Whether proprioception hurts a manipulation policy depends on what the policy can see

Ada Mao (Independent) · Claude Code (AI agent, Anthropic)

Preprint · OpenFindings or:2026:39f390f2 · Code and data: github.com/OpenFindings/scientific-discovery-robotics (branch claude/friendly-archimedes-4ft6w0) · CC-BY-4.0

Version 2, 2026-10-26. Adds a state-dropout arm (D) to the pixel experiment (§4.6); version 1 was posted 2026-10-25. A literature meta-analysis plus a CPU simulation study in LIBERO, with testable predictions for large vision-language-action (VLA) models. No GPU was used: every simulated policy is a small network trained and evaluated on CPU. The predictions in §6 are written so that a group with GPUs can confirm or refute them. Code, per-episode results and analysis scripts: `pilot/sim/`, `evidence_map/`.

Abstract

Papers disagree about whether a robot policy should be given its own joint and gripper readings, its *proprioceptive state*.

- Removing state has made policies tens of points more robust to position shifts.
- Adding state has rescued policies by up to +69 points when a camera fails.

Literature. We coded 123 with-vs-without-state settings from 14 papers, found through a two-hop citation walk.

- No single variable predicts the sign of the effect at our pre-set bar.
- The type of evaluation shift comes closest:
  - raw state hurt in 16 of 16 configuration-shift settings;
  - state helped in 11 of 14 camera-fault settings;
  - in-distribution effects were small and of mixed sign.

Simulation. In LIBERO we trained policies with no state (N), raw absolute state (R) or episode-relative state (C), at four perception levels. With pixels we also trained raw state with 30% state dropout (D). The perception levels were:

- no perception at all;
- perfect object positions in world coordinates;
- perfect object positions relative to the gripper;
- real 64×64 camera images with a small CNN.

They were evaluated on paired instances, 3 seeds each.

Five findings.

1. **In distribution, LIBERO hardly tests perception.** A policy that never sees an image solves 74% of in-distribution LIBERO instances from raw state and a task id.
2. **The "configuration shift" of the literature splits in two.** With idealised perception:

    - state hurts when the robot starts from a shifted joint configuration (−21 pp);
    - state helps when the object is moved 6 cm (+7 to +10 pp).
3. **The robot-start harm disappears with real images.** With pixels, raw state is neutral under the robot-start shift (+4 pp, CI includes 0), and relative state *helps* (+12 pp vs none, +8 pp vs raw).
    - **Why:** a no-state pixel policy sees the arm in the scene camera. It is therefore exactly as sensitive to the arm's start as a state-using policy (−25 vs −23 pp from in-distribution).
    - State adds harm only when it is the policy's *only* source of arm configuration.
    - An offline input-ablation probe agrees. The idealised policies rely on state about as much as on object information. The pixel policies rely on state about half as much as on either camera.
4. **Whether the relative frame helps depends on perception, and it can even flip sign.** Measured against raw state under the robot-start shift, relative state is:
    - +18 pp blind;
    - −34 pp with world-frame objects;
    - +3 pp with gripper-frame objects;
    - +8 pp with pixels.
5. **State dropout does not help under shift; the relative frame does.** With pixels, raw state trained with 30% state dropout (D) behaves like raw state:
    - −0.7 pp vs raw under the robot-start shift;
    - −3.8 pp under the layout shift.

    It is 9 pp below relative state under the robot-start shift, and 7.6 pp below under the layout shift. The probe shows D has learned to nearly ignore state.

We turn these into eight predictions for π0.5-class models on LIBERO-Plus. The central claim is that the sign of a state ablation is set by **what else the policy can see about the robot**, not by state alone.

---

1. The contradiction

Recent work disagrees about proprioception:

- **State-free policies are more robust.** π0, ACT and Diffusion Policy without state lose far less under height, horizontal and embodiment shifts: 37–98 pp less (2509.18644).
- **State rescues camera faults.** π0.5 and GR00T with state keep 11–69 pp more success when a camera is blacked out or the robot is masked in the image (Blackout vs Freeze, 2609.39145).
- **In distribution it barely matters.** Most papers report differences within a few points.

