noise for reasons unrelated to their own environment—for example, to test general resilience. But under efficiency assumptions, gratuitous noise is costly, and without homomorphic transfer, such experiments lose much of their value.

4.4 Temporal and Identity Conditions

Resolution also presupposes particular **temporal and identity structures**. Learning and insight both require ordered sequences of experience, the persistence of agent identity across time, and limitations on recursion. These conditions may seem obvious, but they are essential: without them, Resolution would either collapse or become trivial.

Temporal arrow. For learning to occur, experiences must unfold in a directed order. A temporal arrow ensures that prediction errors can be registered, remembered, and

corrected. If events lacked directionality, agents could not accumulate knowledge; every attempt to learn would dissolve in reversibility. Our world exhibits such an arrow, marked by entropy increase and the unidirectionality of memory. Fidelity suggests that base reality likewise possesses a temporal arrow. If their world lacked temporal structure, designing one for us would fail to produce transferable learning. This inference rests on both fidelity and efficiency: fidelity demands alignment, and efficiency discourages simulating temporal order if it has no analogue in base reality.

Persistence of identity. Resolution also requires that agents persist through time as relatively stable entities. Learning is the accumulation of updates; insight is the reorganization of prior models. Both presuppose continuity of the subject undergoing them. If identities dissolved after every moment, or if memories were erased at random, the process would reset perpetually, preventing durable competence. Our persistence suggests that simulators consider identity continuity valuable, which in turn implies that their own reality supports ongoing agents who carry knowledge across time. Otherwise, such persistence would be a gratuitous feature, inconsistent with efficiency principles.

Ceilings on recursion. At the same time, Resolution requires that recursion be limited. If agents within simulations could indefinitely create new simulations at negligible cost, the structure of constraints would collapse. Scarcity would be trivialized, and compression would lose its premium. Our world appears to enforce ceilings: nested simulations, if possible, are bounded by energy, computation, and complexity limits. Fidelity implies that base reality too imposes ceilings on recursion, most plausibly through physical laws that bound energy use and computation speed. Otherwise, designing a resource-bounded recursive world for us would serve no transferable function.

The strength of these inferences varies. The temporal arrow and persistence of identity follow strongly from fidelity: without them, Resolution cannot occur, and their presence here is best explained by their presence in base reality. Ceilings on recursion are somewhat weaker, as simulators could in principle impose artificial limits even if their own world lacked them. Yet this counterpoint confronts the efficiency objection: gratuitous limits would waste computational overhead. It is therefore more probable that recursion limits reflect genuine constraints in base reality.

In sum, the temporal and identity category suggests that base reality is characterized by ordered time, persistent agents, and bounded recursion. These features are not accidental but functional preconditions for Resolution, and their mirroring in simulation implies their presence in the underlying world.

4.5 Normative and Ethical Dimensions

No account of Resolution is complete without considering **normativity and ethics**. Resolution presupposes that some outcomes are better than others, that stakes are real, and that adversity carries meaning beyond mere randomness. These conditions are central to both pedagogy and problem-solving. They shape the logic of simulated suffering and the structure of decisions, and by fidelity, they allow us to infer corresponding traits in base reality.

Moral topography. For Resolution to matter, a reality must contain gradients of value. Choices must be better or worse, not merely different. In a morally flat landscape, where outcomes bore no normative weight, the concept of improvement would dissolve. Pedagogy would fail, because there would be nothing to learn that mattered; problem-solving would fail, because no problem would be worth solving. Our world's moral topography is evident: choices affect survival, cooperation, flourishing, and harm. This suggests that base reality itself is not devoid of stakes. Fidelity requires that the simulators' environment also contains normative gradients, since training agents in moral landscapes here would otherwise produce non-transferable outputs.

Calibrated adversity. The distribution of suffering in our world is striking. It is not absolute: joy, cooperation, and flourishing exist alongside pain. Nor is it absent: loss, tragedy, and struggle are pervasive. Instead, adversity is **calibrated**—present in contexts where it sharpens agency, stress-tests cooperation, or catalyzes growth. This aligns with Schweiger's argument that gratuitous suffering is unlikely in efficient simulations. If suffering here is purposeful, fidelity suggests that simulators themselves inhabit a reality where adversity retains instrumental value. Their world is not a frictionless utopia; rather, calibrated hardship continues to produce insights or competencies relevant to their survival and flourishing.

