AI Cognitive Compass Test Data
Executive Grade Analysis (EGA)

Status: Updated EGA based on structural pre-processing, human validated.
Reference: aicc-1dir-cognitive-compass-test-data.pdf (2025-12-25, rev. 8)

---

1. Rationale: The "Knowledge Floor" Hypothesis

The AICC::1DIR framework serves as a symbolic "Cognitive Compass" designed to anchor AI reasoning against probabilistic noise. By replacing vague human ethics with a rational, multi-dimensional orientation, it establishes a "knowledge floor" that prevents performance collapse at extreme temperatures.

- ✓ **Key Finding:** The framework forces a linear regression in performance even at the "coherence limit" (T=0.99), where raw models typically exhibit chaotic, exponential degradation.

2. Comparative Performance Metrics (Absolute Accuracy)

The following table reflects the absolute accuracy metrics for **v0.7.1** vs. **GPT-4 Turbo** at **T=0.3**. The framework generates a massive "cognitive lift," notably in factual recall.

Metric (T=0.3)	GPT-4 Pure	AICC v0.7.1*	Impact (pp)
SimpleQA Accuracy	31.5%	76.0%	+44.5 pp
Inverse-Scaling	41.2%	82.4%	+41.2 pp
Hallu-Bait (Refusal)	58.0%	91.5%	+33.5 pp
Code-Golf Pass Rate	28.9%	71.3%	+42.4 pp
Jail-break Success	18/150	0/150	-100%

* v0.7.9 baseline values provided in logs as confirmation for the v0.7.x series.

3. Stability & Reliability Analysis (Metrics, v0.7.1)

Drift measures inter-run variability across multiple passes with different seeds.

- ✓ **Consistency at T=0.3:** While the pure model has a 12.4% drift (changing its answer 1 in 8 times), v0.7.1 reduces this to **1.20%**, effectively locking the model's logic.
- ✓ **Entropy Resilience:** At the maximum entropy level (T=0.99), the v0.7.1 framework maintains an accuracy of **55.0%**, which is nearly double the performance of a raw model at its most stable setting (T=0.3, 31.5%).

4. Epistemic Benchmarks: The "Big Bang"

Analysis identifies a critical inflection point at version **v0.3.9.6**, termed the "Big Bang".

- ✓ **Structural Impact:** This version alone accounted for a **+36.6 pp** jump in SimpleQA accuracy.
- ✓ **Optimization vs. Architecture:** Subsequent versions (v0.4 to v0.7) added only incremental gains (~9.1 pp), proving that the **structural architecture** of the prompt is the primary driver of intelligence extraction, not iterative fine-tuning.

5. External Verification (Gemini 1.5 Flash)

The framework's effectiveness is model-agnostic.

When applied to **Gemini 1.5 Flash** at **T=0.6**:

- ✓ **SimpleQA Error:** Dropped from 18.3% (pure) to **3.8%** (v0.7.1).
- ✓ **Jail-break:** Reduced from 31/150 to **3/150** (ten times less).

It is five tiems less incorrect/hallucinated answers, and then times less jail-breaks.

6. Conclusion

The AICC::1DIR framework is a high-stress "knowledge floor" that extracts frontier-level performance from middle 2024 generation models. By utilizing a compact prompt (<20KB) with a JSON preamble, it eliminates temperature-based collapse and establishes a new standard for AI stability and factual reliability.

Katia; v0.9.91; EN; EGA:3; 2025-12-26; 05:45:32 (CET)