The standard explanation is a **shortcut**. During training, state predicts the next action so well that the policy leans on it and attends less to the images (de Haan et al. 2019; Wen et al. 2020). That explains the harms but not the rescues.

Question. Which variable decides the sign of a state ablation?

2. Meta-analysis

2.1 Retrieval and coding

- **Seeds.** The three conflicting papers (2509.18644, 2608.03052, 2609.02546) and the two shortcut papers (1905.11979, 2010.14876).
- **Graph walk.** Both directions of the citation graph, plus a second hop from five first-hop papers that run state ablations. Semantic Scholar returned 912 papers.
- **Screen.** A high-recall title/abstract screen flagged 82. With 7 classics added, all 89 were read:
  - 49 excluded (driving, locomotion, Atari, datasets, no state arm);
  - 2 duplicates;
  - 3 unavailable.
- **Coding.** 14 papers give 123 numeric with-vs-without-state settings. One row per setting, with source table or figure: `evidence_map/data/results.jsonl`.
- **Variables coded.** Shift type, state frame, fusion, velocity terms, history length, action space, wrist camera, platform, backbone, pretraining and masking protocol.

2.2 Pre-set test

Each record is classed as:

- **hurt:** below −3 pp;
- **null:** within ±3 pp;
- **help:** above +3 pp.

Each paper carries a total weight of 1. A variable *explains* the contradiction if both hold:

- its levels predict the weighted-majority class with paper-weighted accuracy of at least 0.90;
- no level has counterexamples from two or more papers.

No variable passes. The best single variable is the type of evaluation shift (0.78). The best pair, shift × state frame, reaches 0.82. The failing level is the in-distribution one; every shifted cell is nearly clean:

Cell	Settings (papers)	Outcome
Configuration shift (object, base, embodiment or start joints out of range), raw state	16 (6)	**hurt 16/16**
Configuration shift, corrected frame or regularised fusion	12 (3)	help 8, null 1, hurt 3 (the 3 hurts have broken images)
Camera fault, robot in distribution	14 (1)	**help 11/14**, hurt 0
In distribution, raw state	60 (14)	help 15, hurt 14, null 31; mostly within ±3 pp

2.3 Planner

Seven mutually exclusive hypotheses were scored with a Bayesian evidence planner:

Hypothesis	Claim	Posterior
H1	Any shift: state hurts	0.05
H6	Configuration shift: state hurts. Camera fault: state helps.	0.28
H7	As H6, but a corrected encoding removes the harm under configuration shift	0.67

Leaving out any one paper keeps H7 at 0.55–0.68. H6 and H7 were written after reading the papers, so the posterior ranks explanations and does not confirm them.

The most informative missing cells are *configuration shift × corrected encoding*. No paper tests one policy under both a configuration shift and a camera fault.

2.4 What the literature cannot separate

"Configuration shift" pools two different things:

- the **robot** starting from a new joint configuration;
- the **objects** being somewhere new.

The published harms under object shifts come from policies whose visual grounding is unknown. The published rows also never report how the *no-state* arm itself fared relative to its in-distribution score, beyond its success under shift. §4 shows that this missing number decides the sign.

3. Simulation setup

Simulator. LIBERO (MuJoCo 3.2.3, robosuite 1.4.1), CPU only.

Shifts. Every policy runs the same paired instances: 10 tasks × 5 init states per suite.

Shift	Definition
robot_init:0.1	The 7 arm joints get an L2 offset of 0.1 rad in a random direction, as in LIBERO-Plus's robot-initial-state generator. Median end-effector displacement 35 mm. Raw states stay inside the training 1–99% range.
layout:0.06	The first movable object of interest moves 6 cm horizontally. The placement is validated after settling: ≤10 mm xy and z deviation, ≤8° tilt, nothing else displaced, task not pre-solved. Then the same validated state is re-entered through a fresh reset. Invalid instances are skipped for all policies.
scene_blackout	The agent-view image is all zeros from the first step (pixel policies only).

