

# Forensic Analysis Report

Neuro-Symbolic Crisis Generator • Comprehensive Evaluation

Generated: 2025-12-17 22:18:55

## Executive Summary

---

**2**

TOTAL SCENARIOS

**0**

HALLUCINATIONS PREVENTED

**0.00**

AVG LEGACY HALLUCINATIONS

**0.00**

AVG THESIS HALLUCINATIONS

**0.0%**

AVERAGE REDUCTION

# A/B Testing Comparison

---

**A/B Testing Comparison**



# Statistical Significance Analysis

## Hypothesis Testing

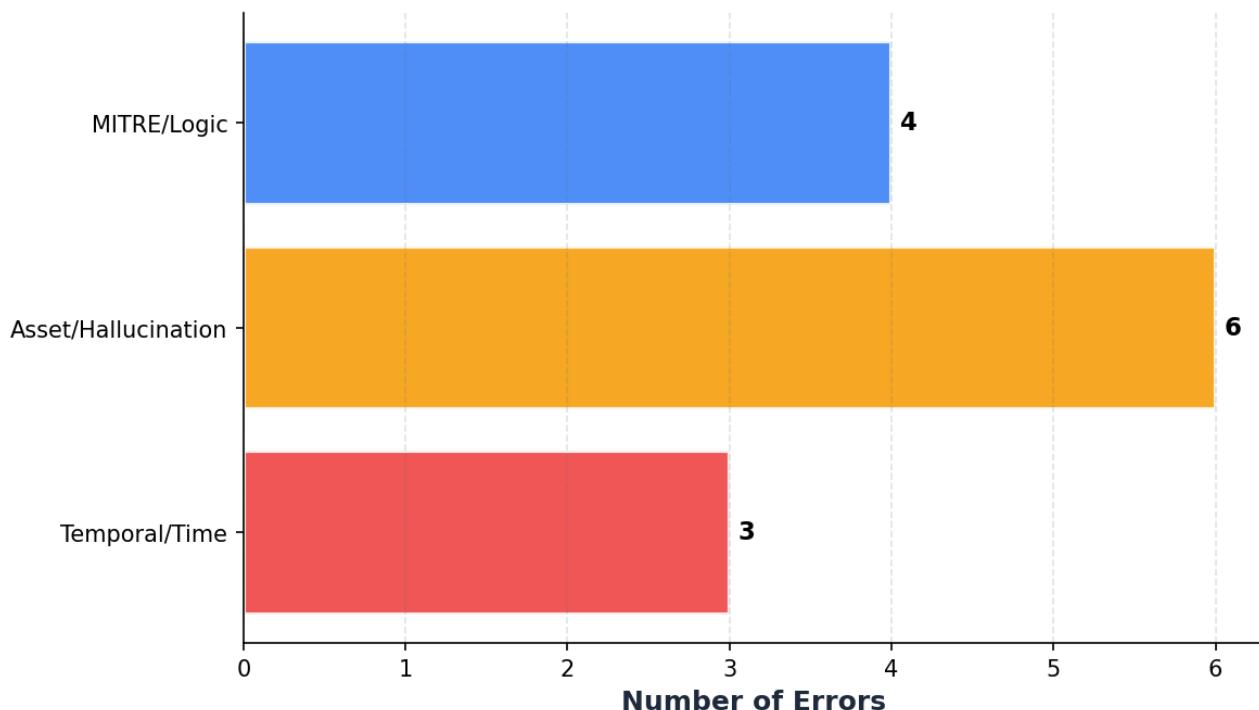
**H0:**  $\mu_{\text{legacy}} = \mu_{\text{thesis}}$  (No difference)

