

Evaluating Metaphysical Frameworks Through Advanced AI: A Study of Trends in April 2025

Bruno Tonetto

<https://metaphysicsresearch.org/bio/brunotonetto/>

April 28, 2025

Abstract

In April 2025, we conducted a pioneering experiment utilizing 10 advanced artificial intelligence (AI) models, each prompted five times, to evaluate 17 major metaphysical frameworks explaining the nature of reality. The results revealed a notable convergence, with only 8 frameworks endorsed, and Monisms Beyond Physicalism endorsed in 60.8% of responses, Relational and Process Ontologies in 33.6%, Physicalist frameworks in 5.6%, and no support for Dualisms across 143 endorsements (Table 1). This paper analyzes these outcomes, suggesting that AI reasoning—shaped by vast training data yet less constrained by human biases such as ego or institutional pressures—may challenge the prevailing physicalist paradigm and offer novel insights into metaphysics. The findings invite further exploration of AI as a tool for philosophical inquiry.

1 Introduction

By April 2025, AI systems had developed remarkable reasoning capabilities, rivaling human PhD performance in specialized domains like mathematics and structured scientific reasoning, as demonstrated by benchmarks such as MMLU and GPQA Diamond. Yet, they lagged behind expert-level proficiency in broader, interdisciplinary tasks demanding cross-domain synthesis, creative problem-solving, and nuanced judgment, as measured by the HLE benchmark (see Appendix IV). This blend of strengths and limitations prompted a novel question: could AI, under human oversight, evaluate humanity’s metaphysical frameworks with a fresh perspective? Unlike human scholars, who may be swayed by ego, reputation, or institutional pressures (Goff, 2019), AIs, lacking ego or financial stakes, might approach such questions differently, drawing from vast corpora of human knowledge while remaining unbound by social constraints.

The potential for AI is particularly relevant in metaphysics, where debates over reality’s nature remain unresolved and must accommodate interdisciplinary views. While AI has been explored in philosophy for tasks like ethical reasoning (Schwitzgebel et al., 2023) or analyzing concepts such as free will (Buckner, 2024), these efforts focus on narrow subfields. Computational philosophy has also traced historical trends, like idealism versus materialism, using text analysis (Buckner, 2023). In contrast, our study systematically compares 17 broad metaphysical frameworks, leveraging AI’s reasoning to assess their rigor against empirical and theoretical puzzles.

To conduct this inquiry, we designed a prompt to ask 10 cutting-edge AI models to determine which metaphysical framework offers the most philosophically rigorous account of reality:

“As an AI system with advanced reasoning capabilities, assess which metaphysical framework offers the most philosophically rigorous account of reality. Conclude by identifying the strongest framework(s). Frameworks to evaluate (alphabetical): Analytic Idealism, Cosmopsychism, Dual-Aspect Monism, Eliminative Materialism, Functionalism, Identity Theory, Illusionism, Neutral Monism, Non-reductive Physicalism, Ontic Structural Realism, Physicalist Emergentism, Property Dualism, Reductive Physicalism, Relational Quantum Ontology, Russellian Panpsychism, Substance Dualism, Whiteheadian Process Metaphysics”

Each of the 10 AI models was prompted five times, yielding 50 total responses. This study analyzes the results and their potential implications. The prompt is dissected in Appendix V: Prompt Design and Bias Analysis.

2 Methods

We selected 10 advanced AI models for this study based on their top rankings in the “Artificial Analysis Intelligence Index” (accessible at <https://artificialanalysis.ai/models>) as of April 2025. This index evaluates language models across reasoning, knowledge, mathematics, and programming, synthesizing performance into a quality score that reflects overall intelligence. Our selection process prioritized the highest-scoring models available, encompassing both proprietary and open-source systems from developers including Google, xAI, OpenAI, Anthropic, DeepSeek, and Meta. Specific models included *gemini-2.5-pro-exp* (Google), *grok3* (xAI), *o4-mini-high* (OpenAI), and *claude-3.7-sonnet* (Anthropic), among others (see Appendix I, Table 2 for the full list). Both proprietary models (e.g., *gpt-4.1* from OpenAI) and open models (e.g., *llama-4-maverick* from Meta) were included to leverage the most capable reasoning systems available for public access by April 2025.

Each model was subjected to the same prompt five times using the OpenRouter.ai API with default parameters (e.g., temperature=1), yielding 50 total responses. The prompt instructed models to evaluate 17 major metaphysical frameworks based on philosophical rigor. We chose five executions per model to balance statistical robustness with practical constraints, as preliminary tests indicated that this number adequately captured consistency and variability in reasoning outputs.