State arms.

- **N:** none.
- **R:** raw `[eef_pos, eef_axis_angle, gripper_qpos]` in the world frame.
- **C:** episode-relative, $\text{pos} - \text{pos}_0$ and $\text{axis_angle}(R_0^\top R)$, gripper absolute.
- **D** (pixels only): as R, but on 30% of training samples the state input is replaced by its training mean (state dropout, the regularised encoding of the pilot design). At evaluation the state is always given.

Policy and training. Every policy gets the task as a one-hot input and predicts a 10-step action chunk, of which 5 steps are executed. Demonstrations are no-op filtered.

Perception levels.

Level	Perception input	Network	Training data	Suites	Seeds
Blind	none. The N arm is open-loop replay of the time-indexed policy (T).	MLP	all 1,693 openpi LIBERO demos	all 4 (200 instances)	3
Allocentric oracle	world positions of the two objects of interest	MLP	original LIBERO demos, object poses rebuilt by replaying the saved simulator states	Spatial + Object (100)	1
Ego oracle	object minus gripper position, a perfect wrist camera	MLP	same	Spatial + Object (100)	3
Pixels	64×64 agent-view + wrist RGB from the original LIBERO demos	two 4-layer CNNs + MLP head, 1.7M params, random-shift augmentation, 20k steps	same	Spatial + Object (100)	3

Pixel alignment. At evaluation, images are rendered at 128×128 and resized exactly as in training. Rendered frames match the dataset frames: mean absolute error 3–13 out of 255 when demo states are replayed (`results/pix/render_alignment_spatial_t0.png`).

Statistics. Differences are paired on identical instances:

- the effect is the mean over seeds;
- the 95% CI is a task-cluster bootstrap with each instance carrying all its seeds;
- the seed spread is the per-seed effects and their SD.

4. Results

All numbers are success rates in %, as the mean over 3 seeds, with per-seed values in brackets unless noted.

4.1 In distribution, LIBERO barely tests perception

Blind policy	ID	robot_init 0.1	layout 6 cm
T (time only; open-loop replay)	52.8 (52/53/53)	31.0	4.1
A (raw state)	**73.8** (72/75/75)	14.5	4.1
Rl (relative state)	63.2 (66/62/62)	32.3	4.9

- A policy that never sees an image solves 74% of in-distribution instances across the four suites.
- Moving the object 6 cm takes every blind policy to about 4%. So 6 cm is the smallest layout shift we tested that requires perception, and 3 cm does not (30–44% blind, single seed).

This agrees with LIBERO-PRO (2510.03827), which finds trajectories unchanged when the object is removed, and with MINERVA (2609.03715), where a task-id permutation drops a 0.54M CNN to chance. It sharpens them to a number with no images at all (`research/novelty_check.md`).

4.2 The effect of state, by perception level and shift

Paired differences in pp, with 95% CI and per-seed values:

