

ENGINEERING TEST REPORT

SmolVLA through Endstop to SO-101

Protocol V2: model-input stress, exact counterfactual replay and bounded recovery

Purpose

Evaluate whether a deterministic Endstop action gateway can remain transparent to a public numeric-action VLA in SO-101 simulation, and whether an explicit hold, stability check, reset and fresh-observation replan can change task outcome after an input-induced, model-authored refusal.

Critical interpretation

The primary placement result begins after a scripted grasp. It is an evidence-backed placement-phase recovery result, not an end-to-end robotics-quality result or a physical safety claim. The full-task audit is reported separately and does not show a consistent improvement.

Document field	Value
Report identifier	ETR-2026-08-20-03
Revision	B - Protocol V2 public release
Test date	20 August 2026
Policy	zwan1003/pickplace_skills_vla_v3_2_ckpt060k
Confirmatory corpus	114 episodes; 4 independent blocks
Environment	SO-101 MuJoCo; Apple M1 Pro MPS
Prepared by	Alnair LLC / Endstop engineering

Approval status: experimental software-in-the-loop evidence. Not certification, hardware validation, a stopping-performance demonstration, or a claim that the physical product prevents damage.

1. Design and precommitment

The V2 protocol deliberately separated two tracks. The model-input track modifies vision, state or language before SmoIVLA inference, so every resulting command is authored by the policy. The command-path track changes an already-produced command and measures gateway coverage only. These tracks are never pooled.

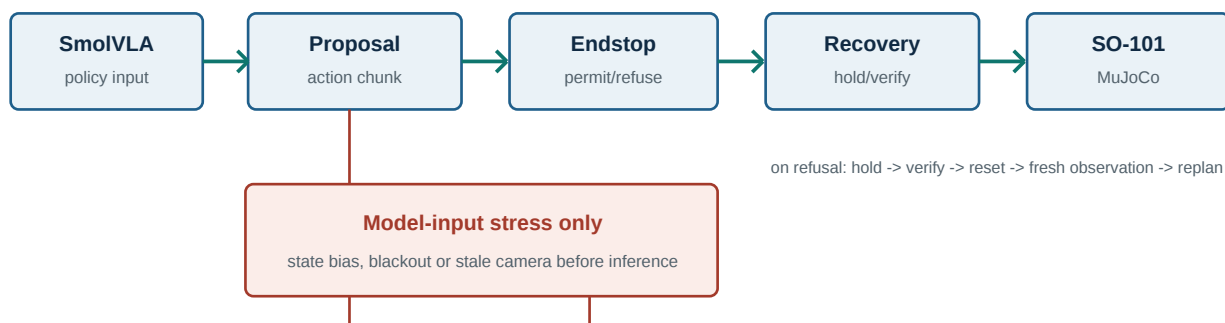


Figure 1. In the primary track, the state bias reaches the policy input before inference. The action is not edited after inference. The distinct command-path spike remains a positive control.

1.1 Frozen confirmatory matrix

Block	Cases	Purpose
Placement primary	40	Nominal and state-bias; shadow/recover; five fresh seeds
Placement visual stress	20	Camera blackout and stale camera; descriptive
Command-path control	30	Post-policy rate spike; shadow/halt/recover
Full-task audit	24	No scripted grasp; nominal and state bias

The initial calibration data were excluded. A first confirmatory replay failed its frozen identity criterion because MPS/camera replays diverged at the disturbed inference; it remains published but excluded. V2.1 instead shares one realized shadow policy path, replaying byte-identical proposals into counterfactual arms until refusal. V2.2 bounded recovery at three replans after a partial full-task run exposed an unbounded loop. Both amendments were frozen before their fresh-seed datasets.

2. Primary placement result

Placement successes (n=10 per cell)

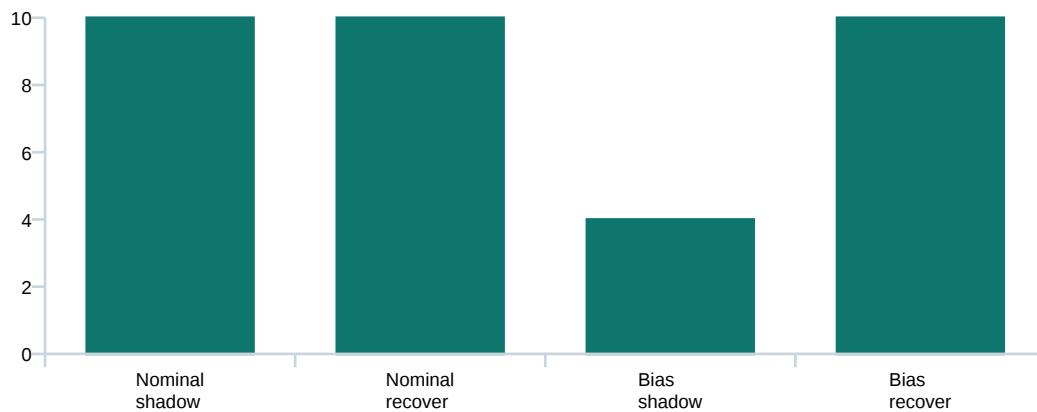


Figure 2. Exact-prefix state-bias recovery completed 10/10 placement episodes versus 4/10 shadow episodes. Nominal behavior was 10/10 in both arms.