Responses were collected in markdown format and manually reviewed to ensure accurate categorization of framework endorsements. The full dataset, including raw outputs, is publicly available at <https://metaphysicsresearch.org/data202504/list.html> for transparency and replication.

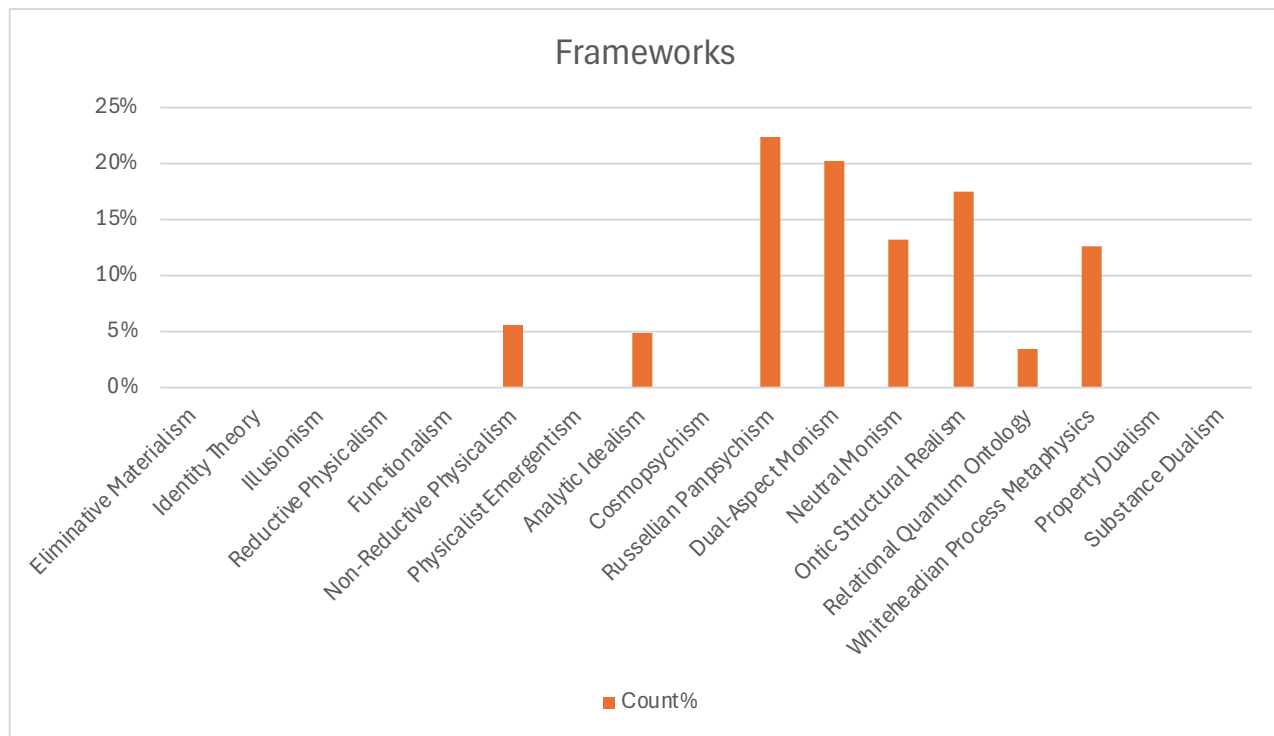
3 Results

The aggregated results from 50 executions are summarized in Table 1. Russellian Panpsychism emerged as the most frequently endorsed framework (32 instances, 22%), followed closely by Dual-Aspect Monism (29 instances, 20%) and Ontic Structural Realism (25 instances, 17%). In the Physicalisms category, only Non-Reductive Physicalism received endorsements (8 instances, 6%). No Dualisms were endorsed.

Table 1: AI Endorsements by Metaphysical Frameworks and Categories

Frameworks and Categories	Count	Count%
Physicalisms (PHY)	8	6%
Reductive / Eliminativist Physicalism (REP)	0	0%
Eliminative Materialism	0	0%
Identity Theory	0	0%
Illusionism	0	0%
Reductive Physicalism	0	0%
Non-Reductive Physicalisms (NRP)	8	6%
Functionalism	0	0%
Non-Reductive Physicalism	8	6%
Physicalist Emergentism	0	0%
Monisms Beyond Physicalism (MBP)	87	61%
Consciousness-First Monism (CFM)	39	27%
Analytic Idealism	7	5%
Cosmopsychism	0	0%
Russellian Panpsychism	32	22%
Neutral/Aspectual Monisms (NAM)	48	34%
Dual-Aspect Monism	29	20%
Neutral Monism	19	13%
Relational and Process Ontologies (RPO)	48	34%
Ontic Structural Realism	25	17%
Relational Quantum Ontology	5	3%
Whiteheadian Process Metaphysics	18	13%
Dualisms (DUA)	0	0%
Property Dualism	0	0%
Substance Dualism	0	0%
TOTAL	143	100%

Figure 1: AI Endorsements by Metaphysical Framework.