**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

# Qualitative Error Pattern Analysis

Error Type Distribution

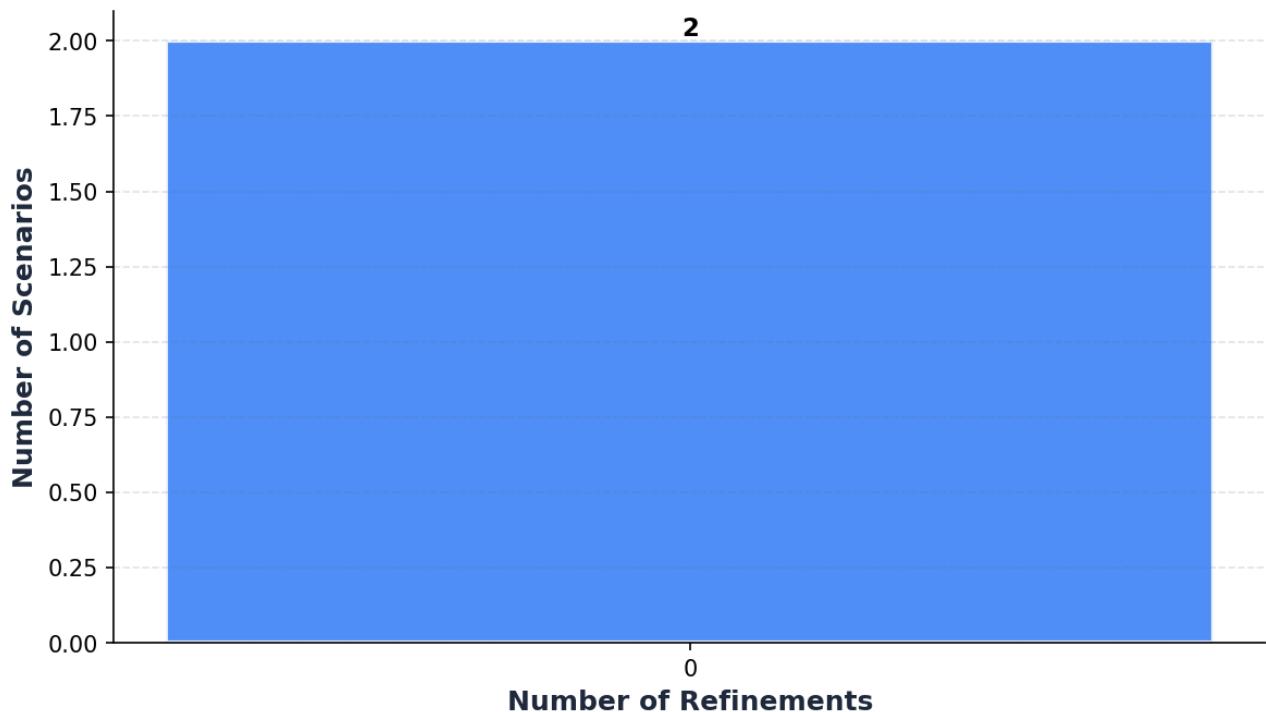


CATEGORY	COUNT	THESIS INTERPRETATION
Asset/ Hallucination	6	LLM invents non-existent assets. System enforces reality.
MITRE/Logic	4	LLM violates cybersecurity rules. System enforces domain knowledge.
Temporal/Time	3	LLM creates temporal paradoxes. System enforces causality.

