

Stage 3a Calibration Report

STAGE 3a VERDICT

Protocol Decisions & Evidence Synthesis for Stage 3b

RUN METADATA

- **Controllers:** SAC, TD3, TAD3
- **Run ID:** offpolicy-phase3a-v3
- **Completed At:** None

EVALUATION GRID PROTOCOL

- **Design:** 15 Sobol configs × 7 checkpoints
- **Checkpoints:** 50k, 100k, 200k, 300k, 500k, 750k, 1M
- **Replication:** 5 seeds × 2 conditions × 2 eps = 20/cp

1. Executive Protocol Verdicts

The table below consolidates the pre-registered decision rules and combination policies specified in ExperimentDesignRevision.md. Stage 3b search consumes these exact parameters.

Decision main	Do-	Calibrated Verdict	Evidence & Operational Rationale	Status
Timestep get	Bud-	750,000 steps	Max plateau across controllers (SAC, TD3, TAD3). Preserves fair computational allotment.	AUTO
Pruning	Strat-	PRUNE (ASHA) (200 trials, $\eta = 2$)	Combined across controllers. Pruning signal: progress_metric. Requires unanimous consensus across algorithms.	AUTO
Disturbance Aggregation	Ag-	composite_only	Min $\rho(\text{no_dist, comp}) = 0.507$. Threshold: $\rho < 0.8 \rightarrow \text{composite_only}$.	AUTO
Objective Metric		Manual Sign-Off	Requires manual review of per-algorithm correlations below.	REVIEW
Episodes per Cell	per	Manual Sign-Off	Requires manual review — compare se_of_config_new against between_config_sd, per algorithm.	REVIEW

2. Data Quality & Sweep Integrity Audit

Stage 3a protocol requires strict verification that evaluations are balanced and not floor-camping on crash states before correlations can be trusted.

Algorithm: SAC — 2,700 total episodes recorded

HEALTHY SWEEP

- **Floor-Camping Fraction:** 26.7% of configurations achieved < 5% success rate.
- **Sweep Balance:** No balance departures detected.

Algorithm: TD3 — 2,700 total episodes recorded

HEALTHY SWEEP

- **Floor-Camping Fraction:** 26.7% of configurations achieved < 5% success rate.
- **Sweep Balance:** No balance departures detected.

Algorithm: TAD3 — 2,700 total episodes recorded

HEALTHY SWEEP

- **Floor-Camping Fraction:** 33.3% of configurations achieved < 5% success rate.
- **Sweep Balance:** No balance departures detected.

3. Analysis (a) — Objective Metric & Disturbance Sensitivity

Analysis (a) tests whether episodic return tracks true flight precision (RMSE, attitude stability, control smoothness) or operates merely as a binary crash/survival detector.