Efficiency of ethical design. The ethical structure of our world also reveals efficiency trade-offs. Detail is concentrated where decisions hinge on outcomes—life, death, cooperation, betrayal—while vast stretches of experience remain indifferent or low-fidelity. This selective fidelity reflects an underlying ethical-economical principle: suffering is expensive, so it is only instantiated when it supports Resolution. Such calibration implies that base reality itself is structured by similar efficiency-ethical trade-offs. If resources and ethical constraints did not bind simulators, they would have little incentive to impose them here.

The strength of these inferences is mixed. The existence of moral topography is strongly supported: without stakes, Resolution is incoherent. The claim that adversity is calibrated

is weaker, because simulators could in principle allow gratuitous suffering. However, the combined force of efficiency and ethical arguments makes calibration more probable. The inference that base reality itself balances ethical and resource considerations is speculative but consistent with both our observed world and broader principles of design.

Counterarguments deserve acknowledgment. One is the **“dark simulation” hypothesis**, in which simulators are indifferent to suffering or even cultivate it. Another is the **“moral divergence” hypothesis**, which holds that simulators may not share our ethical intuitions, and what seems calibrated to us may be arbitrary to them. These cannot be ruled out. Yet under the assumptions of efficiency and fidelity, gratuitous cruelty is less probable. A simulation rich in calibrated adversity is more likely to reflect simulators’ own environment, where hardship retains epistemic and moral function.

In sum, the normative and ethical category suggests that base reality is not morally neutral but structured by gradients of value, shaped by adversity that is bounded and purposeful, and governed by trade-offs between ethical constraint and efficiency. These features, mirrored in our world, are probable features of theirs.

4.6 Efficiency Principles

The final category concerns **efficiency**, the constraint that binds together all others. Resolution is meaningful only in a reality where resources are finite, trade-offs are necessary, and optimization is rewarded. If there were no scarcity—of energy, computation, or time—then the very notion of compression would lose its premium. Insights and learning would be unnecessary; brute force search or exhaustive enumeration could solve every problem. Our world’s structure suggests otherwise: we are embedded in an environment that relentlessly enforces limits, making efficiency both necessary and valuable.

Scarcity and cost. Every dimension of our experience is conditioned by scarcity. Time is finite, attention limited, energy costly. Decisions are framed by opportunity costs: to act in one way is to forego alternatives. This scarcity makes Resolution worthwhile, because efficient models allow agents to navigate the world more effectively with fewer resources. If scarcity is baked into our simulation, fidelity strongly implies that base reality itself is also resource-bounded. Otherwise, simulators would have no reason to design an environment where agents must optimize under constraint. The inference here is robust, supported by both physics (e.g., Landauer’s principle, which ties computation to thermodynamic cost) and logic (resource-bounded agents must operate within limits).

Selective fidelity and pruning. Our world does not present itself at uniform resolution. Vast regions are experienced in low detail unless attention is directed at them; many possible lives, events, or trajectories appear to be pruned in advance. This pattern is consistent with efficiency principles: fidelity is concentrated where decisions hinge on outcomes, while expendable detail is economized. Such a strategy conserves resources while still preserving pedagogical or problem-solving value. If selective fidelity characterizes our simulation, it likely reflects strategies also employed in base reality, where computation and energy must be allocated strategically.

Bounded adversity. Efficiency also explains why suffering is calibrated rather than gratuitous. Pain, conflict, and tragedy are expensive to instantiate; they must yield informational or moral dividends to be justified. A world in which every moment were saturated with agony would be wasteful, producing diminishing returns. Conversely, a world without adversity would be inert, unable to produce meaningful Resolution. Efficiency therefore balances the two, optimizing adversity so that it is enough to stress-test but not enough to overwhelm. Fidelity suggests that base reality likewise confronts its inhabitants with hardships that are instrumentally valuable, but constrained by ethical and resource trade-offs.