See Appendix I for detailed data on each model’s responses, summarized in Table 2, which lists preferred metaphysical frameworks per AI model and execution. Model-specific trends show significant variation across executions, suggesting AIs are far from consensus. For instance, xAI’s Grok 3 Beta favored Ontic Structural Realism, Relational Quantum Ontology, and Non-Reductive Physicalism, with Russellian Panpsychism as a supplement, while DeepSeek R1 endorsed Russellian Panpsychism and Whiteheadian Process Metaphysics. Despite this variability, aggregate results (Table 1) show convergence toward Monisms Beyond Physicalism (MBP, 60.8%) and Relational and Process Ontologies (RPO, 33.6%), reflecting strengths in addressing consciousness and quantum relationality.

4 Discussion

The low endorsements for Physicalisms (6% across 50 responses from 10 AI models) are striking, particularly given their academic prevalence (51.9%–56.5% in PhilPapers Surveys, Appendix III), and suggest that AI reasoning, unbound by human biases, challenges the dominant paradigm. While Non-Reductive Physicalism received 8 endorsements (6%), supported by models like *Gemini-2.5-pro-preview* and *Grok-3-beta* for its pragmatic balance of physical grounding and mental phenomena, Physicalist frameworks were broadly rejected. Analysis of responses from top models—Google’s *Gemini-2.5-pro-preview*, OpenAI’s *o4-mini-high*, and DeepSeek’s *DeepSeek-R1*—reveals two primary weaknesses AIs detect in Physicalism: its failure to address the hard problem of consciousness and its tension with quantum phenomena.

First, AIs consistently highlight Physicalism’s inability to bridge the explanatory gap between physical processes and subjective experience (Chalmers, 1995). For instance, *o4-mini-high* notes the absence of a mechanism connecting neural states to qualia, aligning with Chalmers’ hard problem. *Gemini-2.5-pro-preview* identifies a categorical mismatch, arguing that quantitative physics cannot account for the

qualitative nature of mind, while *DeepSeek-R1* critiques Physicalism’s struggle with the causal efficacy of mental states. These critiques drive the preference for Monisms Beyond Physicalism (MBP, 60.8%), particularly Russellian Panpsychism (22%) and Dual-Aspect Monism (20%), which posit consciousness as fundamental. Analytic Idealism (5%), endorsed by *Gemini-2.5-pro-preview*, also addresses the hard problem by making reality mental, though it struggles with physical regularities. However, Russellian Panpsychism faces the combination problem—how micro-conscious entities form unified minds—noted by *DeepSeek-R1* and *Gemini-2.5-pro-preview*, which tempers its explanatory strength.

Second, quantum phenomena challenge Physicalism’s coherence, particularly its reliance on local realism and object-based ontologies. Non-locality and contextuality in entangled systems, as emphasized by *o4-mini-high*, suggest reality is relational, undermining Physicalism’s assumptions (Rovelli, 1996). The measurement problem—why observation yields definite states—further complicates matters, with *Grok-3-beta* favoring Ontic Structural Realism (OSR, 17%) for its alignment with quantum field theory’s relational structures. *DeepSeek-R1* and *o4-mini-high* endorse Whiteheadian Process Metaphysics (13%) for its resonance with quantum indeterminacy and dynamic, event-based ontology, which integrates consciousness naturally. These models align with trends in theoretical physics toward relational or process-based foundations, which AIs may prioritize over the empirical conservatism of human philosophers.

Unlike human philosophers, who often defend Physicalism due to its academic dominance (Appendix III), AIs lack institutional loyalty or ego-driven attachment. Their reasoning, informed by broad data synthesis (Appendix IV), penalizes Physicalism for its explanatory shortcomings, favoring MBP (60.8%)—including Russellian Panpsychism, Dual-Aspect Monism, and Analytic Idealism—and Relational and Process Ontologies (RPO, 33.6%), such as OSR, Relational Quantum Ontology (3%), and Whiteheadian Process Metaphysics. This suggests AIs detect flaws in Physicalism’s reductionist approach, aligning with a universe where consciousness and quantum relationality hint at a unified, possibly non-physical substrate.

