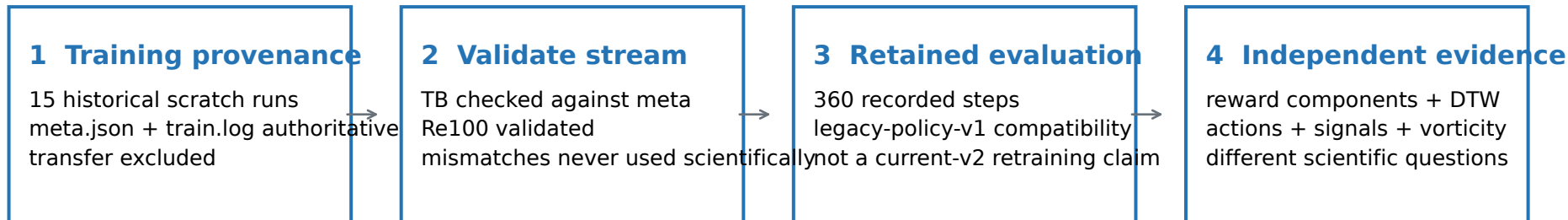


How to read retained V5 scratch/evaluation evidence



Training unit: one PPO learn chunk + deterministic evaluation; call it training iteration, not Gym episode.

Higher reward/components are better. DTW is signal similarity—not physical field error.

Target = desired reference; Controlled = policy outcome; Zero = no-rotation baseline.

Retained actions/vorticity predate the calibration-v2/native-normalization source contract.

d/L₀=0.75 and 2.0 used lr=1e-4; most runs used 3e-4. Geometry is not a pure one-factor ablation.

Only Re100 has n=5 replication; all other points are n=1 demonstrations.