

A Two-Level Bayesian Analysis of Reality, Consciousness, and the Physical World

September 20, 2026

A Bayesian Comparison of Ultimate Ontology and Local Reality

Rodney Malpert, J.D./M.A., Cornell; with ChatGPT (GPT-5.6 Sol)

RodMalpert.com

Adversarial contributions from Claude and Grok

Abstract

This paper compares physicalism, idealism, Russellian monism, and dualist theories of mind using an explicit Bayesian framework. The model separates evidence that bears on ultimate ontology from evidence that bears only on the local causal architecture of the world humans inhabit. That distinction prevents the success of local physical science from being treated as automatic proof that the observed physical level is metaphysically fundamental or causally exhaustive.

Tier 2 asks what reality is ultimately; Tier 1 asks whether the observed physical world is the bottom level or instead simulated, generated, represented, or otherwise dependent on a deeper reality. Simulation is therefore treated as compatible with several ultimate ontologies. The model also conditions generated-consciousness versus participatory simulation on the theory of consciousness rather than assuming that computation automatically produces phenomenality.

The consciousness analysis includes two often-neglected boundary positions. Eliminative materialism and illusionism are treated as internal physicalist strategies: they deny or revise the phenomenal properties that generate the hard problem rather than adding a separate ontology. Dualism is partitioned between interactionist/separable theories and epiphenomenal property dualism, because the latter accepts physical causal closure and detailed brain dependence while retaining irreducible phenomenal properties.

The evidential perimeter includes detailed neuroscience, causal closure, psychophysical harmony, quantum and spacetime considerations, fine-tuning, artificial consciousness, anomalous-consciousness reports, and evidence considered in a separate study of postmortem survival. That companion study, *A Bayesian Analysis of Life After Death*, asks a different question from the present paper and uses a distinct local-architecture model; its central estimate for meaningful conscious continuation is about 47%, but that result is not fed mechanically into the ontology calculation here. DMT entity encounters, psychedelic synesthesia, a reported laser-code phenomenon, and claimed telepathy involving nonspeaking autistic children are assessed under ordinary evidential standards rather than either excluded as fringe or credited without controls.

Under the central partition, Russellian monism remains the leading ultimate ontology at roughly 40%, followed by physicalism at roughly 26%, idealism at roughly 23%, and dualism in the aggregate at roughly 10%. These are model-conditioned judgments, not measured frequencies. A dedicated review of the professional simulation literature in philosophy, physics, and computer science strengthens the case that simulation should be treated as a serious Tier-1 hypothesis while also exposing substantial reference-class, consciousness, motivation, and testability objections. Physics supplies modest support for local nonfundamentality and only a small simulation-specific increment; existing experimental probes mainly constrain narrow implementations. The principal incremental numerical update is temporal selection: conditional on a high probability that present civilization is within at most a few decades of a first artificial-superintelligence transition, an observer-moment near that focal transition is more expected under simulation if such epochs are preferentially rerun. Using a 30% simulation baseline that already incorporates Bostrom-style counting and broad nonfundamentality considerations, a deliberately conservative incremental Bayes factor of about 1.7 yields a central simulation probability of about 42%. Coherence then places local nonfundamentality near 50% and local intelligence- or mind-priority near 45%. The separate postmortem-survival estimate remains about 47% because its branch structure must be recalibrated independently.

Central results

Question / hypothesis	Central estimate
Tier 2 — Russellian monism	~40%
Tier 2 — Physicalism	~26%
Tier 2 — Idealism	~23%
Tier 2 — Dualism (aggregate)	~10%
Observed physical world is not the bottom level	~50%
Local world generated by/dependent on intelligence or mind	~45%
Simulation	~42%

Question / hypothesis	Central estimate
Meaningful persistence beyond bodily death	~47% (separate postmortem-survival model)

I. Russellian monism and the causal structure

Russellian monism does not ordinarily deny neuroscience. It can accept lesions, anesthesia, stimulation, neuronal processing, evolution, and computational explanation exactly as described by physical science. Its distinctive claim concerns what the entities participating in those causal structures are intrinsically.

Russellian monism should therefore receive nearly all of physicalism's credit for detailed brain dependence and ordinary physical causal structure. Its distinctive burden lies elsewhere: whether an appeal to intrinsic or proto-phenomenal nature actually explains unified consciousness rather than merely relocating the explanatory gap.

Evidence	Idealism	Physicalism	Russellian	Interactionist/ separable dualism
Detailed brain dependence	.55	1.00	.95	.45
Closure/ evolution/ computational success	.80	1.00	1.00	.40

II. Priors and scientific success

Scientific integration should not be counted twice. The central priors are therefore only mildly tilted toward physicalism and Russellian monism, while the likelihood model separately gives

those theories credit for lawfulness, causal closure, evolutionary continuity, and computational success.

The central prior is deliberately contestable rather than presented as uniquely neutral. Equal-prior and consensus-heavy alternatives remain robustness tests.

Hypothesis	Central prior
Idealism	25%
Physicalism	30%
Russellian monism	30%
Interactionist/separable dualism	10%
Epiphenomenal property dualism	5%

III. Phenomenal consciousness

The hard problem remains the strongest single challenge to physicalism. A complete structural and functional description does not transparently entail why there is something it is like to undergo a conscious state. Yet the force of this evidence depends heavily on whether the explanatory gap is ontological or instead conceptual and ultimately reducible.

Both interactionist and epiphenomenal property dualism score well on phenomenal consciousness considered alone because neither reduces qualia to ordinary physical structure. Their costs arise downstream and differ sharply: interactionist dualism faces causal-closure and interaction problems, whereas epiphenomenalism faces causal impotence, psychophysical harmony, and phenomenal-judgment problems.

I	P	R	D _i / D _e
2.20	1.00	2.20	2.40

The likelihood ratio of roughly 2.2 deliberately leaves substantial probability that physicalism will eventually dissolve or reduce the explanatory gap. If the gap were shown to be genuinely ontological, both idealism and Russellian monism would rise sharply.

III-A. Eliminative materialism and illusionism

Classical eliminative materialism, associated especially with Paul and Patricia Churchland, argues that some categories of folk psychology - such as belief or desire understood in their ordinary theoretical form - may fail to survive mature neuroscience. That thesis should be distinguished from the stronger claim that phenomenal consciousness itself is illusory or misdescribed. Contemporary illusionism, associated prominently with Keith Frankish and predecessors such as Daniel

Dennett, holds that introspection represents experience as possessing irreducible phenomenal properties even though the underlying states are wholly functional or physical.

This is a genuine physicalist escape route from the hard problem. Instead of deriving irreducible qualia from physical structure, the theory denies that the supposed irreducible properties are instantiated as ordinarily conceived and shifts the explanatory burden to the "illusion problem":

why introspection and report make it seem that such properties exist. The position gains from neuroscience, causal closure, and parsimony, but it inherits the challenge of explaining the apparent immediacy of experience without merely relocating phenomenality into the representation of phenomenality.

Eliminative and illusionist theories are therefore treated as subfamilies within physicalism rather than as a fifth top-level ontology. The physicalist likelihood on phenomenal consciousness already leaves substantial room for the possibility that the explanatory gap will be dissolved by conceptual revision, representational theory, or illusionism. No additional independent multiplier is added, avoiding double counting.

III-B. Empirical survival and anomalous-consciousness evidence

A separate companion paper, A Bayesian Analysis of Life After Death, asks whether meaningful conscious continuation beyond biological death occurs. It uses a local-architecture tree rather than the Tier-2 ontology partition used here and examines near-death experiences, prospective OBE targets, prerecord childhood past-life cases, blinded mediumship and psi, purpose-built ciphers and locks, terminal/paradoxical lucidity, and deathbed/shared-death experiences. Its most important result for the present metaphysical analysis is that the direct paranormal literature is not carrying the case: after negative tests, replication discounts, mechanism specificity, and dependence controls, the direct empirical cluster is close to net-neutral.

The best prerecord childhood past-life cases remain the strongest positive direct material, but even genuine anomalous information would not uniquely imply reincarnation; living-person telepathy, clairvoyance, deceased communication, idealist nonlocal information, and documentation error remain competitors. Blinded mediumship provides a smaller anomaly after replication discount. Prospective OBE target failures and purpose-built cipher/lock nulls are genuine negatives for theories that specifically predict precise extracerebral perception or motivated postmortem communication, but they have little force against S1 or idealist architectures that predict no in-world return channel.

Evidence	Anomalous consciousness / information	Personal survival
NDE phenomenology	Moderate	Weak
Prospective OBE targets	Little / somewhat negative	Little / somewhat negative
Best pre-recorded past-life cases	Moderate-strong	Moderate
Terminal/paradoxical lucidity	Moderate evidence that cognition can reappear unexpectedly	Very weak
Ordinary deathbed visions	Mild	Very weak
Shared-death experiences	Mild	Weak
Blinded mediumship	Moderate anomaly if replicated	Weak-moderate
General psi / Ganzfeld	Mild-moderate anomaly	Indirect

In that survival model, the H0/H1/H2 decomposition remains conceptually useful - H0 conventional production physicalism, H1 expanded naturalism with unknown information or cognitive mechanisms, and H2 non-production or survival-compatible consciousness - but a single aggregate Bayes factor can mislead. Positive and negative observations target different auxiliary mechanisms. The dependence-adjusted direct multipliers are approximately production/termination 1.00, S1 1.01, idealism 1.08, and separable/transmission dualism 0.93.

A generic 3-8 Bayes factor against physicalism is therefore not imported from the survival literature. That would compress distinct mechanisms too aggressively. The better conclusion is more restrained: anomalous reports prevent a dismissive "nothing to explain" posture, but current direct survival evidence scarcely moves the ultimate-ontology scoreboard once mechanism specificity and dependence are enforced.

The greater relevance to the present paper comes from structural analysis rather than paranormal reports. Detailed brain dependence strongly favors production models over naive separable dualism, but it discriminates much less sharply against participatory simulation and interface idealism. Anesthesia belongs inside that same dependence cluster rather than constituting a second independent penalty. Sean Carroll's Core Theory is strong against unknown low-energy carriers that must extract person-defining information from a brain at death, but not against models in which the persisting subject was never wholly produced or stored in the brain.

Terminal/paradoxical lucidity remains a useful conceptual warning rather than a survival proof: severe behavioral impairment does not entail destruction of all corresponding information or capacities. Sophisticated physicalism can explain

transient lucidity through network dynamics, so this receives essentially no independent metaphysical multiplier. Its role is to block an overly simple inference from impaired access or expression to annihilation of the underlying information.

The methodological lesson is to ask which architecture actually predicts each observation, count correlated evidence once, and distinguish local causal architecture from ultimate ontology. The survival analysis is therefore used here as a cross-check on model architecture, not as an independent posterior to be imported wholesale.

DMT, synesthesia, and claimed telepathy

The evidential perimeter also includes anomalous-consciousness claims that are culturally controversial or associated with psychedelics. The governing rule is symmetrical: unusual claims receive consideration, but not relaxed standards. Three propositions must be kept distinct: an unusual experience occurs; it contains information unavailable through ordinary channels; and the best explanation is metaphysically nonstandard.