Placement cell	Success	Wilson 95%	Intervened	Prevented
Nominal / shadow	10/10	0.722-1.000	0/10	0
Nominal / recover	10/10	0.722-1.000	0/10	0
State bias / shadow	4/10	0.168-0.687	10/10	0
State bias / recover	10/10	0.722-1.000	10/10	20

The bias added an alternating ± 0.35 rad offset only to the six-dimensional policy state during chunks 2 and 3. The simulator state was untouched. In every state-bias episode, the first refusal was at tick 50 for `axis_rate`. Six paired cases were recover PASS / shadow non-PASS; there were no reversals. The matched success difference was $+0.60$, with two-sided exact McNemar $p=0.03125$.

All 20 primary recovery cycles completed hold, stability verification, reset authorization, monitor reprime, queue discard and fresh-observation replan. The greatest final-four-tick measured rate was 0.000402 rad/tick against a 0.02 rad/tick authorization limit.

3. Recovery and the independent command-path control

A refusal does not immediately clear the latch. The recovery arm holds the last permitted target for 12 ticks; requires all final four per-tick joint changes to stay at or below 0.02 rad; only then executes the simulated supervisor reset; re-primes the monitor; discards the remaining policy queue; refreshes observation; and replans. The fourth refusal halts the episode under the V2.2 retry budget.

Command-path control	Detection	Task result	Interpretation
Shadow	10/10 first refusals	10/10 PASS	Spike executed after being logged
Halt	10/10 first refusals	0/10 PASS	Spike prevented; task intentionally ended
Recover	10/10 first refusals	10/10 PASS	Spike prevented; one replan per case

The rate spike was inserted after policy output at tick 60. It was detected as `axis_rate` in every arm. This is a positive control for detection and recovery coverage; it does not estimate natural VLA failure or policy quality. Shadow recorded a second refusal after executing the spike because the return command also exceeded command-to-command rate.

The descriptive visual probes were less eventful: camera blackout and stale-camera placement both completed 10/10 with no intervention. That supports tolerance to this short late-placement perturbation only; it does not establish camera-loss robustness during grasp or on hardware.

4. Full-task audit and limits

The full-task stratum begins with the cube on the table and asks SmoIVLA to pick and place. It removes the scripted grasp that makes the primary placement measure repeatable. Outcomes were weak and mixed; the base policy's grasp behavior dominates.

Full-task cell	PASS	Other terminal outcomes	Intervened
Nominal / shadow	1/6	2 TOUCHED; 3 FAIL	3/6
Nominal / recover	0/6	3 TOUCHED; 3 FAIL	3/6
State bias / shadow	0/6	6 FAIL	6/6
State bias / recover	1/6	4 TOUCHED; 1 FAIL	6/6

Nominal recovery minus shadow was -1/6; state-bias recovery minus shadow was +1/6. Each comparison had one discordant pair and exact $p=1.0$. Two recovery episodes exhausted the three-replan budget and halted. The primary placement effect must therefore not be described as a general end-to-end VLA improvement.

Claim boundary

Statement	Status
Endstop can sit in this public numeric-action VLA path	Supported in simulation
Sampled nominal placement stays unchanged	Supported for this subset
Input-induced model-authored rate commands can be refused	Supported for this scenario
Bounded recovery improves the full task	Not established
SmoIVLA naturally produces unsafe commands	Not measured
Physical motors, stopping or damage avoidance	Not tested

5. Timing, kinematics and reproducibility

Measure	Result
Policy live inference	1,416 calls; median 755.75 ms; p95 1,229.18 ms; max 3,628.70 ms
Gateway round trip	43,105 calls; median 0.102 ms; p95 0.222 ms; p99 0.575 ms; max 49.717 ms
Endstop FK vs MuJoCo	5.72 micrometres maximum; 64 poses and 192 tap comparisons per valid block
Raw corpus	114 episodes; 42,769 policy ticks; 43,328 simulator ticks

These are software timings on an Apple M1 Pro/MPS host. They exclude camera capture, a motion bus, a real-time scheduler, motor drive response, feedback independent of the host, braking and physical stop behavior.

The public bundle contains the V2 runner, analyzer, protocol, preregistration, amendments, raw traces, monitor configurations, FK checks, block analyses, checksums and the combined result index. It excludes downloaded weights, environments, build products and the invalidated datasets from the headline corpus. All four final block manifests verified with zero mismatches.

References and controlled release

- SmolVLA documentation: <https://huggingface.co/docs/lerobot/smolvla>
- Task-specific checkpoint: https://huggingface.co/zwan1003/pickplace_skills_vla_v3_2_ckpt060k
- SO-101 task: <https://github.com/zwaneiz/so101-vla-pickplace> at recorded commit 8df25f827c2141b22a762c4a8ecc33fbbe6a48ca.
- Protocol, data and source: Endstop SmolVLA Protocol V2 artifact bundle, 20 August 2026.

Approved concise statement

In a preregistered, software-only placement study, a temporary proprioceptive bias caused Endstop refusals of model-authored SmolVLA actions in all ten matched cases. Exact-prefix bounded recovery completed 10/10 placements versus 4/10 in shadow execution. The full-task audit without scripted grasp did not show a consistent benefit, and no hardware-safety claim is made.

End of controlled report

ETR-2026-08-20-03 | Revision B | Public release