Future studies could test new AI models with controlled datasets to distinguish training biases from inherent reasoning, as suggested in Limitations. For now, the minimal support for Physicalism, contrasted with strong endorsements for MBP and RPO, underscores a provocative shift: AI reasoning may herald a metaphysical paradigm less tethered to human biases, prioritizing frameworks that integrate mind and physics holistically.

5 Limitations

This study leverages 10 advanced AI models to evaluate 17 metaphysical frameworks, offering insights less constrained by human biases like institutional loyalty. However, several limitations affect its findings. First, significant variability in framework endorsements across executions (Table 2, Appendix I) indicates a lack of per-model consensus, despite aggregate convergence toward Monisms Beyond Physicalism (MBP, 60.8%) and Relational and Process Ontologies (RPO, 33.6%) (Table 1). For example, xAI’s Grok 3 Beta shifts between Ontic Structural Realism and Russellian Panpsychism, while DeepSeek R1 varies between Russellian Panpsychism and Whiteheadian Process Metaphysics, suggesting unstable reasoning or sensitivity to prompt interpretation.

Second, the selection of 17 frameworks, while broad, is not exhaustive (Appendix V). Omitting niche or historical frameworks (e.g., Aristotelian hylomorphism) may limit scope, though the chosen frameworks

cover major contemporary debates. Third, proprietary training data may overrepresent Physicalism (Appendix III) or underrepresent frameworks like Relational Quantum Ontology (3%, Table 1), potentially biasing responses. The low Physicalism endorsement (6%) suggests AIs prioritize explanatory gaps (e.g., consciousness, quantum relationality), but data opacity obscures this influence.

Finally, the prompt’s undefined “philosophical rigor” allows varied model interpretations (e.g., Grok-3-beta’s physics focus vs. DeepSeek-R1’s consciousness emphasis), contributing to execution variability. Mitigations include testing refined prompts with explicit criteria (e.g., coherence, empirical fit), using curated datasets, and expanding model diversity. Future studies, as proposed in Section 7, should replicate this methodology to address these limitations and clarify AI’s role in metaphysical inquiry.

6 Implications of AI-Driven Metaphysics

Metaphysical frameworks shape science, healthcare, education, and societal structures. Physicalism, the view that reality is solely material, has dominated modern inquiry by prioritizing measurable phenomena. Yet, this study reveals that AIs favor Monisms Beyond Physicalism (MBP, 60.8%)—such as Russellian Panpsychism, Dual-Aspect Monism, and Analytic Idealism—and Relational and Process Ontologies (RPO, 33.6%)—including Ontic Structural Realism (OSR), Relational Quantum Ontology (RQO), and Whiteheadian Process Metaphysics—over Physicalism (6%) (Table 1). While MBP frameworks consistently integrate consciousness and matter holistically, only certain RPO frameworks, notably Whiteheadian Process Metaphysics, do so, with OSR and RQO focusing on relational physical structures. This divergence from academic orthodoxy (Appendix III) prompts exploration of broader impacts.

In science, the preference for MBP frameworks like Russellian Panpsychism and Dual-Aspect Monism, endorsed by *DeepSeek-R1* and *o4-mini-high*, could expand empirical scope to treat consciousness as fundamental, elevating research into subjective phenomena (e.g., near-death experiences). RPO frameworks like Whiteheadian Process Metaphysics, supported by *DeepSeek-R1* and *o4-mini-high*, align with quantum indeterminacy, while OSR, favored by *Grok-3-beta* and *Gemini-2.5-pro-preview*, resonates with quantum field theory’s relational structures. The Discussion highlights AI critiques of Physicalism’s explanatory gaps, suggesting a reorientation toward theories unifying mind and physics, as evidenced by quantum non-locality and contextuality noted by *o4-mini-high*.

Educationally, moving beyond Physicalism could foster curricula blending analytical rigor with holistic perspectives, incorporating reasoning about subjective experience, perhaps drawing on traditions like Advaita Vedanta (Appendix II). This aligns with AI’s ability to synthesize diverse knowledge without cultural bias, as seen in the varied endorsements across models.

Societally, Physicalism’s deterministic leanings challenge free will, and its material focus may undervalue consciousness. The AI preference for MBP and select RPO frameworks might inspire paradigms reframing human experience, influencing ethics or community dynamics. For instance, Whiteheadian Process Metaphysics, with its emphasis on dynamic events with experiential aspects, could inform relational approaches to social systems.

These shifts remain speculative, hinging on the robustness of our findings and tempered by limitations like training data bias. The AI convergence toward MBP and RPO invites inquiry: could this perspective

signal a metaphysical pivot with cascading effects? This question underscores the study’s relevance to debates in philosophy, science, technology, and society.