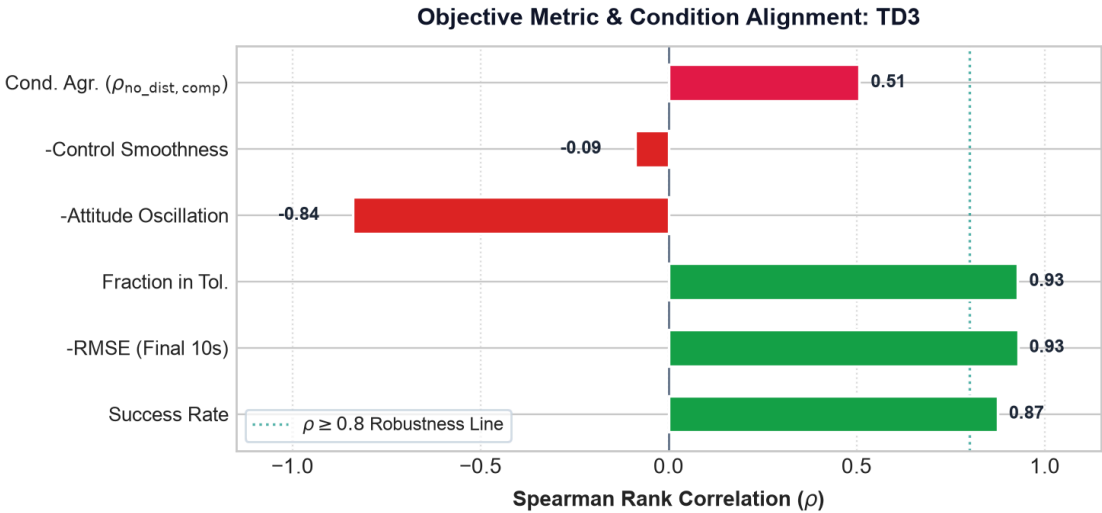


TD3 — Flight Metrics vs Return at 1,000,000 Steps

Metric Name	ρ vs Return	p -value	Interpretation
Success Rate	0.875	0.000	Airborne survival proxy
-RMSE (Final 10s)	0.929	0.000	Hover spatial precision

Fraction in Tolerance	0.928	0.000	Target cylinder containment
-Attitude Oscillation	-0.839	0.000	Angular flight stability
-Control Smoothness	-0.089	0.751	Actuator jerk / wear penalty

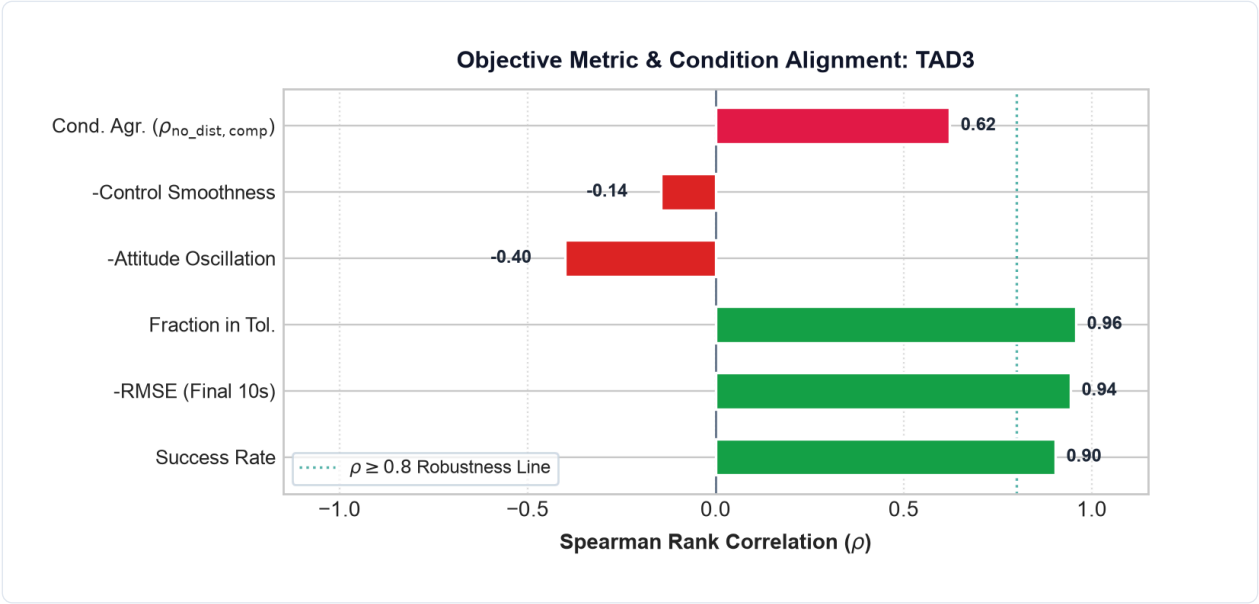
Condition Consistency: $\rho(\text{no_dist}, \text{comp}) = 0.507$ ($p = 0.054$).