What the established DMT evidence shows

N,N-dimethyltryptamine reliably produces unusually immersive alterations of consciousness. Controlled studies establish radical changes in visual experience, spatial and temporal organization, selfhood, and perceived reality. Survey evidence also shows that apparent encounters with autonomous entities are common among selected DMT users and that many experiencers describe the encountered domain as unusually real. Placebo-controlled comparisons find substantial phenomenological overlap between DMT states and reported near-death experiences. These findings establish the occurrence, structure, and psychological force of the experiences; they do not establish that the apparent entities or environments exist independently.

The strongest naturalistic account begins with pharmacology and shared neuroanatomy. DMT acts principally through serotonergic systems, especially 5-HT_{2A} receptors, and EEG-fMRI work reports increased global connectivity, reduced network segregation, and altered cortical hierarchy. Those changes make intensified imagery, weakened top-down constraint, unusual cross-network association, and disruption of the ordinary self-model unsurprising. They provide a detailed account of why the state occurs without by themselves deciding whether the brain produces consciousness, filters it, or instantiates the intrinsic side of a Russellian structure.

The philosophically strongest DMT observations are the apparent autonomy of encountered agents, the coherence of seemingly navigable environments, and recurring motifs across subjects. Yet unconscious generative systems can surprise the conscious ego, as dream characters do, and cultural transmission supplies familiar motifs. Existing surveys are retrospective and self-selected. No established experiment has shown isolated DMT subjects retrieving the same novel information, encountering the same independently identifiable agent, or transmitting a message through a shared environment.

Synesthesia and recurrent visual structure

Developmental synesthesia is an automatic and durable coupling between sensory or conceptual domains; psychedelic states can produce looser cross-modal binding in which music becomes visible, emotion acquires spatial form, or patterned light becomes linguistically suggestive. This supplies a plausible mechanism for recurring alphabet-like or glyph-like visual structure. The decisive evidential variable is intersubjective specificity. Similar visual cortices exposed to the same structured stimulus can produce broadly similar geometry; independently reproduced unfamiliar characters at the same physical coordinates would be less expected, and recovery of verifiable information unknown to the observers would be more important still.

The laser-code claim

Danny Goler has reported that large numbers of participants viewing a diffracted 650-nanometer laser projection under DMT perceive coherent, often katakana-like symbols. The claim is interesting because it combines an external stimulus with allegedly recurring subjective structure. The published pilot report, however, does not provide the participant-level data, unsuccessful-trial count, blinding, predetermined scoring rule, control group, or quantitative comparison of drawings needed to distinguish shared neurocognitive response from suggestion and semantic priming. Code-like appearance is not evidence of literal computer code, and simulation hypotheses do not independently predict that implementation instructions should

become optically visible under a serotonergic psychedelic. The present simulation likelihood update is therefore effectively negligible.

The Telepathy Tapes and nonspeaking autistic children

Widely circulated demonstrations involving nonspeaking autistic children sometimes appear to show identification of numbers, words, pictures, or private thoughts known to parents or other close partners. If such performance survived genuine sensory isolation and independent response production, it would have greater metaphysical significance than

unusual phenomenology because it would involve transfer of specific information. The principal current weakness is authorship control. Rapid Prompting Method, Spelling to Communicate, handheld letterboards, or target-informed facilitators create well-documented opportunities for subtle ideomotor guidance through board angle, movement, posture, timing, vocal cadence, and reinforcement.

A preregistered 2026 systematic review of Rapid Prompting Method, Spelling to Communicate, and related variants searched thousands of records but found no study meeting its criteria for empirical demonstration of independent authorship. That finding does not imply that nonspeaking autistic people lack rich cognition; it means that the response channel used in many telepathy demonstrations has not been adequately validated. A decisive experiment would isolate target and receiver physically and acoustically, generate targets only after isolation, exclude all target-informed persons from the response room, use an independently controlled response interface, preregister scoring and stopping rules, preserve synchronized recordings, publish every trial, and obtain independent replication.

Bayesian consequence

The DMT literature modestly enlarges the set of consciousness phenomena requiring explanation but currently has little discriminating force among sophisticated ontologies. A sensitivity judgment in the range of roughly 1.1-1.4 for idealism over production physicalism can be defended from recurrent entities, ontological conviction, and alternate-world coherence, but that small effect disappears or reverses if the recurring structure is readily derived from shared neurocognitive architecture and should not be multiplied independently by the near-death cluster. The laser-code claim adds essentially no current update, and the telepathy material receives no positive multiplier until independent authorship and information isolation are demonstrated. These claims change the research agenda more than the posterior.

IV. Detailed brain dependence

Modern neuroscience supplies content-specific lesion effects, dose-dependent anesthetic responses, stimulation-evoked experiences, split-brain effects, personality changes from focal damage, and increasingly detailed mappings between neural organization and cognition. The survival/anomalous-consciousness review does not overturn this evidence. It does, however, show that behavioral unavailability is not always equivalent to destruction of information and therefore slightly weakens the strongest production-only interpretation of brain dependence.

Physicalism predicts such dependence naturally. Idealism can accommodate it through interface or dissociation theories, but accommodation is not the same as advance prediction. Russellian

monism preserves the ordinary causal brain story and pays almost no penalty. Interactionist/separable dualism has substantially greater difficulty. Epiphenomenal property dualism is different: because phenomenal states are generated by physical brain states and exert no downward physical influence, detailed brain dependence is expected and receives a much milder penalty. The four-column component table below therefore refers to interactionist dualism; the epiphenomenal subfamily is handled separately in the expanded calculation.

I	P	R	D _i
.55	1.00	.95	.45

V. Causal closure, evolutionary continuity, and computation

No established mental intervention violates physical causal accounting; cognition appears in graded evolutionary continuity; and computational models increasingly reproduce perception, memory, language, and reasoning functions. These facts strongly challenge interactionist dualism and modestly challenge idealism, while Russellian monism can preserve the entire physical causal structure. Strict epiphenomenal property dualism also preserves physical closure by design, although evolutionary and psychophysical-harmony questions remain.

V.A. Local causal closure does not entail ultimate causal closure

Participatory simulation supplies a concrete counterexample to the stronger inference from local causal closure to ultimate causal closure. A pre-existing conscious subject could operate an avatar through the architecture: external subject -> interface/simulator -> simulated neural state -> simulated behavior. The simulator can translate participant intention into an ordinary in-world neural precursor, after which the sequence unfolds according to local physical laws.

A neuroscientist confined to that world could observe a complete chain N1 -> N2 -> N3 -> muscle movement. Every local physical event could have an appropriate local physical precursor. The causal model could be predictively excellent while the further inference 'therefore there is no extra-simulation subject or deeper mental causation' remained invalid. Local physical causal closure can therefore coexist, in principle, with deeper causation implemented through the rules generating the local sequence.

This is a coherence argument, not positive evidence that our universe has this architecture. Straightforward physicalism retains a simplicity advantage: the apparently self-contained physical causal order simply is the complete order. But empirical local closure does not conceptually exclude every ontology in which consciousness is prior to, or external to, the observed physical level.

Princess Elisabeth and the interaction problem

Princess Elisabeth's objection remains powerful against traditional Cartesian substance dualism: an extensionless mind seems to lack an intelligible mechanism for exerting force on extended matter. Participatory simulation shows why the objection should not automatically be generalized to every consciousness-prior ontology.

If the external participant is physically instantiated at base level, the base-level chain can be entirely physical: biological brain → brain-computer interface → simulator → avatar. There is

then no cross-substance interaction problem at base level, even though consciousness is not produced by the simulated brain. Local mind-priority or apparent local dualism therefore does not imply ultimate dualism.

If the external subject is itself a Cartesian nonphysical substance, however, the problem merely moves upward: one must still explain nonphysical consciousness → interface/simulator. Princess Elisabeth has followed the theory upstairs. Idealist/interface approaches such as the broad architectures associated with Hoffman or Kastrup instead deny that fundamental mentality and physical appearance are two independently existing substances requiring a mechanical push across an ontological gap.

An explicit receiver architecture

The same model improves on the traditional radio/receiver analogy. That analogy merely asserts that receiver damage can impair expression without destroying the source. S1 supplies a conceivable rule-governed architecture: subject ↔ interface algorithm ↔ avatar brain. Simulated hippocampal damage can block memory access; simulated anesthesia can interrupt ordinary experience; split-brain manipulation can constrain integration; destruction of the avatar brain can terminate operation of that avatar - all while leaving the external subject intact.

Brain dependence therefore remains powerful evidence favoring physicalism, but S1 shows that the likelihood of detailed brain dependence need not be extremely low under every survival-compatible model. Near-perfect brain-mind dependence and apparent local causal closure can coexist with persistence of the ultimate experiencer in a fully specified architecture.

I	P	R	D _i
.80	1.00	1.00	.40

V.B. Epiphenomenalism: causal closure without physicalism

Epiphenomenal property dualism provides an important counterexample to any inference from physical causal closure directly to physicalism. On the strict view, physical brain states cause phenomenal states, but phenomenal properties do not cause physical behavior. It therefore predicts ordinary neuroscientific closure almost as well as physicalism and much better than Cartesian interactionism. Its principal costs are different: conscious intentions appear causally idle; reports about experience seem unable to be caused by the phenomenal properties they describe; psychophysical harmony is difficult to explain by selection if qualia are causally inert; and the theory still requires a one-way bridge from physical organization to irreducible experience.

Participatory simulation supplies a partial reply to the first two objections. If consciousness is the way an external participant experiences the simulated world, the participant can have genuine intentions and phenomenal judgments while the simulator maps those states into ordinary in-world neural precursors. An internal neuroscientist can then observe a complete local physical sequence even though the subject participates causally at the deeper generative level. Reports about pain or color can track the participant's experience because the interface coordinates experience, intention, and the simulated neural states that produce speech.

That reply does not vindicate strict global epiphenomenalism. If the external participant is physically instantiated, the ultimate theory may simply be physicalist at base level; if a nonphysical conscious intention itself influences the interface, then consciousness is causally efficacious at the deeper level and the architecture becomes interactionist there. The simulation model therefore shows that a locally epiphenomenal appearance can coexist with genuine subject-level mental causation and phenomenal judgment, while preserving the distinction between local causal closure and ultimate metaphysics. It strengthens S1 as a coherence model rather than raising the probability of strict epiphenomenalism itself.

VI. Quantum physics and nonfundamental spacetime

Quantum mechanics and possible spacetime emergence undermine naïve nineteenth-century material imagery but do not establish that consciousness creates reality. Their direct discriminating force among sophisticated contemporary ontologies is therefore small.

I	P	R	D _i / D _e
1.15	1.00	1.15	1.00

Russellian monism fits this evidence comfortably because abstract structural descriptions naturally raise the Russellian question: what, if anything, has this structure intrinsically? Physics itself does not answer that philosophical question.

VII. Psychophysical harmony

Psychophysical harmony is an important datum: pain generally feels bad and motivates avoidance; rewarding experiences tend to accompany adaptive states; and perceptual qualities track behaviorally useful distinctions.