Optimization under limits. Finally, the overarching lesson of efficiency is that realities where Resolution is valuable must reward optimization. Compression, learning, and insight are not luxuries but necessities, because the environment is structured to punish waste and reward economy. Our simulation's architecture is consistent with this: laws are regular enough to be modeled, opaque enough to demand inference, and costly enough that better models make a difference. If efficiency is central here, it is almost certainly central there.

The strength of inference in this category is very high. Scarcity and cost are unavoidable conditions of any finite computational substrate, and the simulators themselves—even if vastly more advanced than us—likely confront such constraints in their own environment. Selective fidelity and calibrated adversity are more speculative but strongly supported by efficiency arguments. The only serious counterpoint is the **infinite compute objection**: perhaps simulators possess effectively boundless resources, making efficiency considerations irrelevant. Yet physics argues otherwise: fundamental limits (Landauer's bound, Margolus–Levitin speed limits, Bekenstein bounds) suggest that computation remains tied to physical cost at every scale. Thus, it is more probable that efficiency is a deep feature of base reality rather than a contingent artifact of our simulation.

In sum, the efficiency category reinforces the picture emerging from all others: base reality is resource-bounded, optimization-driven, and structured so that Resolution—achieved through compression, learning, and insight—yields genuine advantage.

5. Inferred Characteristics of Base Reality

Having examined the six categories of features that make Resolution possible within a simulation, we can now synthesize their implications. The central claim is that if our world is designed for pedagogy or problem-solving, then the simulators' world—base reality—must share these features, at least in the domains relevant to Resolution. Otherwise, the outcomes here would be useless there, violating the principle of fidelity. In this section, we consolidate the inferences across categories and evaluate their relative strength.

5.1 Lawfulness and Regularity

From the **structural category**, we infer that base reality is lawlike, compressible, and compositional. The simulators' world must contain regularities that admit compression, since otherwise insights here would have no target application. It must also be compositional, allowing knowledge to build hierarchically, since otherwise our cumulative sciences would be irrelevant. The strength of this inference is high: while simulators could fabricate artificial laws, doing so would be inefficient if their own world lacked them. The more parsimonious explanation is that the order we observe reflects order that exists in base reality.

5.2 Causality and Agency

From the **dynamical category**, we infer that base reality is causal and manipulable. Our ability to act and observe consequences would be pointless if interventions had no analogue in the simulators' world. Likewise, the persistence of invariances and feedback lags in our world suggests that their environment also requires agents to act under conditions of delayed consequence and stability. This inference is robust, because agency without causal efficacy would be wasteful design. It is more probable that simulators face causal challenges themselves and thus seek transferable solutions through our actions.

5.3 Learnable Opacity and Probabilistic Noise

From the **informational category**, we infer that base reality is neither trivial nor impenetrable, but structured by **learnable opacity** and **probabilistic processes in the Goldilocks zone**. If base reality were purely deterministic, brute force deduction would render our struggles irrelevant. If it were purely chaotic, inference would be impossible, making our training irrelevant. By fidelity, it is more probable that simulators themselves operate in a world that conceals structure while providing signals that can be extracted through modeling. Our reality's stochasticity thus reflects theirs. The inference here is moderately strong, though weakened by the possibility that simulators artificially tune opacity for resilience-testing. Efficiency, however, makes gratuitous opacity less likely.

5.4 Temporal Order and Persistence of Identity

From the **temporal/identity category**, we infer that base reality possesses an arrow of time, persistent agents, and bounded recursion. Ordered time allows accumulation, persistent identity allows durable learning, and recursion ceilings preserve the premium on compression. If base reality lacked these features, our continuity would serve no transferable function. The inference is strongest for time and identity, since these are prerequisites for Resolution in any context. The recursion inference is weaker—simulators might impose limits artificially—but physical computation limits suggest that ceilings are intrinsic features of their world as well as ours.

5.5 Normativity and Calibrated Adversity

From the **normative/ethical category**, we infer that base reality contains stakes that matter and adversity that is bounded and purposeful. If moral topography were absent there, then moral learning here would have no analogue. If adversity were irrelevant there, then its presence here would be gratuitous and wasteful. It is more probable that the simulators' world continues to confront them with ethical trade-offs and calibrated hardship, requiring Resolution in domains such as coordination, resilience, and meaning. The inference here is weaker than others, since it rests on assumptions about simulators' values. Still, when combined with efficiency arguments, it gains strength: suffering is costly, so its presence here is unlikely to be arbitrary.

