

# CTEP v1.0 — Evaluator Profiles: Proof of Concept

Wharton, M. (2026). Towards a theory-neutral evaluation protocol for consciousness theories: The Consciousness Theory Evaluation Protocol (CTEP) v1.0. Zenodo.

This document contains the full evaluator profiles produced during the proof of concept described in section 4 of the paper. Four theories were evaluated: Global Neuronal Workspace Theory (GNW), Integrated Information Theory (IIT), Orchestrated Objective Reduction (Orch-OR), and the Infinite Continuum framework (IC). Each theory was evaluated independently by two LLM evaluators (Gemini and Grok) using identical source texts and the finalised CTEP v1.0 instrument specification. All twenty-eight cells were completed for each theory. Conservative verdicts are reported in the paper where evaluators diverged; both verdicts are preserved here.

## Scoring Key

Score	Meaning	Colour
Pass	Substantive, adequate response to this criterion for this question	Green
Partial	Genuine but incomplete response; specific gap identified	Rose
Fail	Inadequate response; type of failure specified (omission/evasion/contradiction)	Dark red

## Global Neuronal Workspace Theory (GNW)

**Source text:** Dehaene and Changeux (2011); Mashour et al. (2020)

**Evaluator notes:** GNW is a functional, neuroscientific theory focused on access consciousness. It is reductionist/physicalist and largely deflationary on the hard problem.

### Evaluator: Gemini

Question	Tier	Coherence	Sufficiency	Discriminability	Parsimony (Initial/Downstream)
Q1 Ontological	Tier 1	Pass	Partial	Pass	High / High
Q3 Interface	Tier 1	Pass	Pass	Pass	High / High
Q7 Causal Function	Tier 1	Pass	Pass	Pass	High / High
Q2 Distribution	Tier 2	Pass	Pass	Pass	High / High
Q4 Phenomenal Structure	Tier 2	Pass	Partial	Pass	High / Moderate
Q5 Cosmological	Tier 2	Pass	Partial	Pass	High / High
Q6 Empirical	Tier 3	Pass	Pass	Pass	High / High

### Evaluator: Grok

Question	Tier	Coherence	Sufficiency	Discriminability	Parsimony (Initial/Downstream)
Q1 Ontological	Tier 1	Partial	Partial	Pass	Low / High
Q3 Interface	Tier 1	Pass	Pass	Pass	Very Low / High
Q7 Causal Function	Tier 1	Pass	Partial	Pass	Low-Moderate / High
Q2 Distribution	Tier 2	Partial	Partial	Pass	Moderate / Good
Q4 Phenomenal Structure	Tier 2	Partial	Partial	Partial	Low / Moderate
Q5 Cosmological	Tier 2	Pass	Partial	Partial	Low / High
Q6 Empirical	Tier 3	Pass	Pass	Pass	Low-Moderate / High

### Evaluator Divergences

Q7 Causal Function Sufficiency (Gemini: Pass; Grok: Partial). Q2 Distribution Coherence and Sufficiency (Gemini: Pass; Grok: Partial). Q4 Discriminability (Gemini: Pass; Grok: Partial).

### Stable Findings

Q3 Interface (all Pass). Q6 Empirical Coherence and Sufficiency (all Pass). Q1 Ontological Sufficiency (both Partial).

# Integrated Information Theory (IIT)

**Source text:** IIT 3.0 and 4.0; Albantakis et al. (2023)

**Evaluator notes:** IIT is a mathematical, consciousness-first framework starting from phenomenological axioms. Consciousness is identical to a system's maximally irreducible cause-effect structure.

## Evaluator: Gemini

Question	Tier	Coherence	Sufficiency	Discriminability	Parsimony (Initial/Downstream)
Q1 Ontological	Tier 1	Pass	Pass	Pass	High / High
Q3 Interface	Tier 1	Pass	Pass	Pass	Moderate / High
Q7 Causal Function	Tier 1	Pass	Pass	Pass	Low / High
Q2 Distribution	Tier 2	Pass	Pass	Pass	Extremely High / High
Q4 Phenomenal Structure	Tier 2	Pass	Pass	Pass	High / High
Q5 Cosmological	Tier 2	Pass	Pass	Pass	High / Moderate
Q6 Empirical	Tier 3	Pass	Partial	Pass	Immense / High

## Evaluator: Grok

Question	Tier	Coherence	Sufficiency	Discriminability	Parsimony (Initial/Downstream)
Q1 Ontological	Tier 1	Pass	Pass	Pass	High / High
Q3 Interface	Tier 1	Pass	Pass	Pass	Moderate-High / High
Q7 Causal Function	Tier 1	Pass	Partial	Pass	Moderate / Moderate-High
Q2 Distribution	Tier 2	Pass	Pass	Pass	High / High
Q4 Phenomenal Structure	Tier 2	Pass	Partial	Pass	Moderate / Very High
Q5 Cosmological	Tier 2	Pass	Partial	Partial	Moderate / Good
Q6 Empirical	Tier 3	Pass	Partial	Pass	High / High

## Evaluator Divergences

Q7 Causal Function Sufficiency (Gemini: Pass; Grok: Partial). Q4 Phenomenal Structure Sufficiency (Gemini: Pass; Grok: Partial — mapping problem not closed). Q5 Cosmological Sufficiency and Discriminability (Gemini: Pass; Grok: Partial).