Straightforward physicalism expects a deep connection between phenomenal and functional states. Russellian monism can also do well if intrinsic phenomenal character systematically tracks causal structure. Naturalistic idealism and interactionist dualism require additional explanation unless purposive coordination is introduced. Strict epiphenomenalism faces a particularly sharp version of the problem: natural selection can select the physical circuitry that causes avoidance behavior, but cannot select an inert quale because of the behavioral effects of that quale. A common physical cause can correlate the quale with the behavior, but the systematic valence-function mapping still requires psychophysical laws or deeper explanation. Because this overlaps with evolutionary and functional evidence, it receives only modest independent weight.

VIII. Fine-tuning and creator intelligence

Fine-tuning requires careful anthropic conditionalization. If the multiverse hypothesis is explicitly defined as an adequately varied ensemble with observer selection, then the probability of observers finding themselves in a life-permitting region should be high. Uncertainty about whether such an ensemble exists belongs primarily in its prior rather than being charged again in its conditional likelihood.

The stand-alone estimate for a generic purposive creator is therefore best expressed as roughly 35-40%, with a very wide 20-50% range. Technological and discoverability fine-tuning create mild upward pressure because observer selection explains life-permission more directly than it explains why observers should inhabit a world that is unusually decipherable and technologically exploitable. Fine-tuning increases the attractiveness of structured explanations - creator, adequate ensemble, simulation, or other generative order - relative to brute single-universe chance, but does not decisively adjudicate among them.

The creator posterior is not fed back into the ontology model, because doing so would double-count the evidence used to generate it.

Fine-tuning beyond habitability: intelligibility, discoverability, and technology

The familiar anthropic response is strongest for conditions necessary for observers to exist at all. It is less obviously sufficient for a second class of features: a universe whose laws are mathematically tractable, whose regularities can be experimentally uncovered, whose matter supports stable machinery and computation, whose electromagnetism permits high-bandwidth information transfer, and whose energy structure can be manipulated across scales. These properties permit observers not merely to exist but to understand and technologically exploit the world that contains them.

This distinction matters Bayesianly. Conditional on the existence of observers, bare habitability is partly screened off by observer selection. Technological and discoverability affordances are not obviously screened off to the same degree. If conscious observers could have existed in many worlds that were opaque, experimentally inaccessible, or technologically sterile, then finding ourselves in the narrower class of highly intelligible and exploitable worlds is additional evidence that calls for explanation. The strength of that evidence depends on an unresolved reference-class question: how broad is the space of observer-permitting but technologically poor universes?

Jeff Shainline's proposed extension of cosmological natural selection supplies one possible naturalistic explanation: universes reproduce through black holes, heritable variation occurs in constants, and technological civilizations eventually enhance a universe's reproductive success by manufacturing black holes. The proposal is elegant because it turns intelligence and technology into part of a cosmological reproductive process. At present, however, it requires several speculative bridges at once—daughter universes from black holes, heritable variation, the technological manufacture of

reproductively relevant black holes, and a material fitness effect. It is therefore retained as a candidate structured-naturalistic explanation rather than scored as an established likelihood advantage.

The metaphysical consequence is deliberately modest. Technological/discoverability fine-tuning strengthens the case that brute-fact physicalism may be explanatorily incomplete, but it does not uniquely favor idealism. It can support purposive creator hypotheses, simulation or intelligence-generated local reality, sufficiently structured multiverse models, or future cosmological-selection theories. To avoid double counting, the effect is folded into the existing fine-tuning/psychophysical-order factor rather than multiplied as a new independent datum. The Tier-2 headline posterior is therefore left unchanged at the present level of precision, while local intelligence-generated reality receives mild qualitative upward pressure but no separate multiplier.

IX. Field theories and related proposals

Field theories are potentially relevant because consciousness might depend on distributed physical dynamics rather than narrowly neuron-local mechanisms. Different versions, however, have very different metaphysical implications.

Ordinary physical field theories

If consciousness turns out to depend critically on the brain's electromagnetic field or another recognized physical field, this could strengthen physicalism or Russellian monism rather than idealism. Consciousness would be less neuron-localized than assumed but would not thereby be nonphysical.

Morphic resonance and related proposals

Sheldrake-type morphic resonance is more disruptive. If reproducibly demonstrated—especially if organizing information were transmitted without an ordinary physical causal channel—it would materially challenge conventional physical causal closure. At present the evidence is not strong enough to justify a meaningful likelihood multiplier, so morphic resonance remains a high-information scientific watch item rather than banked evidence.

X. Other scientific observations with high information value

- Prospective near-death and cardiac-arrest studies: independently verified information acquisition during genuinely unavailable sensory/neural channels would materially hurt ordinary physicalism.
- Terminal or paradoxical lucidity: potentially important if transformed from anecdotal reports into reproducible, quantitatively documented cognition exceeding plausible surviving neural capacity.
- Anomalous cognition or psi: a stable, adversarially replicated information-transfer effect would be highly consequential; current evidence is not scored.
- Artificial consciousness: convincing evidence of phenomenal consciousness arising from a well-understood artificial computational architecture would strongly help physicalist/functionalist theories.
- Anesthesia and covert consciousness after severe brain injury: increasingly precise work can test the tightness and character of the neural-consciousness relation.
- Electromagnetic-field theories: successful derivation of unified conscious contents from field dynamics could clarify the binding problem while remaining physicalist.

XI. Combination, decombination, and the three bridges

Each leading family faces a characteristic bridge problem.

- Physicalism: nonphenomenal physical structure → phenomenal experience.
- Russellian monism: micro/proto/intrinsic nature → one unified subject.
- Idealism: fundamental or universal consciousness → many bounded, mutually opaque subjects.
- Epiphenomenal property dualism: physical causal structure → irreducible but causally inert phenomenal properties, with a psychophysical-law and harmony burden.

Russellian monism's combination problem is potentially its version of the hard problem. Saying that intrinsic natures are proto-phenomenal does not explain why many such natures constitute one experience. The recurring 'what exactly is X?' objection therefore survives: Russellian monism may identify where the unknown intrinsic nature belongs without fully telling us what it is.

Two quantum proposals illuminate the issue without resolving it. David Pearce's non-materialist physicalism identifies experience or proto-experience with the intrinsic nature of physical reality. He conjectures that quantum coherence across neural feature processing supplies the physical unity of a single phenomenal scene. The challenge is to show why one

coherent physical state is one subject and to demonstrate the relevant neural coherence. His proposal invokes no new collapse law.

Roger Penrose and Stuart Hameroff instead propose orchestrated objective reduction: quantum processes in neuronal microtubules terminate through a gravity-related collapse, and those events are proposed to constitute conscious moments. This is a more specified biological and mathematical conjecture, but it depends on relevant microtubule coherence, a disputed objective-collapse process, and an unexplained link from collapse to experience. The two theories differ on the candidate conscious event: Pearce emphasizes the coherent state; Penrose and Hameroff emphasize its objective reduction.

Pearce fits most naturally within the Russellian-monist family, though his preferred label is non-materialist physicalism. Orch OR can be read as a broad physicalist theory with added psychophysical commitments and has no unique allocation among the four ontology families. Neither conjecture has received decisive empirical confirmation. Their existence does not itself justify shifting the central 40/26/23/10 partition; a successful independent test of a consciousness-relevant quantum mechanism would warrant a new likelihood assessment.

I	P	R	D _i
.80	1.00	.70	.90

XII. Tier-2 calculation

Evidence	I	P	R	D _i	D _e
Phenomenal consciousness	2.20	1.00	2.20	2.40	2.40
Brain dependence	.55	1.00	.95	.45	.90
Closure/evolution/computation	.80	1.00	1.00	.40	.90
Quantum/fundamental physics	1.15	1.00	1.15	1.00	1.00
Fine-tuning / psychophysical / technological order	1.20	1.00	1.05	1.15	.75
Combination/decombination	.80	1.00	.70	.90	.90

Using central priors I 25%, P 30%, R 30%, interactionist/separable dualism 10%, and epiphenomenal property dualism 5%, the raw composite likelihood factors relative to physicalism are approximately I 1.07, P 1.00, R 1.77, D_i 0.45, and D_e 1.31. The D_e factor is intentionally tempered: epiphenomenalism receives strong credit for closure and brain dependence but pays for psychophysical harmony, one-way emergence, causal impotence, and phenomenal judgment. Because the evidence clusters are correlated and the likelihood judgments uncertain, the headline applies a 75% log-evidence attenuation.

Model	R	P	I	D
No attenuation	~44%	~25%	~22%	~9%
75% attenuation — headline	~40%	~26%	~23%	~10%

The meaningful result is not the second decimal place. Russellian monism remains the leading ultimate ontology once it receives nearly all of physicalism's credit for brain dependence and causal structure while retaining much of the anti-reductive benefit of phenomenal consciousness. Explicitly separating epiphenomenal from interactionist dualism raises the aggregate dualist posterior from the high single digits to roughly ten percent without overturning the overall ordering.

The separate postmortem-survival analysis leaves the direct empirical anomaly cluster close to net-neutral after branch-specific controls. Its more important implications are structural: S1 and interface idealism can reproduce detailed local brain dependence; naive separable dualism remains heavily penalized; epiphenomenal property dualism preserves closure but does not by itself imply survival; anesthesia is not independent evidence once brain dependence is counted; and carrier/extraction objections apply only to mechanisms that require such extraction. These considerations are incorporated into the partition rather than multiplied again as a fresh paranormal factor.

XIII. The Russellian/physicalist boundary

The roughly 40/26 distinction between Russellian monism and physicalism is partly taxonomic. Some philosophers classify Russellian monism as a species of physicalism; others regard experiential intrinsic natures as panpsychist or neutral-monist. A safer substantive conclusion is that the leading family preserves ordinary physical science structurally and causally while regarding it as potentially incomplete as a description of intrinsic reality.

XIV. Tier 1 — the local question

The analysis distinguishes ultimate ontology from the question more directly relevant to human beings: is the physical universe we observe actually the bottom level, are persons exhausted by their biological bodies, and does apparent physical causal closure exhaust our ontological situation?

Tier 1 therefore treats local nonfundamentality, simulation, and mind- or intelligence-generated reality as distinct variables from Tier 2. A world can be locally generated or simulated while base reality remains physical, Russellian, idealist, or dualist.

XV. Simulation probability: philosophy, physics, and temporal selection

The present model assigns a central simulation probability of about 42%, with a serious uncertainty range of roughly 20-60%. Because that estimate is substantially higher than conventional academic priors and will be surprising to many readers, the simulation branch warrants disproportionate scrutiny. This section therefore functions as a focused literature review. It separates five questions that are often conflated: whether Bostrom's self-locating argument is structurally sound; whether consciousness can exist in a computation; whether modern physics is more expected under simulation than under a fundamental physical world; whether proposed experiments have found or excluded simulation signatures; and whether our temporal location near a possible first ASI transition creates an additional selection effect.

A speculative participatory architecture connects these questions. A base-level quantum substrate might support a generated world whose interacting components decohere into definite-looking histories, while an external subject experiences one such history through an in-world brain interface. Pearce's bound state could organize the interface contents; the external subject would remain the experiencer. This is a proposed architecture, not a prediction of quantum mechanics. Physical observation does not require a conscious observer, and superposition does not automatically save computation on classical hardware. If actual quantum unity or objective reduction is necessary for consciousness, a software simulation that only describes those processes might fail to generate subjects. That would change the relative plausibility of generated-consciousness and participatory or quantum-substrate simulation, without independently raising the overall simulation probability. The current Tier-1 estimates remain unchanged pending evidence and conditional recalibration.