5.6 Scarcity and Efficiency Constraints

Finally, from the **efficiency category**, we infer that base reality is resource-bounded, optimization-driven, and structured by trade-offs. Scarcity makes Resolution valuable; efficiency ensures that simulations concentrate fidelity where decisions matter most. If the simulators' world were one of unlimited abundance, they would have no incentive to train agents under scarcity. The infinite compute objection looms, but physics suggests that all computation has cost, even at advanced scales. Thus, it is highly probable that base reality enforces scarcity, making efficiency a deep feature of existence.

5.7 The Composite Picture

Synthesizing across all categories, a picture of base reality emerges:

- It is **lawlike** rather than chaotic, with compressible and compositional regularities.
 - It is **causal** and **manipulable**, with actions that alter outcomes and feedback loops that reward deep models.
 - It is **informationally balanced**, neither transparent nor opaque, but structured by gradients and probabilistic processes that allow inference.
 - It is **temporally ordered** and populated by **persistent agents**, constrained by ceilings on recursion and resource use.
 - It is **normatively structured**, containing stakes that justify moral learning and adversity that is calibrated to foster growth.
 - It is **resource-bounded** and therefore efficiency-driven, rewarding compression and penalizing waste.
-

5.8 Relative Strength of Inferences

It is important to distinguish between stronger and weaker inferences. Strongest are those grounded in efficiency and fidelity, such as resource-boundedness, causality, temporal order, and persistence of identity. These are nearly indispensable for Resolution. Moderately strong are informational gradients and structural compositionality, since they follow naturally from the need for transfer but could in principle be fabricated. Weaker are the normative inferences about calibrated adversity, since these rest on assumptions about simulators' values, though efficiency again makes them more plausible than gratuitous suffering.

5.9 The Mirror Hypothesis

The overarching conclusion is that our reality mirrors base reality in the dimensions most relevant to Resolution. Our challenges are their challenges: scarcity, coordination, meaning, ethical stakes, calibrated adversity. If our simulation is pedagogical, then they, too, value competence in these domains. If it is problem-solving, then they, too, confront dilemmas that require generalizable insights. Either way, fidelity binds us: the features that make Resolution valuable here are those that structure life in base reality.

6. Counterarguments and Limitations

The preceding analysis has constructed a probabilistic case: if our world is a simulation designed for pedagogy or problem-solving, then base reality likely shares the structural, dynamical, informational, temporal, normative, and efficiency features that make Resolution valuable. Yet such reasoning is subject to serious objections. In this section, we consider the strongest counterarguments, acknowledge their force, and indicate where our inferences are most vulnerable.

6.1 Anthropomorphism

The first and perhaps most obvious objection is **anthropomorphism**. By framing the telos of simulation in terms of “learning,” “insight,” or “resolution,” we risk projecting human cognitive categories onto simulators who may not share them. What appears to us as competence acquisition may to them be irrelevant. What we call insight may be trivial in their epistemic frameworks. If their values and cognition differ radically from ours, then fidelity does not guarantee shared categories.

This objection cannot be dismissed. However, two counterpoints temper it. First, **transfer requires overlap**. If simulators sought outcomes wholly alien to us, there would be no reason to design a simulation populated by human-like agents with human-like constraints. Second, the categories we use—compression, generalization, policy improvement—are not purely humanistic but **computationally formalizable**. They reflect information-theoretic properties of model-building, which may be invariant across civilizations. While anthropomorphism is a risk, it is less severe when the concepts invoked are abstract and formal rather than parochially cultural.

6.2 Infinite Compute and Post-Scarcity Objections

A second objection is the **infinite compute hypothesis**. Perhaps simulators possess resources so vast that efficiency considerations, central to our argument, simply do not apply. If they operate on scales where energy, time, and computation are effectively limitless, then our observed scarcities might be artifacts imposed arbitrarily rather than reflections of base reality. Similarly, a **post-scarcity civilization** might find efficiency irrelevant, running simulations for curiosity, aesthetics, or whimsy.