TAD3 — Flight Metrics vs Return at 1,000,000 Steps

Metric Name	ρ vs Return	p -value	Interpretation
Success Rate	0.902	0.000	Airborne survival proxy
-RMSE (Final 10s)	0.943	0.000	Hover spatial precision
Fraction in Tolerance	0.957	0.000	Target cylinder containment
-Attitude Oscillation	-0.400	0.140	Angular flight stability
-Control Smoothness	-0.145	0.607	Actuator jerk / wear penalty

Condition Consistency: $\rho(\text{no_dist}, \text{comp}) = 0.621$ ($p = 0.013$).

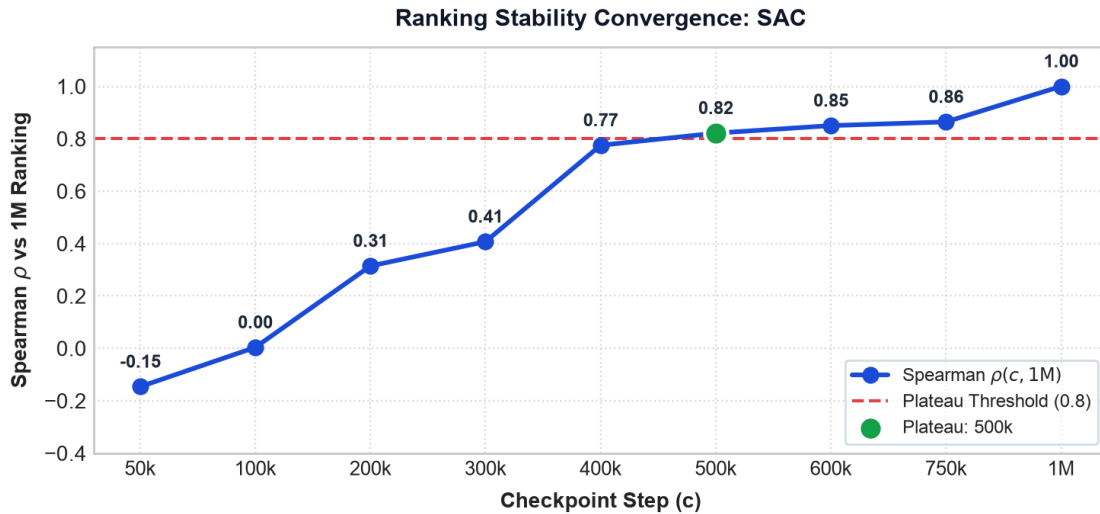


4. Analysis (b) — Timestep Budget & Ranking Convergence

Analysis (b) tracks the rank correlation $\rho(c, 1M)$ across intermediate checkpoints. The search budget is chosen as the earliest checkpoint where rankings stabilize ($\rho \geq 0.80$ sustained).