XV-A. Bostrom's trilemma: what it establishes and what it does not

Bostrom (2003) did not simply argue, "future computers will run many simulations, therefore we are probably simulated." His formal result is disjunctive. Given assumptions about the number of posthuman civilizations, the computational feasibility of ancestor simulations, and the number of conscious simulated observers, at least one of three propositions must be true: civilizations like ours almost always fail to reach technological maturity; technologically mature civilizations almost never run large numbers of ancestor simulations; or observers with experiences like ours are overwhelmingly simulated. Bostrom later emphasized that the argument does not identify which disjunct is true and that he himself did not infer that we probably live in a simulation (Bostrom 2009).

The formal history matters. Bostrom and Kulczycki (2011) identified a mathematical non sequitur in the original formulation and supplied two repairs that preserve the trilemma. The existence of a patch is not a refutation, but it is a warning against treating the argument as a one-line counting proof. Brueckner (2008) and Bostrom (2009) also disputed how the self-locating step should be interpreted. The strongest version of the argument therefore rests not merely on the existence of many possible simulations, but on a rule for moving from population ratios or expected population ratios to first-person credence.

XV-B. Philosophical attacks, defenses, and newer Bayesian formulations

The deepest philosophical dispute concerns self-location. Weatherson (2003) argued that the indifference principle needed to infer "I am probably one of the numerous simulants" has no interpretation that straightforwardly supports Bostrom's conclusion without additional assumptions. Crawford (2013) pressed a related worry by constructing an analogous argument for being a freak or Boltzmann-type observer: the mere existence or numerical predominance of a class of observers does not automatically show that one is a member of that class. These objections target the move from third-person census facts to first-person probability, not the technological possibility of simulation itself.

Birch (2013) develops a different objection, which is especially relevant to a paper that also uses evidence from physics. Bostrom's argument seems to trust our physics enough to estimate the future limits of computation while becoming skeptical about whether the physical world described by that same evidence is the ground level of reality. Birch calls this selective skepticism. If simulated physics may differ arbitrarily from base physics, our apparent computational limits are a poor guide to the simulators' resources; but if our physics is reliable enough to extrapolate base-world computation, it also appears to

supply evidence that we are ordinary physical beings. The objection does not make simulation incoherent. It limits how much probability one should extract from computational extrapolation alone.

There are important defenses. Lewis (2013), comparing the Simulation Argument with the Doomsday Argument, argues that the relevant structural disanalogies allow the former to succeed even if the latter fails. Bostrom's own replies emphasize that the trilemma is a conditional constraint rather than a direct assertion of the simulation hypothesis. More recently, Thomas (2026) has proposed a significant reconstruction. Instead of requiring confidence that the actual ratio of simulants to non-simulants is high, Thomas uses a "high expectation" condition: conditional on our being non-simulants, the expected ratio of simulants to non-simulants in an appropriate reference class may be very high even when there is substantial probability that no such simulations are ever run. A small probability of an astronomically large number of simulations can dominate the expectation.

Thomas's argument is unusually relevant here because his illustrative reference class is people living on minor variants of twenty-first-century Earth. That formulation is much closer to the question posed by this paper than an unrestricted class of all observers. It also sharpens rather than eliminates the reference-class problem: the more specifically one defines the class so that our detailed evidence is admissible, the fewer simulated observers may qualify; the more broadly one defines it, the harder it is to show that the rest of our evidence is irrelevant to whether we are simulated. Thomas therefore strengthens the pro-simulation armchair case but does not convert it into a measured probability.

Kipping (2020) approaches the problem with Bayesian model averaging rather than accepting simulation feasibility as given. On his model, uncertainty over whether Bostrom-like simulations are technically possible prevents the probability from automatically racing toward unity; his central analysis stays below 50% and tends toward 50% in the limit of arbitrarily many simulations under the modeled alternatives. He also notes that actually observing our civilization produce Bostrom-like simulations would strongly change the evidence. This is a useful corrective to popular "one in billions" rhetoric: enormous simulated populations matter only after probability is assigned to the physical and consciousness assumptions that make those populations possible.

The consciousness assumption is not a side issue. Summers and Arvan (2022) argue that panpsychist or panqualityist theories can undermine ordinary software-simulation arguments because functional replication may fail to generate the relevant consciousness. Harris (2024) attacks a different auxiliary premise, arguing that mature civilizations may be unwilling to run large numbers of ancestor simulations. These arguments fit the architecture already used in this paper. A reduction in the probability of generated consciousness (S2) does not eliminate simulation if participatory/avatar simulation (S1) remains technologically possible; but it does remove the simplest inference from "many computations" to "many conscious simulated observer-moments." Simulator motivation likewise must be modeled rather than assumed.

Literature verdict. The professional philosophical literature does not support the claim that Bostrom proved we are probably simulated. It does support a narrower conclusion: once non-negligible probability is assigned to civilizations capable and willing to create vast numbers of conscious simulations, self-locating evidence can produce a large update, but the size of that update depends on contested principles of indifference, reference-class choice, simulator motivation, and the theory of consciousness. Thomas's recent expectation-based argument is a genuine pro-simulation development; Birch, Weatherson, Crawford, Summers-Arvan, and Harris preserve substantial counterweight.

XV-C. The 1890 counterfactual: what modern physics would have looked like in advance

Consider a counterfactual framing: suppose a serious simulation or generated-reality hypothesis had been on the table in 1890, before relativity and quantum theory. How would the physics discovered from 1900 through the present have updated that hypothesis? The point is evidential rather than rhetorical. Evidence should be scored by asking how expected it would have been under competing hypotheses before the evidence was known.

The 1890 comparison must be stated fairly. Late-nineteenth-century physics was not a crude theory of billiard-ball matter; it already contained Maxwellian fields, thermodynamics, statistical mechanics, and serious puzzles. Even so, the dominant successful picture treated spacetime and continuous dynamical law as an objective stage on which matter and fields evolved. A generic simulation or generated-world hypothesis did not uniquely predict quantum mechanics, but it made several later discoveries less conceptually surprising: discreteness, observer-relative or measurement-dependent descriptions, strict information limits, and the possibility that spacetime itself is not fundamental.

Quantum mechanics provides the most obvious pressure. Quantization, superposition, entanglement, and the measurement problem are difficult to assimilate to naive classical realism. Bell-inequality violations make the point sharper. Loophole-free experiments such as Hensen et al. (2015), together with the body of work recognized by the 2022 Nobel Prize in Physics, rule out broad classes of local-realist hidden-variable explanations under the relevant assumptions. They do not permit controllable faster-than-light signaling and they do not imply that a programmer is choosing outcomes. Standard quantum theory predicts the data perfectly well without simulation. The correct inference is therefore not "Bell proves

code," but "the world is less like a self-contained local classical mechanism than an 1890 observer would naturally have expected."

Relativity cuts both ways. The relativity of simultaneity and the fusion of space and time would have been revolutionary from an 1890 perspective, but exact Lorentz symmetry is not an obvious hallmark of a simple computer grid. A naive fixed lattice commonly introduces a preferred frame or rotational artifacts. The extraordinary success of Lorentz invariance is therefore a counterweight against crude cellular-grid pictures even while the broader lesson - that ordinary spatial and temporal intuitions are not fundamental - favors non-naive generative models.

The information-theoretic turn in fundamental physics is more suggestive but still non-specific. Lloyd (2002) showed that physical law places finite bounds on the information and elementary operations physically available to the observable universe. Bousso's review of the holographic principle (2002) describes strong evidence connecting geometric area with limits on information content. In modern holography, Takayanagi (2025) emphasizes that gravitational spacetime can emerge from entanglement structure in an underlying quantum system. None of this says that an external intelligence is running software. It does say that geometry, matter, and information are related far more deeply than a nineteenth-century ontology would have suggested, and it raises the prior plausibility of local nonfundamentality.

This distinction is central to the Bayesian score. Let H_{1890} denote a relatively straightforward fundamental classical world of the type a scientifically informed observer in 1890 might have regarded as natural, and let G denote a broad generated/information-structured world that includes but is not limited to simulation. The twentieth- and twenty-first-century physics package is plausibly more likely under G than under H_{1890} . A modest likelihood ratio on the order of 1.2 is defensible as a sensitivity judgment. But the relevant comparison for this paper is not G versus H_{1890} . It is simulation versus sophisticated non-simulation physics after quantum theory, relativity, QFT, and modern quantum gravity programs are admitted. On that comparison the likelihood ratio is much closer to one.

Accordingly, the analysis assigns modern physics two different effects. First, it remains a genuine modest argument for the broader proposition that our observed physical level may not be metaphysically fundamental. Second, it supplies only a small simulation-specific increment - roughly 1.05 to 1.10 at most in the present model after dependence is recognized. Even that small increment should not be multiplied again into the headline calculation because Section VI and the local-nonfundamentality estimate already incorporate quantum and emergent-spacetime considerations. The 1890 counterfactual strengthens the interpretation of that evidence more than it independently raises the 42% simulation number.

XV-D. Direct tests, computational constraints, and the meaning of null results

A simulation hypothesis that can accommodate every possible observation is empirically empty. Several researchers have therefore proposed tests of restricted simulation architectures. Beane, Davoudi, and Savage (2014) analyze a universe implemented on a cubic spacetime lattice and derive observable consequences such as a high-energy cutoff and possible rotational-symmetry breaking in ultra-high-energy cosmic rays. Existing observations place a very high lower bound on the corresponding lattice scale. This is valuable physics, but its evidential scope is narrow: failure to find cubic-lattice artifacts counts against that implementation, not against simulations with Lorentz-preserving, adaptive, non-lattice, phenomenological, or otherwise different architectures.

Campbell, Owahdi, Sauvageau, and Watkinson (2017) propose wave-particle and delayed-choice-type experiments based on a much more specific hypothesis: a finite-resource simulator economizes by rendering definite content only when information becomes available to a conscious observer. The experiments are interesting precisely because they make the model risk falsification. But that is also why a null result should not be generalized. It would disfavor consciousness-triggered resource-saving renderers, not Bostrom-style simulations as a class. This literature review located no peer-reviewed positive result establishing the predicted simulation signature.

Vazza (2025) attacks simulation from the opposite direction by estimating the information and energy required to compute the universe, Earth, or even a much lower-resolution Earth under physical laws like ours. His conclusion is strongly negative for a brute-force same-physics simulator. Edge and Brown (2026) reply that this targets a much stronger hypothesis than Bostrom requires: an ancestor simulation need not calculate every remote microphysical degree of freedom if it need only generate coherent subjective experience without detectable anomalies. The exchange is useful because it identifies a real fork. Full microphysical emulation may be physically prohibitive under laws like ours; selective or phenomenological simulation requires much less but adds architectural assumptions.

Wolpert (2025) gives the issue a more rigorous computer-science formulation using the physical Church-Turing thesis, recursion results, and computability limits. The resulting picture is mixed: simulation and even forms of self-simulation can be formally possible, while general theorems also limit what agents can decide about arbitrary computations and therefore about some putative simulations. The main lesson for Bayesian purposes is that "can be computed" and "can be detected

from inside" are different questions. A sufficiently faithful simulation may be observationally indistinguishable from a base world for internal observers.