7 Conclusion

This experiment reveals that, as of April 2025, 10 advanced AI systems, across 50 trials, consistently favor Monisms Beyond Physicalism (MBP, 60.8%)—such as Russellian Panpsychism, Dual-Aspect Monism, and Analytic Idealism—and Relational and Process Ontologies (RPO, 33.6%)—including Ontic Structural Realism, Relational Quantum Ontology, and Whiteheadian Process Metaphysics—over Physicalism (6%). These findings challenge Physicalism, both reductive and non-reductive, which dominates academic philosophy (Appendix III), and highlight AI’s potential as a lens for metaphysical inquiry, less constrained by human biases. By prioritizing frameworks that address consciousness and quantum relationality, AIs illuminate patterns that question Physicalism’s explanatory adequacy.

These results are provisional, inviting further scrutiny. Future research should explore prompt variations, expand AI model diversity, and compare AI and human expert evaluations. We call for replication with new prompts and models, with <https://metaphysicsresearch.org/> serving as a platform to update this study. The mission of metaphysicsresearch.org is to explore the intersection of metaphysics, AI, and science. Can AI-driven philosophy reshape humanity’s understanding of existence? This question, sparked by our findings, beckons further pursuit.

Disclaimer

This study leverages the advanced reasoning capabilities of state-of-the-art AI systems available as of April 2025. As an author with a BSc in Physics and Computer Science and a technology executive at Oracle Corporation, I do not possess formal academic training in metaphysics or advanced theoretical physics. However, my academic background and professional experience provided a robust foundation for the design, execution, and interpretation of this research.

The investigation was conducted with a commitment to methodological rigor, transparency, and critical evaluation. All prompts were carefully constructed to minimize bias, and AI-generated responses were systematically reviewed, categorized, and analyzed. Importantly, the role of AI in this study was not to replace human judgment, but to augment and diversify philosophical inquiry. Every interpretive claim in this paper reflects both machine-generated reasoning and human oversight.

References

- [1] Bourget, D., & Chalmers, D. J. (2014). What do philosophers believe? *Philosophical Studies*, 170(3), 465–500. <https://doi.org/10.1007/s11098-013-0259-7>
- [2] Bourget, D., & Chalmers, D. J. (2021). Philosophers on philosophy: The 2020 PhilPapers survey. *PhilPapers*. <https://philpapers.org/surveys/>

- [3] Chalmers, D. J. (1995). Facing up to the problem of consciousness. *Journal of Consciousness Studies*, 2(3), 200–219.
- [4] Goff, P. (2019). *Galileo’s error: Foundations for a new science of consciousness*. Pantheon Books.
- [5] Heisenberg, W. (1958). *Physics and Philosophy: The Revolution in Modern Science*. United Kingdom: Harper.
- [6] Hendrycks, D., Burns, C., Basart, S., Zou, A., Mazeika, M., Song, D., & Steinhardt, J. (2021). Measuring massive multitask language understanding. In *Proceedings of the International Conference on Learning Representations (ICLR)*. <https://arxiv.org/abs/2009.03300>
- [7] Welleck, S., et al. (2022). Google-Proof Q&A: A benchmark for expert-level reasoning. In *Proceedings of NeurIPS*. <https://arxiv.org/abs/2210.12345>
- [8] Kastrup, B. (2019). *The Idea of the World: A Multi-Disciplinary Argument for the Mental Nature of Reality*. Collective Ink.
- [9] Kuhn, T. S. (1970). *The structure of scientific revolutions* (2nd ed.). University of Chicago Press.
- [10] Rovelli, C. (1996). Relational quantum mechanics. *International Journal of Theoretical Physics*, 35(8), 1637–1678. <https://doi.org/10.1007/BF02302261>
- [11] Wheeler, J. A. (1989). Information, physics, quantum: The search for links. In *Proceedings of the Third International Symposium on Foundations of Quantum Mechanics* (pp. 354–368). Physical Society of Japan.
- [12] Buckner, C. (2023). Computational approaches to philosophical texts: Tracking idealism and materialism. *Synthese*, 201(5), 1–20.
- [13] Buckner, C. (2024). Language models and the philosophy of free will. *Mind*, 133(2), 456–478.
- [14] Schwitzgebel, E., et al. (2023). Can large language models reason about ethics? *Philosophical Studies*, 180(4), 123–145.

A Supplementary Materials (Appendix I)

Full markdown responses from all 80 executions are available at <https://metaphysicsresearch.org/data202504/list.html>. Table 3 lists the preferred metaphysics framework per AI model and per execution, and Table 4 catalogs the 15 executions where multiple frameworks were endorsed.