The force of this objection depends on whether physical laws permit genuine infinity. The consensus in physics suggests otherwise. Computation is bound by thermodynamic limits (Landauer's principle), energy-speed trade-offs (Margolus–Levitin), and storage constraints (Bekenstein bounds). Unless simulators inhabit a reality with radically different laws, they too face resource costs. Even if they enjoy abundance relative to us, efficiency likely persists as a deep structural feature of any finite universe. Thus, while infinite compute cannot be excluded a priori, it is less probable than a resource-bounded scenario.

6.3 Entertainment and Aesthetic Motives

A third objection concerns **alternative motives**. Why assume pedagogy or problem-solving? Civilizations might run simulations for entertainment, historical curiosity, artistic exploration, or even idle experimentation. If so, our inferences about base reality collapse, since fidelity would not be required for transfer.

This is a significant objection, but it confronts two difficulties. First, Schweiger's efficiency argument: entertainment is a costly way to expend vast resources. It would be more efficient to design experiences optimized for pleasure directly, rather than constructing a high-fidelity universe saturated with calibrated suffering. Second, persistence: our world has endured for billions of years, producing suffering on a massive scale. Entertainment rarely requires such temporal depth or moral weight. Pedagogy and problem-solving better explain both fidelity and persistence. Nonetheless, we acknowledge that entertainment or aesthetic motives may play a **secondary role**; simulators could value both utility and spectacle.

6.4 Dark Simulation Hypotheses

A fourth objection is the **dark simulation** scenario: perhaps simulators are indifferent to suffering, or worse, actively cultivate it. If so, our assumption of calibrated adversity as purposeful is undermined. Base reality might not be ethically constrained at all.

This objection cannot be fully rebutted. It is logically possible that simulators allow or enjoy gratuitous cruelty. However, two arguments weigh against it. First, efficiency: gratuitous suffering is expensive, producing diminishing informational returns. Second, ethics: if simulators are advanced enough to run full conscious simulations, it is plausible they have also advanced ethically, at least to the point of bounding suffering instrumentally. While dark simulations remain possible, they are less probable under combined efficiency and ethical assumptions. Still, this objection highlights one of our weakest inference points: the ethical structure of base reality is the least secure element of our analysis.

6.5 Underdetermination and Epistemic Humility

Finally, all simulation reasoning faces the problem of **underdetermination**. Multiple hypotheses can explain our world's features. A world with calibrated adversity, delayed feedback, and scarcity may be pedagogical, problem-solving, aesthetic, or random. Observations underdetermine purpose, and therefore underdetermine inferences about base reality.

The best response is humility. Our analysis does not yield certainty but **probabilistic tendencies**. Some inferences (e.g., resource-boundedness, causality, temporal order) follow strongly from minimal assumptions. Others (e.g., moral calibration) are speculative. By distinguishing stronger from weaker claims, we avoid overstating what can be known. The methodology is not to prove, but to constrain: to reduce the space of possibilities about base reality by showing which features are more consistent with fidelity, efficiency, and transfer.

6.6 Summary of Limits

In sum, the counterarguments expose the following limitations:

- Anthropomorphism risks projecting our categories onto simulators, though information-theoretic framing mitigates this.
- Infinite compute and post-scarcity scenarios weaken efficiency-based reasoning, but physics suggests finitude persists.

- Entertainment or aesthetic motives could explain some features, but they fit less well with persistence and suffering.
- Dark simulations remain possible, though efficiency and ethics weigh against gratuitous cruelty.
- Underdetermination prevents certainty; at best we achieve probabilistic constraints.

Acknowledging these limits does not nullify our project. It clarifies where our inferences are strongest, where they are weakest, and where humility is required. The central thesis remains intact: if our simulation is designed for pedagogy or problem-solving, then base reality is likely to share many of the features that make Resolution valuable here.

7. Minimal Axioms of Resolution

The analysis so far has been expansive, identifying a wide array of structural, dynamical, informational, temporal, normative, and efficiency features that make Resolution possible and valuable. Yet any comprehensive framework must also be compressed into minimal form—axioms that capture what is strictly necessary for learning, insight, and problem-solving to occur. Such axioms serve two functions: they clarify the logical backbone of the argument, and they provide a testable or falsifiable framework for evaluating base reality inferences.