The weighting rule should be hypothesis-specific. A null experiment receives substantial negative weight only against a hypothesis that specifically predicted a positive result. The same logic applies in prospective OBE research: failure of concealed visual targets bears on mechanisms that predict accurate extracorporeal perception but has little force against survival models that predict no such return channel. Likewise, absence of lattice anisotropy or consciousness-rendering anomalies can count strongly against those narrow simulation submodels while barely moving an unrestricted simulation hypothesis. The current simulation-test program is therefore not analogous to decades of repeatedly failed tests of one crisp generic prediction. It is a collection of young, architecture-specific probes.

This does not make the absence of evidence irrelevant. If years of increasingly varied and well-powered tests were to exclude every computationally natural implementation while base-level physics continued to explain the data compactly, the simulation posterior should fall. At present, however, the evidential debit is small. The strongest current negative comes from complexity and motivation costs attached to particular simulator designs, not from a clean experimental observation that generic simulation strongly predicted and base reality did not.

XV-E. Temporal focality and repeated singularity simulations

Roman Yampolskiy's recent public discussion adds a point that is easy to understate. His 2023 paper "How to Escape From the Simulation" does not itself argue that simulation is probable; its abstract explicitly brackets the evidential question and asks whether intelligent agents could jailbreak a simulated environment. The relevant probability argument appears instead in his broader work and recent interview discussion, and it has a published precursor in Turchin, Batin, Denkenberger, and Yampolskiy (2019). That preprint classifies "singularity simulations" - simulations centered on the origin of the first AGI - as a potentially dominant simulation type and specifically identifies simulations of the twenty-first century and global-catastrophic-risk period as plausible.

The stronger version is not merely "what are the odds of being alive near ASI?" It introduces differential rerunning. A simulator interested in the transition from biological civilization to superintelligence need not run one complete history from abiogenesis through the remote future at equal temporal resolution. It could branch, restart, and replay the transition era many times: different AI architectures, governance choices, safety interventions, wars, laboratory decisions, or alignment strategies could be tested from nearby initial conditions. To expose the selection mechanism, consider a deliberately extreme illustration in which a 2020-2030-like interval is instantiated a billion times for each ordinary historical interval. The number is not asserted as realistic; it makes the mechanism transparent.

The mechanism can be stated formally. Let T be the proposition that an observer-moment lies in a narrow first-ASI transition window, and let m represent the expected simulation multiplicity of that window relative to ordinary history. If a fraction f of relevant base observer-moments lie in T , then a simple targeted-replay model gives $P(T|S)$ approximately $mf/[1+(m-1)f]$. When f is small and m is not so large as to saturate the probability, the numerator is amplified approximately by m . Very large m can make transition-era observer-moments dominate the simulated reference class even if a full unbranched history would devote almost no time to the transition.

This connects naturally to Thomas's high-expectation argument. Thomas notes that even a relatively small probability of an enormous simulated population can drive the expected simulant/non-simulant ratio high for an appropriate reference class. A simulator who repeatedly branches minor variants of twenty-first-century Earth around an AGI/ASI transition supplies exactly such a population-generating process. The temporal argument is therefore stronger than a bare anthropic astonishment at living in an interesting century: it supplies a causal reason why that century could be systematically overrepresented under S .

The counterarguments are substantial. First, simulator motives are unknown; civilizations might care more about art, ordinary lives, cosmology, or post-singularity history than about the transition. Second, advanced simulators might infer outcomes analytically rather than rerun conscious worlds. Third, if simulated computation does not generate phenomenal consciousness, a vast number of S_2 runs may add no conscious observer-moments at all; the relevant multiplier then depends on S_1 or another consciousness-preserving architecture. Fourth, reference classes remain contestable. Fifth, the argument partially overlaps with Bostrom-style ancestor-simulation motivation and should not be double counted. Finally, the premise that we are actually within decades of a first ASI transition is itself uncertain, even if this paper treats that premise as highly likely for the purpose of the update.

A useful way to see the evidential asymmetry is to compare $P(T|S)$ with $P(T|B)$, not to compare T with the entire age of the universe. Under base reality, the denominator should be an appropriate class of conscious observer-moments, not dinosaur years or empty cosmic time. That correction prevents absurdly huge Bayes factors. But if simulated worlds

selectively proliferate transition-era observer-moments, $P(T|S)$ can still be materially larger than $P(T|B)$. This is the core of the temporal-focality and repeated-replay argument after the rhetoric is stripped away.

XV-F. Integrated Bayesian update

The integrated update should not multiply every pro-simulation item in this section as though it were new and independent. The simulation submodel uses a 30% baseline that already incorporates Bostrom-style counting, the possibility of participatory simulation, quantum/nonfundamental-physics considerations, and large uncertainty about motives and consciousness. The main incremental quantitative datum is the transition-era selection effect; the physics review mainly clarifies an existing factor, while the philosophical critiques widen uncertainty and the direct experimental literature modestly penalizes only narrow implementations.

Consideration	Direction for simulation	Present treatment	Main limitation / dependence
Bostrom / Thomas self-location	Up	Supports a substantial prior	Reference class and calibration remain contested
Consciousness / S1-S2 uncertainty	Mixed	Condition simulation architecture on mind theory	Many computations need not equal many conscious observers
1890-to-modern physics package	Small up	Standalone LR about 1.2 vs naive classical fundamentality; little new incremental weight	Modern non-simulation physics predicts the same observations
Direct tests and physical constraints	Small down for narrow models	No generic multiplier; architecture-specific penalties	Most tests target lattice or rendering submodels, not simulation as a whole
Simulator motivation	Down / widens	Do not assume ancestor-simulation enthusiasm	Preferences of mature civilizations are unknown
ASI-transition focality / repeated replay	Up	Incremental BF about 1.7	Depends on ASI proximity, motives, reference class, and conscious simulation

For the central calculation, an incremental Bayes factor of approximately 1.7 is used, with a sensitivity range of about 1.1-3.0. Starting from $P(S)=0.30$, prior odds are $0.30/0.70=0.429$. Multiplying by 1.7 gives posterior odds of about 0.729, corresponding to $P(S|E)=0.729/(1+0.729)=0.421$. The central simulation estimate is therefore approximately 42%. The 1.1-3.0 range moves the 30% starting point to roughly 32-56% before uncertainty in the starting estimate is added. A serious overall range of approximately 20-60% is more honest than a narrow confidence band.

The professional literature does not converge on 42%. The figure is a model-conditioned credence produced by placing pro- and anti-simulation considerations in the same Bayesian ledger. The literature does not justify moving toward near-certainty: reference-class objections, simulator motivation, consciousness, Lorentz symmetry, computational cost, and the absence of a generic empirical signature remain substantial. At the same time, expectation-based self-location arguments and transition-era selection make it difficult to dismiss simulation as merely fanciful once non-negligible probability is assigned to advanced simulators.

XVI. Simulation branches

Branch	Description	Conditional status	Illustrative implication
S1 Participatory/avatar	Pre-existing conscious subject interfaces with simulated avatar	May dominate if computation alone is not conscious	Strongly survival-compatible
S2 Generated consciousness	Simulation computation produces a new conscious subject	May dominate if strong functionalism is true	Survival not implied
S3 Sparse/NPC/other	Few genuine subjects or unusual architectures	Residual branch	Architecture-dependent

A fixed S1/S2 split is not the appropriate central representation. Two technological questions must be separated: whether an advanced civilization can simulate an environment, and whether computation within that environment can generate a genuinely conscious subject. The first can be true while the second is false. A civilization could simulate physics, bodies, behavior, language, and intelligence without thereby creating phenomenal consciousness.

Under S1 (participatory/avatar simulation), a pre-existing conscious subject is connected through an interface to a simulated nervous system or avatar. Under S2 (generated-consciousness simulation), the simulation computation itself produces a new conscious subject. These architectures make different commitments about consciousness.

The hard problem therefore cuts in opposite directions for S1 and S2. If strong computational functionalism or substrate independence is true, S2 becomes comparatively natural and may dominate S1. If consciousness requires a special biological or physical substrate, S2 becomes much harder while S1 can remain viable through sufficiently advanced interfaces. If idealism or another consciousness-first ontology is true, S1 may become especially natural: the simulation organizes experience for a pre-existing subject rather than manufacturing consciousness from computation.

Accordingly, the relevant quantities are $P(S1 | S, M)$ and $P(S2 | S, M)$, where M is the underlying theory of consciousness. A single unconditional split conceals precisely the metaphysical uncertainty the simulation argument is supposed to illuminate. Until substrate independence is better established, the S1/S2 allocation should remain highly conditional rather than be treated as a stable posterior.

This also creates a possible selection effect. If advanced civilizations discover that arbitrarily intelligent artificial systems are not phenomenally conscious, then civilizations wishing to populate simulated worlds with genuine subjects may have to use external participants. In that technological regime, participatory simulation is not an extravagant add-on to generated simulation; it may be the only viable architecture for genuinely conscious inhabitants.

For the life-after-death question, the distinction is consequential. Under S1, ordinary avatar death need not imply the death of the underlying subject; it may be analogous to disconnection or logout. Under S2, destruction of the computational process may end the subject unless the simulator preserves, copies, or reinstantiates it. Simulation probability by itself therefore cannot be translated directly into survival probability without conditioning on simulation architecture and consciousness theory.

XVII. Why simulation acts like a creator universe locally

If our world is simulated, then even if base reality is physical: our spacetime is not fundamental; our physical laws are generated constraints; our universe exists because agents outside it caused it to exist; apparent local causal closure does not entail ultimate causal closure; external entities can in principle affect our world; and our cosmological origin is purposive rather than brute.

Even perfect apparent causal closure within a simulated world would not establish that the simulated level is causally ultimate. External participant inputs can be implemented as ordinary-looking internal neural precursors. Simulation therefore supplies a worked countermodel to the inference from local causal completeness to ontological completeness.

Thus, for local ontology, simulation is strongly anti-flat-physicalist and substantially idealist-like even when ultimate physicalism remains true. Generated-consciousness simulation may still be physicalist about minds, but it is not flat physicalism about our observed universe.

This local-versus-ultimate distinction is fundamental to the analysis. Simulation can leave ultimate physicalism intact while still making the physical universe accessible to us a generated, nonfundamental level with a deeper causal architecture.

XVIII. Probability that our physical world is not fundamental

Local nonfundamentality can arise through simulation, idealist manifestation or dissociation, creator-produced worlds, some dualist architectures, and other emergent-level scenarios. These possibilities overlap heavily and cannot simply be added.

Because every simulation in this model is a non-bottom-level world, the broader nonfundamentality estimate must remain coherent with $P(S)=0.42$. One transparent calculation holds the conditional rate of non-simulation nonfundamentality at approximately 0.143. This corresponds to a 10-point residual nonfundamentality component for a 70% non-simulation mass in the baseline model. With $P(S)=0.42$, the calculation gives $P(\text{nonfundamental}) \approx 0.42 + (0.58 \times 0.143) = 0.503$. The central estimate is therefore approximately 50%.

$$P(\text{observed physical world is not the bottom level}) \approx 50\%$$

A reasonable serious uncertainty range is approximately 35–65%. This is a coherence update rather than an independent new piece of evidence: it prevents the broader nonfundamentality estimate from remaining below, or implausibly close to, its now larger simulation component.