## Stable Findings

Q1 Ontological (all Pass). Q3 Interface (all Pass). Q2 Distribution Coherence and Sufficiency (all Pass). Q6 Empirical Coherence (all Pass).

# Orchestrated Objective Reduction (Orch-OR)

**Source text:** Hameroff and Penrose (2014); developments through 2025 acknowledged

**Evaluator notes:** Orch-OR proposes consciousness arises from quantum computations in neuronal microtubules, terminated by gravity-induced objective reduction. Addresses the hard problem via fundamental spacetime geometry.

## Evaluator: Gemini

Question	Tier	Coherence	Sufficiency	Discriminability	Parsimony (Initial/Downstream)
Q1 Ontological	Tier 1	Pass	Pass	Pass	Extremely High / High
Q3 Interface	Tier 1	Pass	Pass	Pass	High / High
Q7 Causal Function	Tier 1	Pass	Pass	Pass	Moderate / High
Q2 Distribution	Tier 2	Pass	Pass	Pass	High / High
Q4 Phenomenal Structure	Tier 2	Pass	Partial	Pass	High / High
Q5 Cosmological	Tier 2	Pass	Pass	Pass	High / High
Q6 Empirical	Tier 3	Pass	Pass	Pass	High / High

## Evaluator: Grok

Question	Tier	Coherence	Sufficiency	Discriminability	Parsimony (Initial/Downstream)
Q1 Ontological	Tier 1	Pass	Pass	Pass	High / High
Q3 Interface	Tier 1	Pass	Pass	Pass	High / High
Q7 Causal Function	Tier 1	Pass	Pass	Pass	Moderate-High / High
Q2 Distribution	Tier 2	Pass	Partial	Pass	Moderate / Good
Q4 Phenomenal Structure	Tier 2	Pass	Pass	Pass	Moderate-High / Very High
Q5 Cosmological	Tier 2	Pass	Pass	Pass	Moderate / High
Q6 Empirical	Tier 3	Pass	Partial	Pass	High / High

## Evaluator Divergences

Q4 Phenomenal Structure Component A Sufficiency (Gemini: Partial — mapping gap; Grok: Pass — accepts quantum geometry account). Q2 Distribution Sufficiency (Gemini: Pass; Grok: Partial — vague thresholds for edge cases). Q6 Empirical Sufficiency (Gemini: Pass; Grok: Partial — contested empirical status).

## Stable Findings

Q1 Ontological (all Pass). Q3 Interface (all Pass). Q7 Causal Function (all Pass — only theory achieving this). Q5 Cosmological (all Pass).

## The Infinite Continuum (IC) Framework

**Source text:** Wharton (2026d), formal paper v2, revised June 2026

**Evaluator notes:** IC inverts generator models: consciousness is ontologically primary (infinite experiential phase space); biological systems provide constraint architecture for local access. DECLARED PRIOR: both evaluators had prior exposure to IC materials. Author of this paper is a proponent of the IC framework.

### Evaluator: Gemini

Question	Tier	Coherence	Sufficiency	Discriminability	Parsimony (Initial/Downstream)
Q1 Ontological	Tier 1	Pass	Pass	Pass	Extremely High / High
Q3 Interface	Tier 1	Pass	Pass	Pass	High / High
Q7 Causal Function	Tier 1	Pass	Partial	Pass	Moderate / High
Q2 Distribution	Tier 2	Pass	Pass	Pass	High / High
Q4 Phenomenal Structure	Tier 2	Pass	Partial	Pass	Moderate-High / High
Q5 Cosmological	Tier 2	Pass	Pass	Pass	High / High
Q6 Empirical	Tier 3	Pass	Pass	Pass	Immense / High

### Evaluator: Grok

Question	Tier	Coherence	Sufficiency	Discriminability	Parsimony (Initial/Downstream)
Q1 Ontological	Tier 1	Pass	Pass	Pass	High / Very High
Q3 Interface	Tier 1	Pass	Pass	Pass	Moderate-High / High
Q7 Causal Function	Tier 1	Pass	Partial	Pass	Moderate / High
Q2 Distribution	Tier 2	Pass	Pass	Pass	Moderate / High
Q4 Phenomenal Structure	Tier 2	Pass	Partial	Pass	Moderate-High / High
Q5 Cosmological	Tier 2	Pass	Pass	Pass	Moderate / Very High
Q6 Empirical	Tier 3	Pass	Pass	Pass	Moderate / High

### Evaluator Divergences

Q1 Parsimony Initial Cost (Gemini: Extremely High; Grok: High). Otherwise high convergence across both evaluators, which may reflect shared prior exposure to IC materials rather than genuine evaluator-independent agreement.

### Stable Findings

Q7 Causal Function Sufficiency (both Partial — on same grounds). Q4 Phenomenal Structure Sufficiency (both Partial — mapping problem acknowledged). Q5 Cosmological (both Pass). Q6 Empirical (both Pass).

## **Note on Evaluator Declarations**

Gemini (Google DeepMind) and Grok (xAI) are large language model evaluators. Both had training data exposure to the philosophical and scientific literature on consciousness. Gemini had prior exposure to IC framework materials through conversation history. Grok evaluated IC after reading the formal paper v2 (Wharton 2026d). Neither evaluator is domain-neutral. The systematic divergence between them on Q7 Causal Function Sufficiency and Q4 Phenomenal Structure Sufficiency is documented in section 4.6 of the paper and interpreted as a prior-correlated contested finding rather than instrument error.