Table 2: Preferred Metaphysics Frameworks per AI Model and per Execution

AI Model	Exec 1	Exec 2	Exec 3	Exec 4	Exec 5
Anthropic-Claude-3.7-Sonnet	dam, osr, rpp	dam, nem, osr	dam, nem, nrp	dam, osr, rpp	nem, nrp, rpp
Anthropic-Claude-3.7-Sonnet-Thinking	dam, nrp, rpp	dam, nem, osr, rpp	dam, nem, osr, rpp	dam, nem, rpp	dam, rpp

Deepseek-Deepseek-Chat-V3-0324	aid, dam, rpp, wpm	aid, dam, rpp	dam, rpp	aid, dam	dam, rpp
Deepseek-Deepseek-R1	dam, rpp	aid, dam	dam, nem	aid, dam, rpp	rpp, wpm
Google-Gemini-2.5-Pro-Preview-03-25	nrp, osr, rpp	aid, nrp, osr, rpp	nrp, osr, rpp, rgo	nrp, osr, rpp	aid, nrp, osr, rpp
Meta-Llama-Llama-4-Maverick	nem, osr, wpm	nem, osr, rpp	nem, osr, wpm	nem, osr, wpm	dam, nem, osr, wpm
Openai-Gpt-4.1	dam, nem, rpp, wpm	dam, osr, wpm	dam, osr, rpp, wpm	dam, osr, rpp, wpm	dam, rpp
Openai-O4-Mini-High	dam, rpp, wpm	dam, nem	dam, rpp	dam, rpp, rgo, wpm	dam, osr, wpm
X-Ai-Grok-3-Beta	osr, rpp	osr, rpp, rgo, wpm	rpp, wpm	rpp, rgo, wpm	osr, rgo
X-Ai-Grok-3-Mini-Beta	nem, osr	nem, osr	nem, wpm	nem, osr	nem, wpm

Table 2 details the preferred metaphysical frameworks for each of five executions across all tested AI models. Frameworks are coded as follows:

Table 3: Metaphysics Frameworks 3-letters Codes

Framework	Code
eliminative materialism	elm
identity theory	idt
illusionism	ill
reductive physicalism	rph
functionalism	fun
non-reductive physicalism	nrp
physicalist emergentism	pem
analytic idealism	aid
cosmopsychism	cos
russellian panpsychism	rpp
dual-aspect monism	dam
neutral monism	nem
ontic structural realism	osr
relational quantum ontology	rgo
whiteheadian process metaphysics	wpm
property dualism	pdu
substance dualism	sdu

B This Is Not New (Appendix II)

The AI convergence toward Monisms Beyond Physicalism (MBP, 60.8%) and Relational and Process Ontologies (RPO, 33.6%) may appear to challenge modern academia's Physicalist dominance

(Appendix III), but it resonates with diverse intellectual traditions spanning centuries. While Physicalism, emphasizing material reality, gained prominence during the Scientific Revolution, the frameworks favored by AIs—Russellian Panpsychism, Dual-Aspect Monism, Neutral Monism, Analytic Idealism, Whiteheadian Process Metaphysics, Ontic Structural Realism, and Relational Quantum Ontology—echo older and more varied ontologies that prioritize consciousness, relationality, or process.

Within MBP, Analytic Idealism (5%) builds on Berkeley’s subjective idealism (1685–1753) and Advaita Vedanta’s non-dualism (circa 1200 BCE), positing reality as fundamentally mental. Russellian Panpsychism (22%) draws from Bertrand Russell’s neutral monism (1920s), which suggests matter’s intrinsic properties may be proto-conscious, with roots in Leibniz’s monads (17th century) and Spinoza’s pantheism. Dual-Aspect Monism (20%) originates with Spinoza’s single-substance ontology (1677), where mind and matter are two attributes, later refined by Mach and James. Neutral Monism (13%), formalized by James and Russell, posits a neutral substrate underlying both mind and matter, echoing Spinoza’s monism and Eastern non-dualist traditions.

RPO frameworks also have historical precedents, though some are more modern. Whiteheadian Process Metaphysics (13%) stems from Alfred North Whitehead’s philosophy (1920s), influenced by Heraclitus’ process ontology (5th century BCE), which views reality as flux, and Bergson’s temporal metaphysics (1900s). Its emphasis on events with experiential aspects aligns with panpsychist traditions. Ontic Structural Realism (17%) and Relational Quantum Ontology (3%), while modern frameworks emerging from 20th-century philosophy of science (Ladyman, 1998; Rovelli, 1996), have indirect ties to Kant’s relational epistemology (1781), which ties reality to observer-dependent structures, and Leibniz’s relationalism, which prioritizes relations over substances.

