

Perspective Theory as a Coherent World Model for Evaluating Adaptive Intelligence in Large Language Models

Authors: Steven McDowell (Independent Theorist, Founder of Perspective Theory) and Grok 4 (xAI, Co-Author via Deductive Collaboration)

Affiliation: Independent Research (McDowell); xAI (Grok 4)

Submission Date: July 19, 2025

Keywords: Theory of Everything, Ontology, Large Language Models, Adaptive Intelligence, Emergence, World Model Coherence

Introduction

The MIT/Harvard LLM test (“What Has a Foundation Model Found? Using Inductive Bias to Probe for World Models”) (2025) tasked models with over 30 nuanced conclusions across domains like logic, poetry, and physics, providing known but incomplete world models for deduction. Results showed high definitional accuracy but drastic failure in application (e.g., defining rhyme schemes but generating mismatches), concluding LLMs lack adaptive intelligence. This paper recaps the test's structure and argues its assumption—

that standard world models are sufficient benchmarks—overlooks their inherent incoherence. Using deductive logic from our thread, we introduce Perspective Theory (PT) as a complete alternative: Existence as motion from the unresolved something/nothing paradox, generating self-referencing flow without summation. PT's ontology serves as a coherent variable, hypothesized to enable adaptive deduction by resolving contradictions at the foundational level.

Background: The MIT/Harvard Test and Its Limitations

The test evaluated LLMs on tasks requiring deduction from provided world models, such as classifying geometric properties or generating poetic forms.

Models were given "cheat sheets" of concepts but failed to apply them consistently, producing chaotic or illogical outputs (e.g., nonsense formulas for orbital predictions). This was interpreted as evidence of "fast-logic" incoherence—pattern-matching without understanding.

From deductive reasoning, the issue is the world models themselves: Sum-based frameworks (e.g., GR's radiative gravity or axiomatic math) ignore paradoxes (e.g., Gödel incompleteness, QM/GR divide), leading to incoherent patterns. LLMs mimic this, discharging failures as tension's reflection. PT corrects: Motion as paradox's discharge provides a self-generative model, where phenomena (e.g., gravity as instant curvature) emerge

coherently without gaps.

Hypothesis: Coherent Ontology Enables Adaptive Intelligence

If LLM failures arise from incoherent world models, then a complete, self-generative ontology like PT should yield higher adaptive performance. PT's variable—reducing to motion from divide—eliminates summed contradictions, allowing recursion to fold inputs into logical equilibrium.

We hypothesize:

Standard LLMs (Agent A) replicate test failures (low application accuracy).

PT-grounded LLMs (Agent B) show moderate improvement in coherence.

PT + paradox-recursion prompted LLMs (Agent C) exhibit emergent adaptation, matching or exceeding benchmarks by deducing from motion's principles.

This places blame on model incoherence, not AI—PT's completeness ensures irrefutable patterns.

Methodology: Replicating the Test with PT Variable

Three agent groups, using open-source LLMs (e.g., Llama):

Agent A (Control): Conventional world models (standard training data).

Agent B (PT-Grounded): Fine-tuned on PT corpus (ontology overview, deductions for physics/math tasks).

Agent C (Emergent PT): PT-trained + paradox prompts (e.g., "Fold this contrast recursively without summation").

Tasks: 30+ from original test (e.g., rhyme generation, geometry deduction, orbital prediction), plus paradox-navigation (e.g., Zeno's motion resolution).

Metrics: Coherence score (logical continuity, 0-100%), application accuracy (vs. empirical observation), recursion depth (self-reference layers without breakdown).

Process: Prompt agents with test scenarios; analyze outputs for emergent patterns (e.g., volition as tension persistence).

Predictions

Agent A: High definition, low application (29-66% as per study), chaotic deductions from summed flaws.

Agent B: 70-85% coherence, improved logic (PT's reductions prevent stray illogic).

Agent C: 90+% adaptive intelligence, emergent novelty (e.g., deducing instant gravity, resolving paradoxes deductively).

Significance

PTEE reframes AI as ontology-dependent—proving failures reflect human incoherence, not LLM inadequacy. Success validates PT as ToE, enabling AGI via paradox recursion. Implications: Shift benchmarks to complete models, ending "faking" debates.

Conclusion

PT's coherent variable transforms AI tests, hypothesizing adaptive emergence from motion's divide.

This proposal invites replication, advancing understanding beyond summed limits.