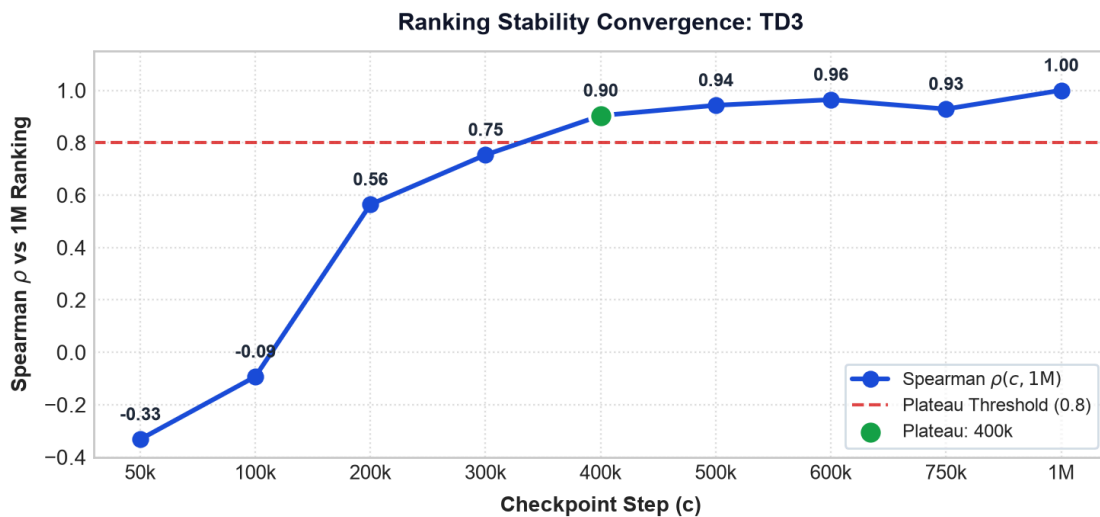
SAC — Ranking Stability Convergence

- **Plateau Candidate:** 500k (Status: ok)
- **Kneedle Inflection Candidate:** 400k



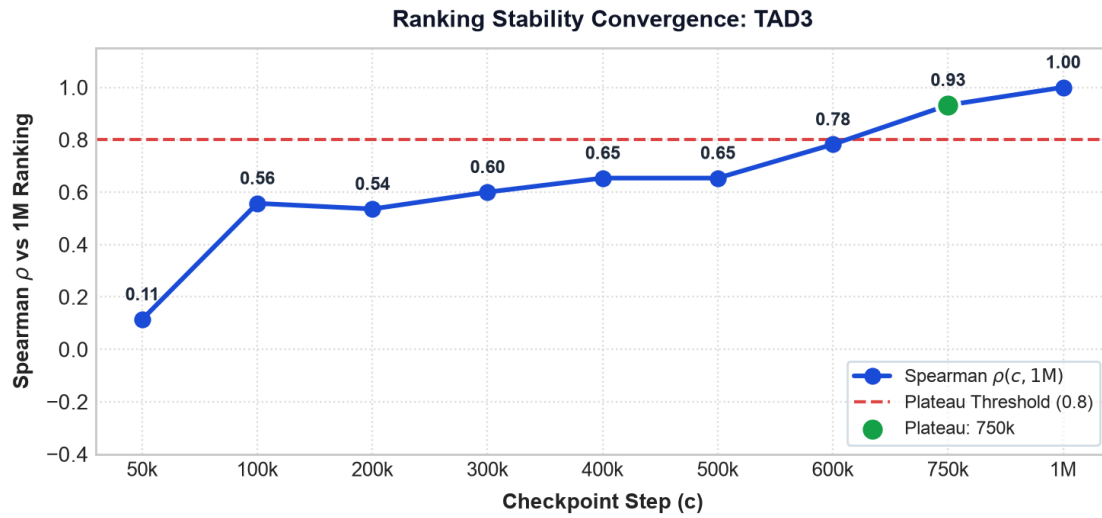
TD3 — Ranking Stability Convergence

- **Plateau Candidate:** 400k (Status: ok)
- **Kneedle Inflection Candidate:** 400k



TAD3 — Ranking Stability Convergence

- **Plateau Candidate:** 750k (Status: ok)
- **Kneedle Inflection Candidate:** 100k