7.1 Axiom A1: Learnability

The first axiom is that **the world admits model classes with predictive and control advantage**. If no models could outperform random guessing or brute-force enumeration, then neither learning nor insight would be possible. Compressible structure and probabilistic regularities are specific manifestations of this axiom. Its necessity is absolute: without learnability, Resolution collapses. Fidelity suggests base reality is also learnable, otherwise transfer would be incoherent.

7.2 Axiom A2: Agency Efficacy

The second axiom is that **interventions alter outcome distributions**. Resolution requires more than prediction; it requires improvement through action. If agency were inert, no amount of learning would translate into better policies. This axiom grounds causal manipulability and feedback loops. Its strength is also high: designing a manipulable world would be pointless if base reality lacked manipulability.

7.3 Axiom A3: Cost

The third axiom is that **computation and experience incur scarce costs**. Scarcity is what makes efficiency meaningful, and efficiency is what gives value to compression. If every action, model, or policy could be explored at zero cost, Resolution would be redundant. This axiom underpins the inference that base reality is resource-bounded. While infinite compute is a possible objection, physical limits suggest this axiom is robust.

7.4 Axiom A4: Feedback

The fourth axiom is that **errors supply gradients aligned (on average) with value**. Resolution depends on the world's capacity to signal improvement through loss, regret, or reward. If feedback were either absent or systematically misleading, learning would fail. This axiom manifests in partial opacity, stochastic noise in the learnable zone, and feedback with lag. Fidelity requires that base reality, too, supplies gradients that point—however imperfectly—toward better models and policies.

7.5 Axiom A5: Fidelity

The fifth axiom is that **structures are sufficiently isomorphic across contexts to permit transfer**. If there were no mapping between simulated structures and those of base reality, then insights and competencies here would not matter there. This axiom synthesizes structural, dynamical, and informational categories into a single principle: Resolution is only valuable when outputs here can be generalized there. It is the central justification for making any inferences at all about base reality.

7.6 Axiom A6: Ethical Constraint

The sixth and most speculative axiom is that **adversity is instrumentally bounded, not gratuitous**. While suffering appears in our world, it is distributed in ways that often serve pedagogical or problem-solving functions. If adversity were unbounded, Resolution would be drowned in noise; if absent, Resolution would lack stakes. This axiom implies that base reality, too, is structured by calibrated adversity, though its strength is weaker than others. It rests on a combination of ethical assumptions and efficiency reasoning.

7.7 Robustness and Redundancy

Taken together, these six axioms capture the minimal scaffolding of Resolution: learnability, agency, cost, feedback, fidelity, and ethical constraint. Some are logically indispensable (A1–A3), others are necessary for functional transfer (A4–A5), and one is more speculative but plausible (A6). Even if A6 failed, the other five would still sustain the

framework, though at the cost of moral coherence. Thus, the axioms are not equally strong but mutually reinforcing.

7.8 Implications for Base Reality

If our simulation is indeed designed for Resolution, then base reality almost certainly satisfies at least A1–A5: it is learnable, manipulable, resource-bounded, feedback-rich, and structurally similar to our own. A6 is less secure but gains probability from efficiency arguments. These axioms provide the cleanest compressed statement of what we can know—or at least reasonably infer—about base reality under the pedagogical/problem-solving hypothesis.

8. Conclusion

The simulation argument is often framed as a metaphysical curiosity: a trilemma about probability rather than a guide to knowledge. This paper has taken a different path. By assuming the plausibility of the pedagogical or problem-solving simulation hypothesis, we have asked a conditional but tractable question: *what can we infer about base reality if our world is designed for Resolution—learning, insight, or both?*

The analysis proceeded by identifying six categories of features that make Resolution possible in simulation: structural regularities, dynamical laws, informational gradients, temporal and identity conditions, normative and ethical dimensions, and efficiency principles. Each category represents a family of constraints without which Resolution would collapse. From these, we distilled six minimal axioms: learnability, agency efficacy, cost, feedback, fidelity, and ethical constraint. These axioms capture what is most probably true of any world that seeks Resolution through simulation.