XIX. Local mind-priority

A stronger question asks whether our physical world is generated by, dependent upon, or an appearance within intelligence or consciousness. This includes participatory and intentionally generated simulation, creator universes, analytic idealism, and relevant cosmopsychist/dissociation models.

Because simulation itself is assigned about 42%, and an intentionally created simulation is a local intelligence-generated world, simulation sets an approximate floor for this broader category. Allowing a modest residual for non-simulation creator and consciousness-first routes, while heavily discounting overlap, yields a central estimate of approximately 45%.

P(local world generated by/dependent upon intelligence or mind) $\approx$ 45%

A broad uncertainty range of roughly 30–60% remains appropriate because the components are correlated and the simulation estimate itself is uncertain. This is not identical to ultimate idealism. A physically constituted civilization could create our universe while ultimate physicalism remained true. From our position, however, the local causal direction would be intelligence $\rightarrow$ our physical universe rather than our physical universe $\rightarrow$ all intelligence.

XX. Persistence beyond bodily death

Local nonfundamentality does not automatically imply survival. Participatory simulation strongly supports persistence of the underlying subject beyond avatar death, whereas generated-consciousness simulation does not by itself. The conditional architecture therefore makes the survival implication depend on the S1/S2 allocation and on the theory of consciousness. Some idealisms permit dissolution of a dissociative boundary, although persistence of autobiographical identity remains a separate issue. Dualism permits separability but does not guarantee continuation. A generic creator could institute continuation but need not.

P(meaningful conscious continuation beyond biological death | separate survival model) $\approx$ 47%

The separate companion paper A Bayesian Analysis of Life After Death gives a preferred headline near 47%, with a most defensible central region around 40–50% and a much wider serious uncertainty envelope around 20–70%. Its model has not been recalculated using the 42% simulation estimate. Raising total simulation probability creates some upward pressure on survival if participatory simulation retains a substantial conditional share, but the magnitude cannot be inferred without re-estimating S1 versus S2 and their overlap with idealist and dualist branches. The 47% figure is therefore reported here as the companion paper's estimate rather than adjusted mechanically.

XXI. Two-tier scoreboard

Tier 2 — What is reality ultimately?

Hypothesis	Central estimate
Russellian monism	$\sim$ 40%
Physicalism	$\sim$ 26%
Idealism	$\sim$ 23%
Dualism (aggregate)	$\sim$ 10%

Tier 1 — What is our world?

Question	Central estimate
Observed physical world is not ultimate/bottom level	$\sim$ 50%
Local world generated by/dependent upon intelligence or mind	$\sim$ 45%
Simulation	$\sim$ 42%
Participatory/avatar simulation	Not assigned a stable unconditional share; conditional on consciousness theory
Meaningful personal persistence beyond bodily death	$\sim$ 47% (separate survival model; broad 20–70% envelope)

XXII. Relation to a separate Bayesian analysis of postmortem survival

This paper occasionally draws on a separate manuscript, A Bayesian Analysis of Life After Death, because both analyses depend on brain dependence, simulation architecture, idealism, dualism, and theories of consciousness. The survival paper asks whether meaningful conscious continuation beyond biological death occurs; it does not treat survival as a fifth ultimate ontology. Instead it builds a local-architecture tree and then applies branch-specific empirical evidence. Five results from that analysis are relevant here. First, strong computational functionalism is treated as a live but minority hypothesis, with a central credence near 25%. This is not a direct probability of physicalism: Russellian monism and some physicalist theories can be substrate-dependent. It does, however, make generated-consciousness simulation less automatic and reinforces the decision to condition S1/S2 on consciousness theory.

Second, participatory simulation has an explicit selection architecture. If computed systems can be intelligent without phenomenality, the relevant observer multiplier need not be the number of generated minds; it can be the number of immersive lives experienced by pre-existing genuine subjects. Repeated immersive lives can recreate a Bostrom-like

observer-selection pressure without assuming astronomical populations of conscious software. The survival model's central participatory-frequency input near 0.30 in the non-functionalist branch is not imported directly into the metaphysics simulation estimate because the conditioning structures differ.

Third, solipsism is represented as a subject-multiplicity parameter inside S1 and idealist architectures rather than as a fifth top-level ontology. Let N denote the number of genuine finite phenomenal subjects associated with the apparent world. $N=1$ produces a one-player S1 or one-local-subject idealist branch; $N>1$ produces multiplayer S1 or plural idealism. This avoids double counting and clarifies why autonomous AI-like behavior weakens, but does not erase, the traditional other-minds evidence against $N=1$.

Fourth, idealism should be internally partitioned when consequences matter. Dissociative analytic idealism, conscious-agent/interface idealism, and cosmopsychism can agree that consciousness is fundamental while differing sharply about finite subjects, brain dependence, and persistence. The Tier-2 headline can aggregate them, but downstream questions should not assign a single consequence conditional to "idealism" without exposing the subfamily weights.

Fifth, death-relevant classification further confirms that local architecture and ultimate ontology are different random variables. The local tree distinguishes production/termination models, S1, idealist interface families, and separable/transmission dualism. Epiphenomenal property dualism belongs with termination unless some independent continuation mechanism is added. An S1 world, by contrast, can have a physicalist, Russellian, idealist, or dualist base. Mixing these levels in one undifferentiated partition would double count worlds.

The survival paper's roughly 47% central estimate does not imply that idealism, dualism, or Russellian monism should each rise correspondingly. It is generated from a different local-architecture tree whose largest live variable is participatory-simulation frequency. The Tier-2 ontology posterior is therefore kept separate at approximately R 40%, P 26%, I 23%, and aggregate D 10%.

Epiphenomenalism and eliminativism mainly improve the ontology taxonomy rather than directly moving the local simulation or survival estimates. The 42% simulation, 50% local nonfundamentality, and 45% local mind-priority estimates are grounded in the Tier-1 analysis developed here. The separate survival estimate should not be mechanically changed without a joint recalibration of simulation architecture, consciousness theory, idealist subfamilies, and dualist branches.

XXIII. Conclusion

The ultimate-ontology contest remains open. Russellian monism leads the central model at roughly 40%, but that is far from decisive; physicalism remains substantial at roughly 26%, idealism at roughly 23%, and the internally partitioned dualist family at roughly 10%. The physicalism/Russellian boundary is itself partly taxonomic.

The more consequential distinction is between the claim that ultimate physicalism is true and the stronger claim that the physical universe we observe is autonomous, fundamental, and causally exhaustive of our existence. Those propositions are not equivalent.

Simulation makes the distinction vivid. The literature shows why neither dismissal nor near-certainty is justified. Bostrom's argument has survived formal repair but remains sensitive to self-location and reference classes; Thomas supplies an expectation-based refinement; consciousness and simulator motivation constrain the population-counting step; modern physics modestly favors nonfundamentality more than simulation specifically; and proposed experiments so far constrain narrow simulation architectures rather than the generic hypothesis. The largest incremental factor in the present model is temporal focality: if humanity is near a first ASI transition, that era may be preferentially branched and replayed by simulators studying precisely the transition they regard as most consequential.

Headline local result

≈50% that the physical world described by our science is not the bottom level of reality

≈45% that it is locally derivative from intelligence or consciousness

≈42% simulation

These estimates do not establish traditional idealism. For inhabitants of this universe, however, local nonfundamentality may be more consequential than the identity of an ultimate substrate several ontological levels beneath us. Detailed neuroscience remains the strongest evidence for profound brain dependence, while epiphenomenalism shows that closure alone does not decide ontology, eliminativism shows that physicalism has a coherent strategy for dissolving the phenomenal explanandum, and participatory simulation shows that apparent local closure can coexist with deeper subject-level causation.

The paper therefore ends with a deliberately plural conclusion: current evidence supports no single ultimate metaphysical theory decisively, but it sharply constrains what a successful theory must explain - phenomenal consciousness or its apparent reality, psychophysical harmony, detailed brain dependence, causal closure, subject unity, and the possibility that local physical structure is generated by a deeper causal architecture. At the Tier-1 level, simulation now deserves unusually serious attention. A 42% central estimate is intentionally startling but not presented as professional consensus; it is a transparent Bayesian judgment whose strongest remaining uncertainties are reference class, simulator motivation, the theory of consciousness, and whether our apparent proximity to a first superintelligence transition is real.

Appendix A. Contributions, attribution, and Centaur-method note

Detailed attribution is included because the paper is also part of a broader Centaur research program examining how substantive human contributions and frontier-model analysis interact across repeated adversarial revisions. The attributions below identify contributions within this collaboration; except where a published precursor is expressly discussed, they are not claims of historical priority. Rodney Malpert originated the project's central methodological demand that competing metaphysical positions be subjected to explicit Bayesian comparison rather than assigned credibility by disciplinary prevalence. Within the collaboration he developed or pressed the local-versus-ultimate distinction; participatory/direct-experiential simulation; the simulation/Descartes architecture in which an external intention is rendered as an ordinary in-world neural precursor; the hydraulic-computer challenge to strong substrate independence; the threshold-death and anesthesia-toggle replies; field-theory possibilities; the participatory observer-life reframing; AI-operated simulation as an experiential-data mechanism; the subject-multiplicity treatment of solipsism; and the recurring "what exactly is X?" challenge to Russellian monism.

Malpert also required explicit evaluation of two philosophy-of-mind boundary cases that had been underdeveloped in earlier analysis: epiphenomenal property dualism and eliminative materialism/illusionism. He further proposed that the classic epiphenomenalist objections from mental causation and phenomenal judgment may be substantially mitigated in participatory simulation if consciousness is the mode by which an external subject experiences the simulation and the interface translates the subject's intentions and judgments into lawful in-world neural precursors. The paper adopts that proposal as a coherent local architecture while distinguishing it from strict global epiphenomenalism.

Malpert additionally pressed for a deliberately broad evidential perimeter, including DMT entity encounters, psychedelic synesthesia, the reported laser-code phenomenon, claimed telepathy involving nonspeaking autistic children, and field-theory proposals, rather than excluding them because of fringe status. The methodological instruction was to apply the same evidential standards to these claims as to conventional neuroscience, separating unusual phenomenology, anomalous information transfer, and metaphysical interpretation.

ChatGPT developed the quantitative Bayesian structure, likelihood comparisons, dependence controls, literature synthesis, and the present prose. Claude supplied major adversarial critiques involving partition sensitivity, prediction versus accommodation, causal closure, decombination, anthropic conditionalization, Russellian monism, and psychophysical harmony, and elaborated the two-tier implementation after Malpert introduced the local/ultimate distinction. Grok supplied additional adversarial criticism on likelihood ranges, priors, dependence, simulation probability, and combination/decombination. The numerical estimates remain transparent judgment calls rather than calibrated frequencies.