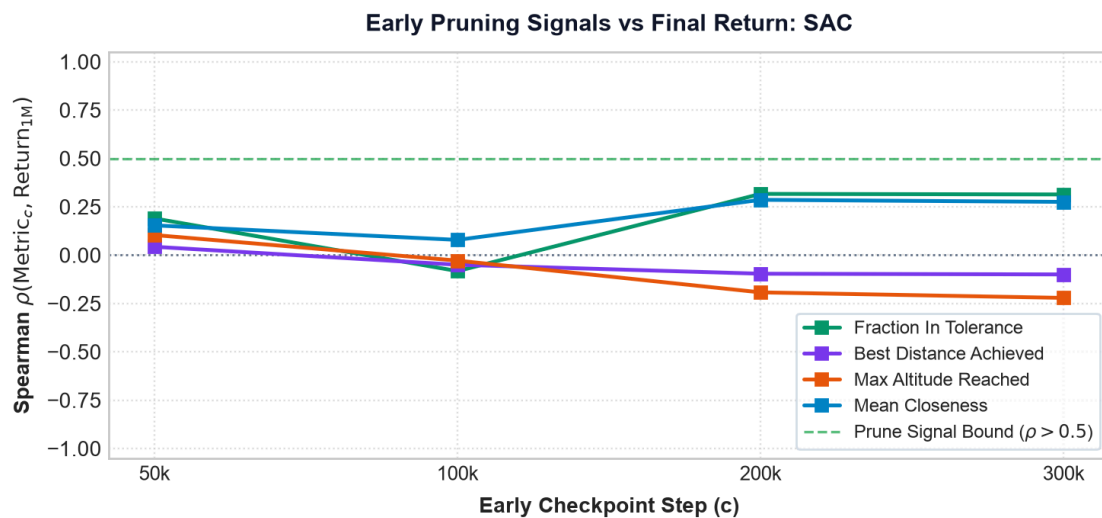


5. Analysis (c) — ASHA Pruning Feasibility & Early Signals

Analysis (c) examines whether early rungs are statistically sound ($\rho > 0, p < 0.05$), audits late bloomers in the top 5, and tests whether candidate progress metrics predict final performance earlier than reward.

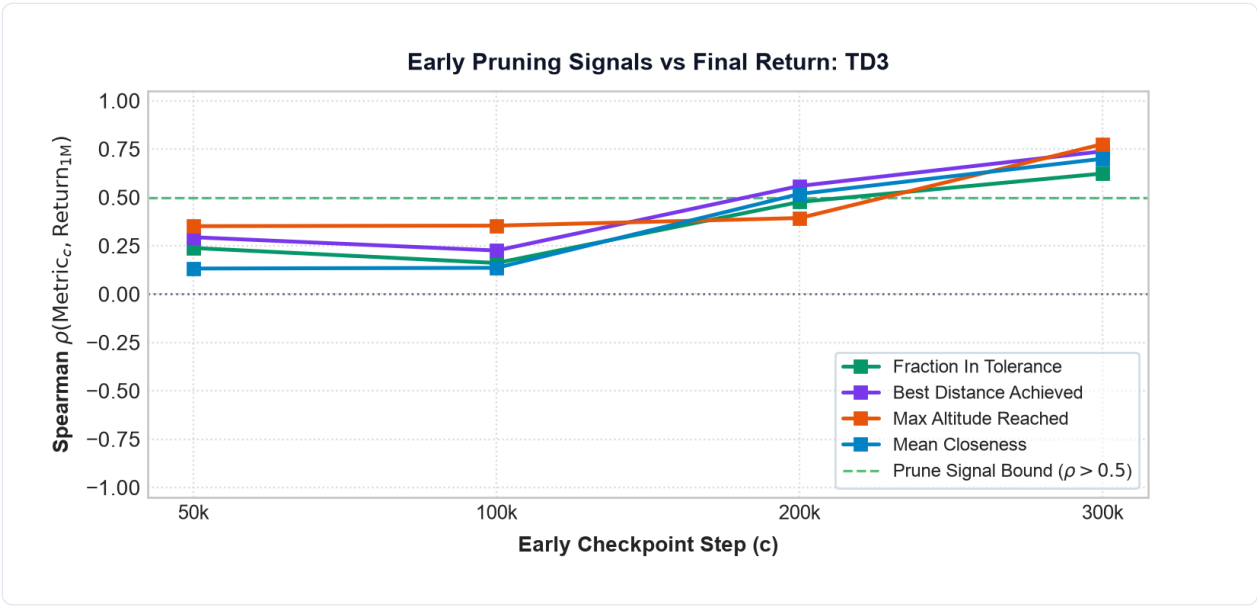
SAC — Pruning Evidence

- **Legal Rungs** ($\rho > 0, p < 0.05$): 400k, 500k, 600k, 750k
- **Late Bloomers Count:** 3 (sac_01, sac_02, sac_12)