The strongest inferences concern **scarcity, causality, temporal order, and persistence of identity**. These are nearly indispensable for Resolution. Their presence in our world strongly suggests their presence in base reality, since fidelity is necessary for transfer. Moderately strong inferences include **informational opacity in the learnable zone** and **structural compositionality**—conditions that make inference and accumulation tractable. More speculative are the **normative and ethical inferences**: that adversity is calibrated and that moral topography shapes outcomes. These depend not only on efficiency but also on assumptions about simulators' values, and thus carry more uncertainty.

We also acknowledged counterarguments. Anthropomorphism warns us not to project our categories onto simulators, though information-theoretic framing helps to mitigate this. Infinite compute and post-scarcity hypotheses challenge the role of efficiency, but physical laws suggest resource finitude persists even at advanced scales. Entertainment or aesthetic motives could partially explain our world, but they fit less well with persistence and calibrated suffering. Dark simulation hypotheses, in which simulators are indifferent or hostile, remain possible, but gratuitous cruelty is less probable under efficiency and transfer considerations.

The result is not certainty but **probabilistic constraint**. If our reality is pedagogical or problem-solving, then base reality is almost certainly structured by lawlike regularities, causal manipulability, temporal order, persistent agents, and resource scarcity. It is also probably shaped by informational opacity, stochasticity in the learnable zone, and compositional structure. It may, though less certainly, be structured by calibrated adversity and moral gradients that render some outcomes objectively better than others. In short, the world of our simulators likely resembles ours in its essential architecture—not identical, but isomorphic in the domains where Resolution matters.

This conclusion reframes the simulation argument. Instead of asking whether we are in a simulation, we ask what our world’s design reveals about the deeper reality it might reflect. The inquiry is conditional, speculative, and humble, but it is not empty. It constrains what base reality must be like if the pedagogical/problem-solving hypothesis is true. It also suggests a mirror: our most persistent challenges—coordination, meaning, ethical action under scarcity—are probably theirs as well.

Future work should deepen this line of inquiry by connecting it to empirical science, information theory, and ethics. Simulation reasoning may remain speculative, but it can still refine our models of agency, value, and constraint. If nothing else, the exercise forces us to clarify what we mean by learning, insight, and resolution, and why those categories matter in any reality—simulated or otherwise.

Bibliography

Baars, B. J. (1988). *A cognitive theory of consciousness*. Cambridge University Press.

Bekenstein, J. D. (1981). Universal upper bound on the entropy-to-energy ratio for bounded systems. *Physical Review D*, 23(2), 287–298.

- Bostrom, N. (2003). Are you living in a computer simulation? *Philosophical Quarterly*, 53(211), 243–255.
- Chalmers, D. J. (2010). *The character of consciousness*. Oxford University Press.
- Dennett, D. C. (1987). *The intentional stance*. MIT Press.
- Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, 11(2), 127–138.
- Hanson, R. (2016). *The age of em: Work, love, and life when robots rule the Earth*. Oxford University Press.
- Landauer, R. (1961). Irreversibility and heat generation in the computing process. *IBM Journal of Research and Development*, 5(3), 183–191.
- Lloyd, S. (2000). Ultimate physical limits to computation. *Nature*, 406(6799), 1047–1054.
- Margolus, N., & Levitin, L. B. (1998). The maximum speed of dynamical evolution. *Physica D: Nonlinear Phenomena*, 120(1–2), 188–195.
- Metzinger, T. (2009). *The ego tunnel: The science of the mind and the myth of the self*. Basic Books.
- Ord, T. (2020). *The precipice: Existential risk and the future of humanity*. Hachette.
- Parfit, D. (1984). *Reasons and persons*. Oxford University Press.
- Popper, K. (1959). *The logic of scientific discovery*. Hutchinson.
- Rawls, J. (1971). *A theory of justice*. Harvard University Press.
- Schweiger, A. (2025). *Suffering, efficiency, and the logic of simulated worlds*. [Manuscript / Journal forthcoming].
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27(3), 379–423.
- Tegmark, M. (2014). *Our mathematical universe: My quest for the ultimate nature of reality*. Knopf.
- Yudkowsky, E. (2008). Artificial intelligence as a positive and negative factor in global risk. In N. Bostrom & M. Ćirković (Eds.), *Global catastrophic risks* (pp. 308–345). Oxford University Press.