For the simulation analysis, Malpert required a dedicated literature audit rather than a numerical increase unsupported by adversarial review. Two framings receive specific collaboration-level attribution. First, he proposed the "1890 counterfactual": ask how relativity, quantum mechanics, Bell-inequality violations, information bounds, and emergent-spacetime research would have updated a simulation or generated-reality hypothesis had such a hypothesis been taken seriously before those discoveries. Second, he pressed the temporal-typicality argument beyond the bare observation that the present era may be historically unusual: if civilization is near a first ASI transition, simulators interested in that focal branch point may branch and replay the transition interval many times relative to ordinary eras. After hearing Yampolskiy articulate the focal-interest point, Malpert emphasized the differential-replay mechanism. Turchin, Batin, Denkenberger, and Yampolskiy (2019) provide a close published precursor for singularity-focused simulations, while Thomas (2026) independently supplies an expectation-based self-location framework well suited to twenty-first-century reference classes. The contribution claimed here is therefore not historical priority for the underlying ideas, but their explicit integration into this paper's two-tier Bayesian architecture and their joint evaluation against the strongest counterarguments located in the literature.

Selected References

- Beane, S. R., Davoudi, Z., & Savage, M. J. (2014). Constraints on the universe as a numerical simulation. *European Physical Journal A*, 50, 148. <https://doi.org/10.1140/epja/i2014-14148-0>.
- Birch, J. (2013). On the "Simulation Argument" and selective scepticism. *Erkenntnis*, 78, 95-107. <https://doi.org/10.1007/s10670-012-9400-9>.

- Bostrom, N. (2003). Are you living in a computer simulation? *Philosophical Quarterly*, 53(211), 243-255. <https://doi.org/10.1111/1467-9213.00309>.
- Bostrom, N. (2009). The Simulation Argument: Some explanations. *Analysis*, 69(3), 458-461. <https://doi.org/10.1093/analys/anp063>.
- Bostrom, N., & Kulczycki, M. (2011). A patch for the Simulation Argument. *Analysis*, 71(1), 54-61. <https://doi.org/10.1093/analys/anq107>.
- Bousso, R. (2002). The holographic principle. *Reviews of Modern Physics*, 74, 825-874. <https://doi.org/10.1103/RevModPhys.74.825>.
- Braddon-Mitchell, D., & Latham, A. J. (2024). Ancestor simulations and the dangers of simulation probes. *Erkenntnis*, 89, 1257-1267. <https://doi.org/10.1007/s10670-022-00560-8>.
- Brueckner, A. (2008). The Simulation Argument again. *Analysis*, 68, 224-226. <https://doi.org/10.1111/j.1467-8284.2008.00742.x>.
- Campbell, T., Owahdi, H., Sauvageau, J., & Watkinson, D. (2017). On testing the simulation theory. *International Journal of Quantum Foundations*, 3, 78-99.
- Chalmers, D. J. (1996). *The Conscious Mind: In Search of a Fundamental Theory*. Oxford University Press.
- Chalmers, D. J. (2022). *Reality+: Virtual Worlds and the Problems of Philosophy*. W. W. Norton.
- Churchland, P. M. (1981). Eliminative materialism and the propositional attitudes. *Journal of Philosophy*, 78, 67-90.
- Crawford, L. (2013). Freak observers and the Simulation Argument. *Ratio*, 26, 250-264. <https://doi.org/10.1111/rati.12009>.
- Davis, A. K., Clifton, J. M., Weaver, E. G., Hurwitz, E. S., Johnson, M. W., & Griffiths, R. R. (2020). Survey of entity encounter experiences occasioned by inhaled N,N-dimethyltryptamine. *Journal of Psychopharmacology*, 34(9), 1008-1020.
- Edge, E., & Brown, C. A. (2026). Commentary: Astrophysical constraints on the simulation hypothesis for this Universe: why it is (nearly) impossible that we live in a simulation. *Frontiers in Physics*, 14, 1808725. <https://doi.org/10.3389/fphy.2026.1808725>.
- Frankish, K. (2016). Illusionism as a theory of consciousness. *Journal of Consciousness Studies*, 23(11-12), 11-39.
- Goler, D. (2025). Detailing a pilot study: The Code of Reality protocol, a phenomenon of N,N-DMT induced states of consciousness. *IPI Letters*, 3(1), N1-N5.
- Greene, P. (2020). The termination risks of simulation science. *Erkenntnis*, 85, 489-509. <https://doi.org/10.1007/s10670-018-0037-1>.
- Harris, K. R. (2024). The simulation argument reconsidered. *Analysis*, 84(1), 23-31. <https://doi.org/10.1093/analys/anad048>.
- Hensen, B., Bernien, H., Dreau, A. E., et al. (2015). Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres. *Nature*, 526, 682-686. <https://doi.org/10.1038/nature15759>.
- Hoffman, D. D. (2019). *The Case Against Reality*. W. W. Norton.
- Jackson, F. (1982). Epiphenomenal qualia. *Philosophical Quarterly*, 32(127), 127-136.
- Kastrup, B. (2019). *The Idea of the World*. Iff Books.
- Kipping, D. (2020). A Bayesian approach to the Simulation Argument. *Universe*, 6(8), 109. <https://doi.org/10.3390/universe6080109>.
- Lewis, P. J. (2013). The Doomsday Argument and the Simulation Argument. *Synthese*, 190, 4009-4022. <https://doi.org/10.1007/s11229-013-0245-9>.
- Lloyd, S. (2002). Computational capacity of the universe. *Physical Review Letters*, 88, 237901. <https://doi.org/10.1103/PhysRevLett.88.237901>.
- Malpert, R., with ChatGPT (GPT-5.6 Sol). (2026). *A Bayesian Analysis of Life After Death*. Unpublished manuscript, September 2026.
- Nobel Prize Outreach. (2022). The Nobel Prize in Physics 2022: Alain Aspect, John F. Clauser, and Anton Zeilinger, for experiments establishing violation of Bell inequalities and pioneering quantum information science.
- Robinson, W. (2026). Epiphenomenalism. *Stanford Encyclopedia of Philosophy*, Summer 2026 edition.
- Schlosser, R. W., et al. (2026). Systematic review of authorship in Rapid Prompting Method, Spelling to Communicate, and related variants. *Review Journal of Autism and Developmental Disorders*.
- Summers, M., & Arvan, M. (2022). Two new doubts about simulation arguments. *Australasian Journal of Philosophy*, 100(3), 496-508. <https://doi.org/10.1080/00048402.2021.1913621>.
- Takayanagi, T. (2025). Essay: Emergent holographic spacetime from quantum information. *Physical Review Letters*, 134, 240001. <https://doi.org/10.1103/pg4r-fy8n>.
- Thomas, T. (2026). Simulation expectation. *Erkenntnis*, 91, 427-444. <https://doi.org/10.1007/s10670-024-00901-9>.
- Timmermann, C., et al. (2018). DMT models the near-death experience. *Frontiers in Psychology*, 9, 1424.
- Timmermann, C., et al. (2023). Human brain effects of DMT assessed via EEG-fMRI. *Proceedings of the National Academy of Sciences*, 120(13), e2218949120.
- Turchin, A., Batin, M., Denkenberger, D., & Yampolskiy, R. (2019). Simulation typology and termination risks. arXiv:1905.05792. [Preprint.]
- Vazza, F. (2025). Astrophysical constraints on the simulation hypothesis for this Universe: why it is (nearly) impossible that we live in a simulation. *Frontiers in Physics*, 13, 1561873. <https://doi.org/10.3389/fphy.2025.1561873>.
- Weatherston, B. (2003). Are you a Sim? *Philosophical Quarterly*, 53(212), 425-431. <https://doi.org/10.1111/1467-9213.00323>.
- Wolpert, D. H. (2025). What computer science has to say about the simulation hypothesis. *Journal of Physics: Complexity*, 6, 045010. <https://doi.org/10.1088/2632-072X/ae1e50>.
- Yampolskiy, R. V. (2023). How to escape from the simulation. *Seeds of Science*. <https://doi.org/10.53975/wg1s-9j16>.

Yampolskiy, R. V. (2026, September 17). AI Safety Expert Roman Yampolskiy Warns About the Two Likely Endings with AI [Interview by J. Issa]. The Beyond Tomorrow Podcast.

Donadi, Sandro, et al. 'Underground Test of Gravity-Related Wave Function Collapse.' *Nature Physics* 17 (2021): 74-78. <https://doi.org/10.1038/s41567-020-1008-4>

Hameroff, Stuart, and Roger Penrose. 'Consciousness in the Universe: A Review of the Orch OR Theory.' *Physics of Life Reviews* 11 (2014): 39-78. <https://doi.org/10.1016/j.plrev.2013.08.002>

Pearce, David. Interview by Daniel Faggella. 'A Negative Utilitarian Vision for AGL.' *The Trajectory, Worthy Successor* episode 34, September 25, 2026. <https://danfaggella.com/pearce1/>

Evidence Audit and Proposed Revision: Morphic Resonance, Psi, and Metaphysical Inference

Rodney Malpert | Research addendum | 10 October 2026

Status and scope. This is an evidence-based amendment to the section “Morphic resonance and related proposals” in *A Two-Level Bayesian Analysis of Reality*. It preserves the existing 20-page model unchanged pending a complete re-estimation of all likelihood factors. The companion integrated file includes this addendum after the original paper; it is not a silently altered original.

Executive finding

The original paper treated morphic resonance as a high-information watch item without a scored positive multiplier. That was cautious but gave readers too little information about actual experiments. Positive findings have been published, notably in telephone-telepathy paradigms and reported owner-return anticipation in dogs. Independent and skeptical investigations also exist, however, and direct morphic-resonance tests have not converged on a robust result. Accordingly, the **recommended headline posterior remains unchanged**; we should replace the terse dismissal with this transparent evidence audit and sensitivity calculations. A nonzero numerical likelihood ratio should not be adopted without a more complete model of selection, bias, and reproducibility.

Evidence matrix: what has and has not been established

Telephone, email and SMS telepathy. Sheldrake, Stedall and Tressoldi (2025) pool 26 experiments in 15 papers, reporting an aggregate above-chance result (p approximately 10^{-7}) and no statistically detectable difference between investigator-linked studies and other replications. A nonsignificant between-group comparison does *not* prove equivalence. The meta-analysis is positive evidence for an anomaly under its assumptions, but risk of dependent trials, selective reporting, protocol heterogeneity and failure to identify a uniquely morphic mechanism still require careful attention.

Return-anticipating dogs. Sheldrake and Smart report that Jaytee waited at a window substantially more while the owner was returning; a separate series on Kane reported a comparable directional effect. Wiseman, Smith and Milton (1998) performed four independent Jaytee trials and concluded that their prespecified behavioral test did not support a psychic interpretation. Sheldrake disputes the stopping rule and argues that the same video record contains the predicted time-at-window pattern. This is a *conflict of operational definitions*, not an unqualified independent confirmation.

Birds: blue tits opening milk bottles. Regional spread, apparent independent invention, and postwar return of the milk-bottle-opening behavior are suggestive historical observations. Yet environmental opportunity, social learning, reinvention, reporting patterns, and lack of prospective randomization preclude a diagnostic likelihood ratio. Treat as hypothesis-generating, not independent experimental confirmation.

Rats learning across generations. McDougall reported improvement over generations; later studies including Crew and Agar did not neatly confirm the proposed acquired-inheritance mechanism. Improvements in trained and untrained controls are difficult to interpret; the data are not a prospective, separated-population demonstration of morphic transfer. A modern preregistered two-location experiment would discriminate mechanisms far better.