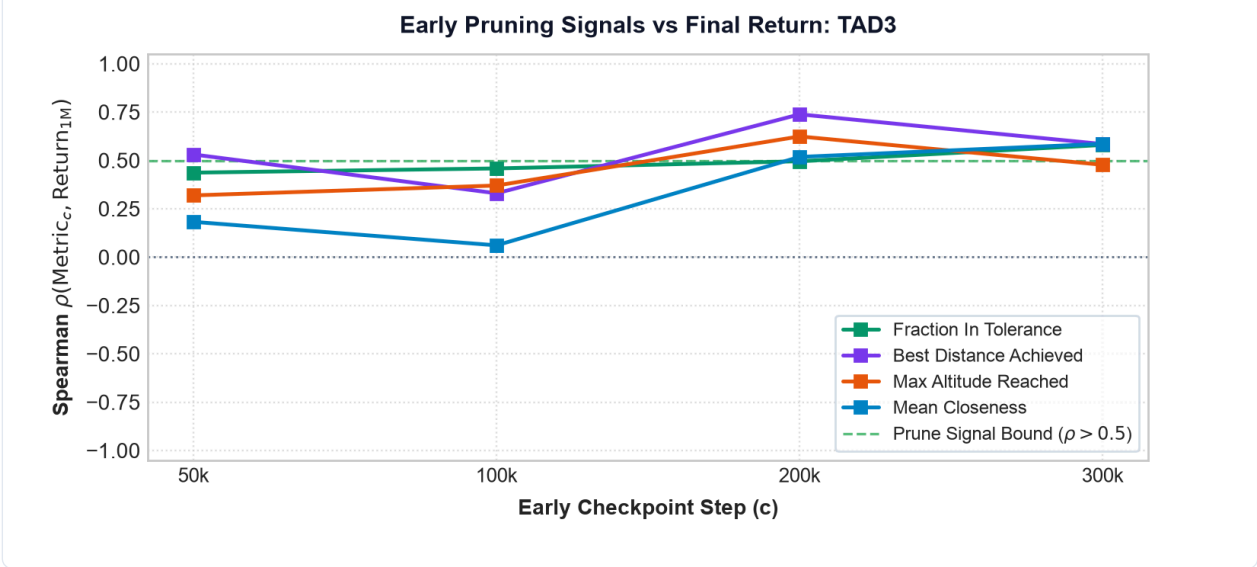
TD3 — Pruning Evidence

- **Legal Rungs** ($\rho > 0, p < 0.05$): 200k, 300k, 400k, 500k, 600k, 750k
- **Late Bloomers Count:** 3 (td3_05, td3_13, td3_14)



TAD3 — Pruning Evidence

- **Legal Rungs** ($\rho > 0, p < 0.05$): 100k, 200k, 300k, 400k, 500k, 600k, 750k
- **Late Bloomers Count**: 2 (tad3_03, tad3_12)



6. Analysis (d) — Variance Decomposition & Evaluation Protocol

Analysis (d) decomposes per-episode returns into σ^2_{config} , σ^2_{seed} , and $\sigma^2_{\text{episode}}$ to determine the standard error $\text{SE}(n)$ of a configuration mean across $n \in \{1, 2, 3\}$ episodes per cell.