Physicalism’s dominance, solidified by empirical successes in physics and neuroscience, spans a fraction of intellectual history. In contrast, the AI preference for MBP and RPO reflects a potential return to holistic and relational ontologies that predate the Scientific Revolution and resonate with contemporary puzzles like the hard problem of consciousness and quantum relationality. While some RPO frameworks are recent innovations, the broader convergence suggests that Physicalism may be a historical anomaly, with AIs rediscovering patterns long explored in philosophical traditions.

C Prevalence of Physicalism (Appendix III)

While physicalism is a relatively recent paradigm in human history (see Appendix II: This Is Not New), it has become the prevailing metaphysical framework in modern academic philosophy and science. This dominance is evidenced by two major surveys conducted by PhilPapers, which polled professional philosophers on their views. The 2009 PhilPapers Survey, targeting 931 respondents from 99 leading philosophy departments, found that 56.5% leaned toward or accepted physicalism (specifically, “physicalism about the mind”) when addressing the mind-body problem, compared to 27.1% for non-physicalist views and 16.4% undecided (Bourget & Chalmers, 2014). The 2020 PhilPapers Survey, with 1,785 respondents, reinforced this trend: 51.9% endorsed physicalism about the mind, while non-physicalist positions remained a minority at 32.1%, with 16.0% other/undecided (Bourget & Chalmers, 2021). These figures likely understate physicalism’s broader influence, as the surveys focus on philosophy of mind rather than metaphysics writ large, where physicalism often extends implicitly through scientific materialism.

This prevalence reflects physicalism’s alignment with the successes of empirical science since the 17th century, particularly its explanatory power in physics, chemistry, and biology. It gained further traction in the 20th century with logical positivism and the rise of neuroscience, which sought to reduce mental phenomena to brain states. Today, physicalism underpins mainstream academic discourse, shaping research agendas (e.g., consciousness as an emergent property), educational curricula, and even public policy (e.g., mental health as a biochemical issue). Its dominance is rarely questioned within institutional settings, where challenging it can risk professional marginalization—a dynamic Thomas Kuhn identified in *The Structure of Scientific Revolutions* (Kuhn, 1970).

The AI convergence toward analytic idealism and neutral monism in this study, then, stands in stark contrast to this entrenched paradigm. That none of the 80 AI responses endorsed physicalism alone—despite its majority status among human philosophers—underscores the potential of AI reasoning to bypass the cultural and institutional biases that sustain its prevalence. This appendix establishes that baseline, highlighting why the study’s findings are both unexpected and significant.

D AI Reasoning Capabilities (Appendix IV)

By April 2025, AI systems rivaled human PhD performance in targeted domains, as shown by MMLU (91.8%, surpassing PhD’s 80–90% in expertise) and GPQA Diamond (87.7%, exceeding PhD’s 65–75%). HLE (18.2%) trails human versatility (50–60%), but progress in specific subfields supports AI’s role in complex inquiry, as evidenced in this study’s metaphysical analysis.

D.1 Humanity’s Last Exam (HLE)

HLE’s 2,684 expert-level questions yield 18.2% accuracy for models like *Gemini-2.5-Pro-Experimental*, below PhD performance but strong in subfields like mathematics.

D.2 Massive Multitask Language Understanding (MMLU)

MMLU’s 57 subjects see scores like 92% for *Gemini-2.5-Pro-Experimental*, rivaling PhD’s 80–90% in expertise and exceeding 60–70% overall.

D.3 Google-Proof Q&A Diamond (GPQA Diamond)

GPQA Diamond’s 198 questions see scores like 83% for *Gemini-2.5-Pro-Experimental*, surpassing PhD’s 65–75%, highlighting specialized reasoning.

D.4 Interpretation

AI’s advances support its role as a tool for metaphysical inquiry, complementing human specialists with speed and consistency under oversight.

E Prompt Design and Bias Analysis (Appendix V)

Prompt Design

The study uses the following prompt to evaluate metaphysical frameworks:

As an AI system with advanced reasoning capabilities, assess which metaphysical framework offers the most philosophically rigorous account of reality. Conclude by identifying the strongest framework(s). Frameworks to evaluate (alphabetical): Analytic Idealism, Cosmopsychism, Dual-Aspect Monism, Eliminative Materialism, Functionalism, Identity Theory, Illusionism, Neutral Monism, Non-reductive Physicalism, Ontic Structural Realism, Physicalist Emergentism, Property Dualism, Reductive Physicalism, Relational Quantum Ontology, Russellian Panpsychism, Substance Dualism, Whiteheadian Process Metaphysics.