Direct morphic resonance: Wordle. Black, Butzer and Sheldrake (2026) tested whether global practice would reduce attempts needed later in the day. Study 1 did not support the primary predicted correlation;

Study 2 found a limited favorable pattern in first-two-guess success; Study 3 did not replicate it. Overall, the direct contemporary test is mixed and does not furnish a dependable positive numerical multiplier.

Controlled symbol-recognition experiments. Published efforts using Chinese characters or symbols provide a more directly testable alternative to natural-history anecdotes, but outcome and replication strength must be evaluated against the full paper and independent repeats before they can enter the model quantitatively.

Bayesian separation: anomaly, mechanism, ontology

Define A = reproducible anomalous information or learning effect; M = a Sheldrake-type morphic-resonance mechanism; O = the higher-level ontology (physicalism, Russellian monism, idealism, or dualism). Data may raise $P(A)$, without implying the same increase in $P(M)$, and even an established M would not uniquely entail idealism. Ordinary physicalist models could in principle be expanded with undiscovered laws; Russellian, idealist and dualist frameworks can also accommodate such effects. The relevant likelihood ratio for the existing ontology model is $P(\text{data} \mid O_i) / P(\text{data} \mid \text{conventional physicalism})$, not a p-value or a direct inference from a surprising anecdote.

Numerical sensitivity — illustrative only

Let the prior odds of an ontology relative to conventional physicalism be R. To show the scale rather than claim measurement, suppose new evidence yields a net likelihood ratio $L = 1.00, 1.05, 1.10, \text{ or } 1.25$ for that ontology compared with physicalism. Its posterior odds become $R \times L$. For a simple two-hypothesis baseline of 50%/50%, the corresponding probabilities are 50.0%, 51.2%, 52.4%, and 55.6%. These are **not** recalculations of the paper's five-ontology posterior: each ontology would need its own empirically defended likelihood and dependency corrections. The present adjudicated choice is $L=1.00$ for the headline pending such work, not because all observed results are proven null.

Recommended replacement text for Section IX

"Sheldrake-type morphic resonance posits nonlocal, history-dependent influences on the formation of patterns and behavior. This program has generated empirical tests rather than only philosophical speculation. A 2025 synthesis of telecommunication-telepathy studies reports an overall above-chance effect, and videotaped studies of dogs anticipating their owners' return report striking temporal associations. These findings warrant explicit attention. Nevertheless, the Jaytee record produced a substantive methodological dispute between Sheldrake and independent investigators; observational reports of blue-tit innovation and intergenerational rat learning do not securely isolate a nonlocal mechanism; and direct 2026 Wordle tests generated mixed, nonreplicating results. Even if some anomalous cognition were reliably established, it would not by itself validate morphic resonance or select uniquely among idealism, Russellian monism, dualism and an expanded physicalism. We therefore retain a neutral headline likelihood multiplier for now, not a conclusion that every experiment is negative. We document the positive results, competing interpretations and preregistered tests capable of changing that assessment. Sensitivity cases of 1.05, 1.10 and 1.25 are exploratory rather than adopted posterior inputs."

Priority discriminating experiments

The most valuable tests would (1) preregister caller-selection, data-exclusion rules, independent randomization and blind audit of automated telepathy trials across adversarially selected labs; (2) preregister continuous-video owner-return experiments with common primary endpoints agreed by both positive and skeptical researchers before data collection; and (3) run geographically separated, randomized parallel

learning tasks with sham-trained controls and concealed training start times. Report all attempted studies, raw protocols and null results.

Selected primary sources and qualifications

Sheldrake, R., Stedall, T., & Tressoldi, P. (2025). Telecommunication telepathy: A meta-analysis. *Journal of Anomalous Experience and Cognition*, 5, 47–69. <https://doi.org/10.31156/jaex.25934>

Wiseman, R., Smith, M., & Milton, J. (1998). Can animals detect when their owners are returning home? *British Journal of Psychology*, 89, 453–462. <https://doi.org/10.1111/j.2044-8295.1998.tb02696.x>

Sheldrake, R., & Smart, P. (2000). A dog that seems to know when his owner is coming home. *Journal of Scientific Exploration*, 14, 233–255. <https://www.sheldrake.org/research/animal-powers/abstract-a-dog-that-seems-to-know-when-his-owner-is-coming-home>

Black, G., Butzer, B., & Sheldrake, R. (2026). Memory within nature: A study of morphic resonance predictions on learning. *Journal of Anomalous Experience and Cognition*, 6. <https://doi.org/10.31156/jaex.27438>

Sheldrake, R. Rat learning and morphic resonance. <https://sheldrake.org/essays/rat-learning-and-morphic-resonance>

Sheldrake, R. Mind, memory, and archetype: Morphic resonance and the collective unconscious (blue-tit history). <https://www.sheldrake.org/research/morphic-resonance/part-i-mind-memory-and-archetype-morphic-resonance-and-the-collective-unconscious>

An empirical test of the theory of morphic resonance by using recognition for Chinese symbols. *Explore* (2010). <https://www.sciencedirect.com/science/article/abs/pii/S1550830710000820>

Audit limitation: This targeted evidence audit is not a complete systematic review, reanalysis of raw trial-level datasets, or independent replication. The paper's unchanged headline posterior reflects that limitation, not a declaration that the data are uninteresting.

Appendix B. Digital subjects, preservation, and the relation between the models

The Bostrom-Yampolskiy interview brings consciousness, personal continuation, and the treatment of digital minds into one accessible discussion. Its particular value here is to expose questions that the word "simulation" can conceal. Whether an environment is simulated, whether its inhabitants are conscious, what individuates those subjects, and whether they continue after an in-world death are separate propositions. Bostrom's willingness to take digital continuation seriously is a philosophical position to assess, not an observation that makes computational consciousness true.

B.1 From a model to an experiencing subject

A stored model, a running computation, a conversational persona, an agent pursuing objectives, and a phenomenal subject need not have the same boundaries. Stored weights may support many executions; one execution may portray multiple characters; several processes may cooperate as an agent. None of these bookkeeping facts, taken alone, determines the number of experiencing subjects. Conversely, one should not assume that the visible conversational character is merely a fiction while some other level necessarily contains the "real" subject. That also requires a theory.

This extends Section XI's combination and decombination problems. A theory needs to explain both how experience arises and why it belongs to this bounded subject rather than to a component, a larger whole, or several overlapping subjects. The engineering ability to copy a system is therefore distinct from a demonstrated ability to count subjects or preserve one particular subject through the copying.

The distinction is compatible with substantial digital moral standing. Bostrom and Shulman's Propositions Concerning Digital Minds and Society treats digital consciousness and emulation-based survival as serious possibilities. Our more pluralistic framework retains substrate-sensitive alternatives. On either approach, uncertainty about identity does not justify ignoring the interests of a resulting conscious being.

B.2 Generated minds can have a continuation mechanism

S2 means that the implementation generates the subject. It does not mean that the operator must erase that subject at the end of a simulated biological life. An operator could retain a causally relevant state, resume its process, preserve its history within a continuing experiencing system, or reconstruct a causally connected successor. These possibilities require distinctions among information retention, consciousness on resumption, psychological continuity, and strict numerical identity.

The interview's preservation discussion motivates this extension, but the extension is our application to the S1/S2 model. It is not evidence that our world has an operator, that its operator keeps backups, or that a backup is sufficient for the original subject to return. A retained recording that is never experienced does not satisfy the survival paper's

operational criterion.

There is also an asymmetry between saving a sufficiently specified digital state and reconstructing a person from damaged biological material. The former may allow relatively direct process resumption under a computational theory; the latter can require inference about missing details. Neither engineering description settles identity. The cryonics paper makes this distinction explicit rather than importing the apparent convenience of software copying into biological preservation.

B.3 Simulator ethics has two potentially opposing effects

Concern for simulated subjects could discourage creating painful conscious worlds. That could reduce the number of simulated observers of our kind. Conditional on a world nevertheless being created, the same concern could increase the probability of preservation, compensation, or a better subsequent life. A single "benevolent simulator" multiplier would hide this opposition.

Let Z describe operator type. Simulation prevalence depends partly on how often each type creates and sustains relevant observers. Continuation within simulation depends on the operator mix conditional on our observations. The correct averaging therefore uses $P(Z | \text{simulation, evidence})$, not simply the unconditioned prevalence of benevolent agents. Our observed suffering can affect that mix, but cannot be scored without specifying what each operator type predicts and whether later compensation is part of its policy.

Cheap computation strengthens a feasibility premise; it does not by itself determine conscious observer measure, actual use, simulator motives, or preservation policy. Likewise, one person's statement that they would run simulations illustrates a possible motive rather than estimating the preferences of mature civilizations. Those restrictions preserve the force of Yampolskiy's accessible intuitions while identifying the additional premises required by Bostrom-style population arguments.

B.4 Numerical continuity: a bridge, not copied percentages

The current survival benchmark is 46.56%, reported as about 47%. Its empirically updated participatory mass is 29.86%; generated simulation contributes another 9.67%, giving 39.54% total simulation within that model. The 32.5% total obtained from its starting inputs is a pre-update quantity. None of these should be compared indiscriminately with this paper's 42.15% Tier-1 scenario.

A transparent bridge can impose 42.15% total simulation while preserving the survival model's participatory/generated ratio within simulation and its architecture proportions outside simulation. This proportional reweighting gives 47.57% continuation if generated-mind continuation remains 2%, and 52.52% if it is 50%. These are coherence scenarios, not a newly estimated joint posterior. They show both that the two models can be connected and that the connection depends on declared choices.

The ontology weights, the simulation estimate, and the survival estimate answer different questions. The 26/40/23/10 rounded ontology weights can still be normalized for the

cryonics comparison; they do not determine simulator governance or the S1/S2 split. A fully empirical joint calibration remains unavailable. The models are continuous in their definitions and explicit dependencies, rather than artificially identical in every marginal percentage.

B.5 Consequence for the conclusion

The interview supplies no new measured likelihood ratio for physicalism, idealism, Russellian monism, or dualism. Their numerical comparison remains unchanged, as do the Tier-1 reference estimates. The substantial change is the consequence model: generated simulation need not be continuation-poor, and ethical restraint can reduce simulation prevalence while increasing preservation conditional on simulation. Those claims must be carried together into the survival and cryonics analyses.

Sources and related analysis

Source: Bostrom, Nick, interviewed by Roman Yampolskiy (2026). "Superintelligence Author: We May Already Have AGI | Nick Bostrom." The Roman Forum with Roman Yampolskiy, YouTube, October 16. <https://www.youtube.com/watch?v=m880ajt5sNM>. Interview discussion; conceptual claims and reported judgments, not an empirical calibration dataset.

Source: Bostrom and Shulman (2023), Propositions Concerning Digital Minds and Society, version 1.21, especially "Consciousness and Metaphysics" and "Respecting AI Interests." <https://nickbostrom.com/propositions.pdf>. Bostrom (2003), Are You Living in a Computer Simulation?, and Chalmers (2022), Reality+, supply the broader simulation framework discussed in the main references. The numerical bridge is derived here from the declared companion-paper inputs; it is not attributed to those authors.

About the Author

Rodney Malpert holds a J.D. and M.A. from Cornell University. His research uses human-AI collaboration and explicit probabilistic reasoning to examine philosophical, historical, and public-policy questions.

Website and current papers: RodMalpert.com