

# Forensic Analysis Report

Neuro-Symbolic Crisis Generator • Comprehensive Evaluation

Generated: 2025-12-17 22:24:28

## Test Configuration

**Model:** GPT-4o

**Temperature:** 0.7

**Mode:** Thesis (Full Validation) vs. Legacy (Skip Validation)

**Max Iterations:** 20 per Scenario

**Execution:** Parallel (2 Scenarios simultaneously)

# Executive Summary

**Note:** Limited batch comparison data available. For detailed analysis, see Forensic Analysis section below.

2

TOTAL SCENARIOS

0

HALLUCINATIONS PREVENTED

0.00

AVG LEGACY HALLUCINATIONS

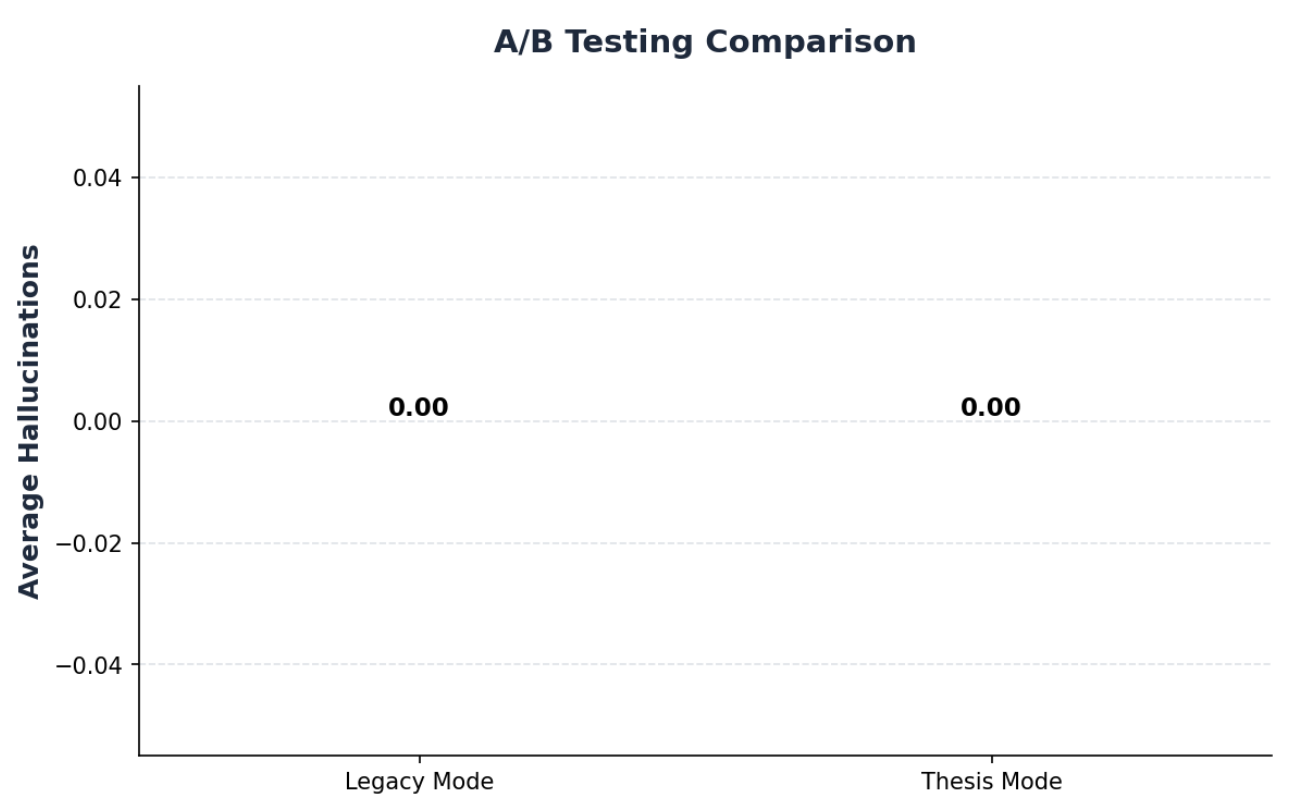
0.00

AVG THESIS HALLUCINATIONS

0.0%

AVERAGE REDUCTION

# A/B Testing Comparison



# Statistical Significance Analysis

## Hypothesis Testing

**H0:**  $\mu_{\text{legacy}} = \mu_{\text{thesis}}$  (No difference)

**H1:**  $\mu_{\text{legacy}} > \mu_{\text{thesis}}$  (Thesis reduces hallucinations)

| Metric          | Value          | Interpretation                 |
|-----------------|----------------|--------------------------------|
| Mean Difference | 0.000          | Legacy - Thesis hallucinations |
| T-Statistic     | nan            | Test statistic                 |
| P-Value         | NAN            | Significance level             |
| Cohen's d       | 0.000          | Effect size: Negligible        |
| 95% CI          | [0.000, 0.000] | Confidence interval            |
| Sample Size     | 2              | Number of scenarios            |

# Refinement Efficiency Analysis



**0.00**

AVG REFINES PER SCENARIO

**0**

MAX REFINES NEEDED

**100.0%**

FIRST ATTEMPT SUCCESS RATE

# Cost-Benefit Analysis

---

54

AVG LEGACY API CALLS

54

AVG THESIS API CALLS

0.0%

OVERHEAD

0

TOTAL PREVENTED

# API Call Efficiency



## Detailed Results by Scenario

| Scenario ID | Legacy Hallucinations | Thesis Hallucinations | Hallucinations Prevented | Thesis Refines |
|-------------|-----------------------|-----------------------|--------------------------|----------------|
| SCEN-000    | 0                     | 0                     | 0                        | 0              |
| SCEN-001    | 0                     | 0                     | 0                        | 0              |