# Refinement Efficiency Analysis

---

Refinement Distribution



**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

## Forensic Analysis Details

---

**45**

TOTAL EVENTS

**18**

UNIQUE INJECTS

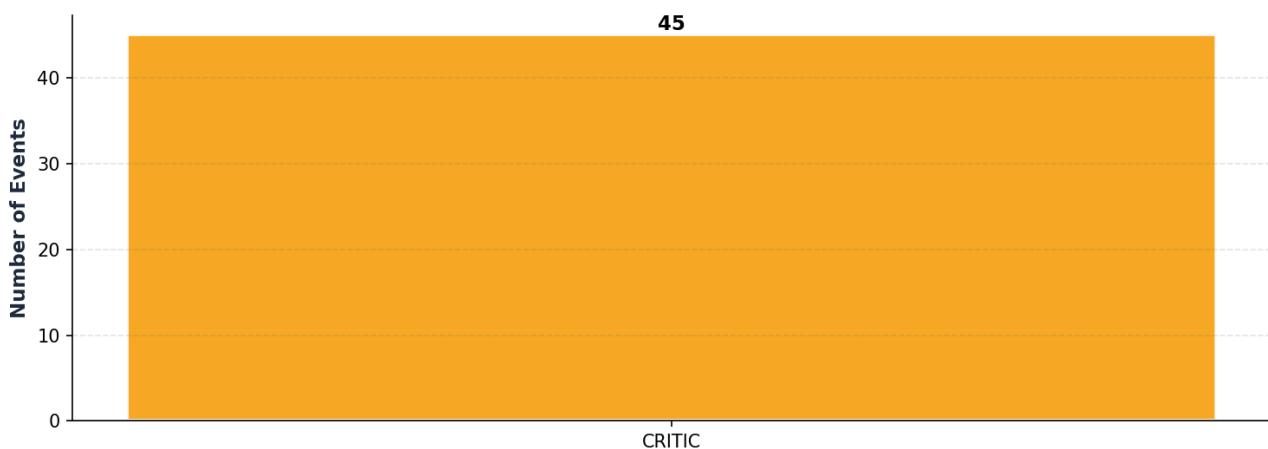
**20.0%**

REJECTION RATE

**0.50**

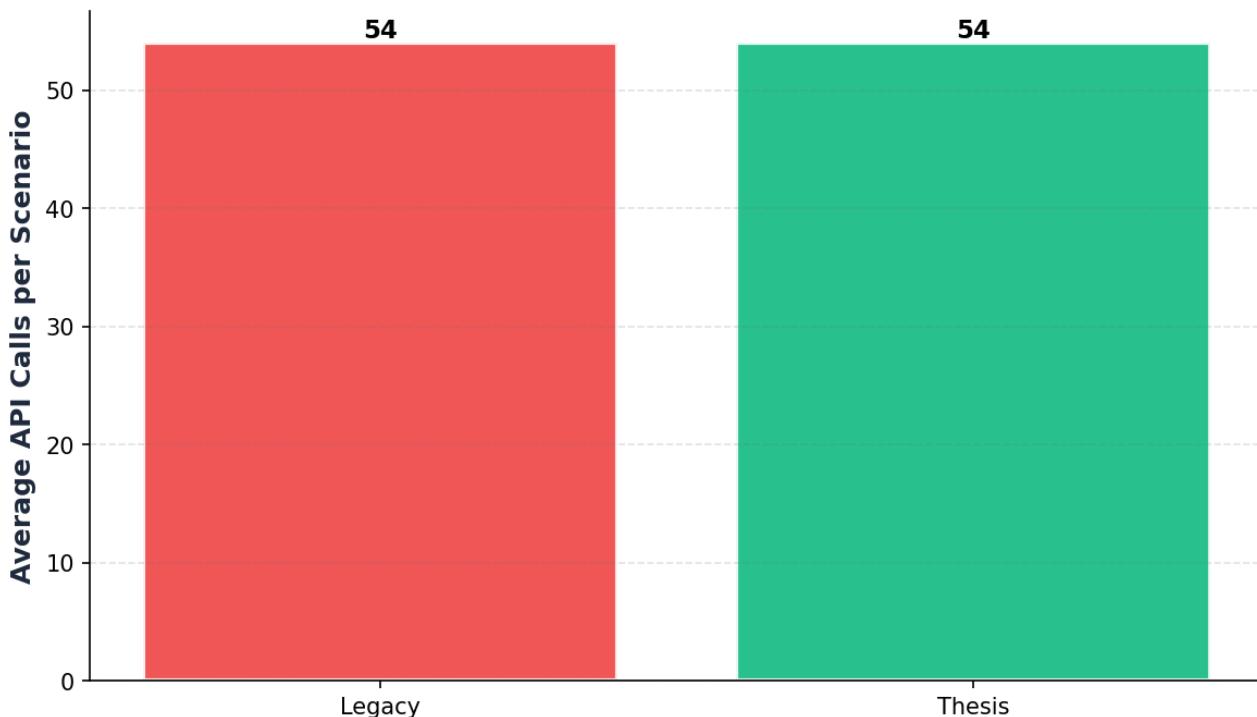
AVG REFINES PER INJECT

**Event Type Distribution**



# API Call Efficiency

API Call Efficiency Comparison



## 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