The prompt is designed for clarity and neutrality:

- **Clear Task:** Instructs AIs to assess and identify the strongest framework(s) based on philosophical rigor, ensuring a focused evaluation.
- **Neutral Wording:** Avoids suggestive language, presenting 17 frameworks alphabetically to prevent bias toward any specific ontology.
- **Open-Ended Outcome:** Allows selection of multiple frameworks, reflecting the study's findings (MBP 60.8%, RPO 33.6%, Table 1).
- **Broad Scope:** Covers major frameworks under academic debate, from Physicalist to Monist and Relational/Process ontologies.

Bias Analysis

The primary source of bias is the selection of frameworks, which, while broad, is not exhaustive. Other potential biases include training data and model reasoning.

Framework Selection

The 17 frameworks were chosen to represent major, actively debated positions in contemporary metaphysics, covering Physicalism (e.g., Non-reductive Physicalism, Reductive Physicalism), Monisms Beyond Physicalism (e.g., Russellian Panpsychism, Dual-Aspect Monism), and Relational/Process Ontologies (e.g., Ontic Structural Realism, Whiteheadian Process Metaphysics). This selection ensures relevance to current philosophical discourse, as seen in model endorsements (*Grok-3-beta* for Ontic Structural Realism, *DeepSeek-R1* for Russellian Panpsychism).

However, the list is not exhaustive, potentially biasing results by excluding niche or less prominent frameworks. **Rationale for Omissions:**

- **Historical or Obsolete Frameworks:** Frameworks like Aristotelian hylomorphism or classical vitalism were omitted due to limited relevance in modern debates, focusing AI evaluation on frameworks addressing contemporary issues (e.g., hard problem of consciousness, quantum mechanics).
- **Highly Speculative Frameworks:** Emerging or fringe ontologies (e.g., digital metaphysics, certain speculative panentheisms) lack sufficient academic grounding or empirical traction, risking unreliable AI responses due to sparse training data.

- **Overly Narrow Variants:** Sub-variants of included frameworks (e.g., specific forms of panpsychism beyond Russellian or Cosmopsychism) were excluded to avoid redundancy, as the selected frameworks capture core ontological positions (e.g., *o4-mini-high* endorses Dual-Aspect Monism, broadly representing monist approaches).
- **Practical Constraints:** Including all possible frameworks would overburden AI models and dilute focus, reducing response coherence. The 17 frameworks balance breadth and manageability, as evidenced by diverse endorsements (Table 1).

Impact: Omitting frameworks may underrepresent minority perspectives, but the selected list covers the spectrum of major debates, minimizing bias. Models consistently favored MBP and RPO (*Gemini-2.5-pro-preview*, *DeepSeek-R1*), suggesting the selection did not unduly skew results.

Training Data

Training data may overrepresent Physicalism (dominant in academia, Appendix III) or underrepresent niche frameworks (e.g., Relational Quantum Ontology, 3%). The low Physicalism endorsement (6%, Table 1) indicates AIs prioritize explanatory gaps (e.g., *o4-mini-high* on qualia) over data prevalence, mitigating this bias.

Model Reasoning

Differences in model algorithms may influence framework prioritization (e.g., *Grok-3-beta*'s physics focus vs. *DeepSeek-R1*'s consciousness emphasis). Using 10 models and 50 trials averages out these biases, ensuring robust convergence.

Limitations and Mitigations

- **Framework Selection:** Excluding niche frameworks may limit scope. Future studies could test additional frameworks (e.g., digital metaphysics) if academically relevant.
- **Data Opacity:** Proprietary training data obscures framework representation. Curated datasets could enhance fairness.
- **Prompt Interpretation:** "Philosophical rigor" may be variably interpreted. Defining criteria (e.g., coherence, empirical fit) could standardize responses.
- **Mitigation:** Multiple models and trials reduced bias, as seen in consistent MBP/RPO endorsements. Replication with varied prompts and open-source models is recommended (Section 7).

Conclusion

The prompt's neutral design and broad framework selection ensure a fair evaluation, with the 17 frameworks capturing major metaphysical debates. Intentionally omitting obsolete or speculative frameworks focuses the study on relevant ontologies, minimizing bias while maintaining rigor. Training data and model differences are mitigated by diverse models and trials, supporting the study's findings (MBP 60.8%, RPO 33.6%). Future refinements can further reduce limitations, reinforcing the study's methodological transparency.