SAC — Variance Components & Resolving Power

Condition	σ^2_{config}	σ^2_{seed}	$\sigma^2_{\text{episode}}$	Config Spread (SD)
No Disturbance	72906.4	89.0	779.7	270.01
Composite	34774.2	0.0	14973.7	186.48
Pooled	44230.6	0.0	7876.7	210.31

Standard Error across Episode Counts: $n = 1 \rightarrow \text{SE} = 28.07$ | $n = 2 \rightarrow \text{SE} = 19.85$ | $n = 3 \rightarrow \text{SE} = 16.20$ (vs Between-Config Spread SD = 210.31).

Variance Decomposition & Evaluation Precision: SAC

Variance Breakdown

Condition	σ^2_{config}	σ^2_{seed}	$\sigma^2_{\text{episode}}$
No Dist.	72906.4	89.0	779.7
Composite	34774.2	0.0	14973.7
Pooled	44230.6	0.0	7876.7

Resolving Power vs Sample Size

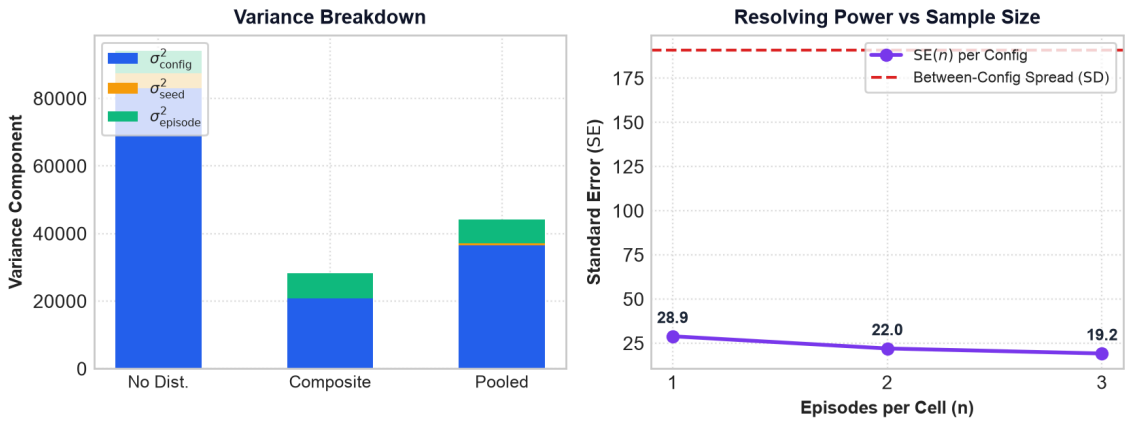
Episodes per Cell (n)	SE(n) per Config	Between-Config Spread (SD)
1	28.1	210.31
2	19.8	210.31
3	16.2	210.31

TD3 — Variance Components & Resolving Power

Condition	σ^2_{config}	σ^2_{seed}	$\sigma^2_{\text{episode}}$	Config Spread (SD)
No Disturbance	82906.9	4462.5	6652.6	287.94
Composite	20800.4	0.0	7422.2	144.22
Pooled	36419.7	664.4	7037.4	190.84

Standard Error across Episode Counts: $n = 1 \rightarrow \text{SE} = 28.92$ | $n = 2 \rightarrow \text{SE} = 22.02$ | $n = 3 \rightarrow \text{SE} = 19.17$ (vs Between-Config Spread SD = 190.84).

Variance Decomposition & Evaluation Precision: TD3



TAD3 — Variance Components & Resolving Power

Condition	σ^2_{config}	σ^2_{seed}	$\sigma^2_{\text{episode}}$	Config Spread (SD)
No Disturbance	102604.3	3.3	17.3	320.32
Composite	35208.2	0.0	6488.7	187.64
Pooled	52912.1	0.0	3253.0	230.03

Standard Error across Episode Counts: $n = 1 \rightarrow \text{SE} = 18.04$ | $n = 2 \rightarrow \text{SE} = 12.75$ | $n = 3 \rightarrow \text{SE} = 10.41$ (vs Between-Config Spread $\text{SD} = 230.03$).

Variance Decomposition & Evaluation Precision: TAD3