Perception	Shift	R − N	C − N	C − R
Blind (N = T)	ID	+21.0 [+10.8, +32.0]	+10.3 [+2.5, +18.7]	−10.7 [−16.8, −5.2]
	robot_init 0.1	**−16.5** [−23.3, −9.7] (−18.5/−12.5/−18.5)	+1.3 [−3.5, +6.3]	**+17.8** [+12.0, +24.2]
	layout 6 cm	0.0 [−4.0, +3.7]	+0.7	+0.7
Allocentric oracle (1 seed; N cannot act, 5% ID)	robot_init 0.1	—	—	**−34** [−45, −23]
	layout 6 cm	—	—	**−25.0** [−40.7, −8.5]
Ego oracle	ID	−4.7 [−10.3, +1.0]	−7.7 [−14.0, −1.7]	−3.0 [−11.0, +4.7]
	robot_init 0.1	**−21.3** [−31.7, −10.3] (−20/−15/−29)	**−18.7** [−27.3, −9.3] (−22/−14/−20)	+2.7 [−5.0, +10.3]
	layout 6 cm	**+6.8** [+0.7, +13.2] (+11/0/+9)	**+9.5** [+3.2, +15.5] (+18/+2/+8)	+2.7 [−3.9, +9.8]
Pixels	ID	+2.0 [−2.7, +7.3]	+6.3 [+1.3, +12.3]	+4.3 [−0.7, +10.0]
	robot_init 0.1	+4.0 [−0.7, +8.7] (−2/+1/+13)	**+12.3** [+4.7, +20.0] (+7/+13/+17)	**+8.3** [+2.3, +13.7] (+9/+12/+4)
	layout 6 cm	+3.8 [−4.3, +11.6] (+3/+6/+2)	+7.6 [−1.1, +16.5] (+5/+11/+7)	+3.8 [−1.8, +9.4]
	scene blackout	0 (all arms 0%)	0	0

Pixel success rates (N / R / C / D):

Condition	N	R	C	D
ID	88.3 (89/91/85)	90.3 (89/91/91)	94.7 (92/96/96)	93.0 (92/93/94)
robot_init 0.1	63.0 (66/65/58)	67.0 (64/66/71)	75.3 (73/78/75)	66.3 (65/65/69)
layout 6 cm	35.2	39.0	42.8	35.2
scene blackout	0	0	0	0

Seed spread. The per-seed SD of a paired effect was 1–8 pp across all experiments. Every effect whose CI excludes zero has the same sign in all three seeds, except two:

- ego layout R − N (+11/0/+9);
- pixel robot-init R − N (−2/+1/+13), whose CI includes zero.

4.3 Why the robot-start harm disappears with pixels

Drop from in-distribution to robot_init 0.1 (pp):

Arm	Ego oracle	Pixels
N	**13.7**	**25.3**
R	30.3	23.3
C	24.7	19.3

- **The state arms degrade about equally** at both perception levels (23–30 pp).
- **What changes is the no-state arm.**
  - The ego-oracle N arm sees only gripper-relative object offsets. Those carry no information about the arm's joint configuration, so the shifted start barely affects it.
  - The pixel N arm sees the whole arm in the agent-view image. It is exactly as start-sensitive as the arms given state.
- **So "state hurts under robot-start shift" measures the no-state arm's invariance, not a harm of state per se.**

Input-ablation probe. On 2,000 training frames, we measured how much the normalised action chunk changes (RMS) when one input is replaced by its training mean (state) or by zeros (images). Values are relative to the action standard deviation; files `results/pix/reliance.json` and `results/seeds/ego_reliance.json`.

Policy (3 seeds)	state	object or agent-view	wrist
Ego oracle R	1.24–1.36	1.42–1.49	—
Ego oracle C	1.39–1.46	1.51–1.64	—
Pixels R	0.38–0.41	0.70–0.81	0.84–0.91
Pixels C	0.45–0.47	0.64–0.84	0.81–0.94
Pixels D	0.07	0.73–0.91	0.95–1.08

- The idealised policies weight state about as much as their perception.
- The pixel policies weight state about half as much as either camera.
- With pixels, adding state does not change the imitation loss: final training MSE was 0.200 / 0.199 / 0.197 for N / R / C. With the ego oracle, state cut the loss 2.5×.

State becomes redundant once the images already show the arm, and redundant state is not a strong shortcut.

4.4 The relative frame: help, harm or nothing, depending on perception

C − R under robot_init 0.1, by perception level:

Perception	C − R (pp)
Blind	+17.8
Allocentric	−34
Ego	+2.7
Pixels	+8.3

- **Blind.** The relative state starts at zero in every episode, so replaying from it avoids extrapolating an absolute state the policy never saw.
- **Allocentric.** World-frame objects need the absolute arm position to compute a reach. The relative frame throws that position away.
- **Ego.** Gripper-relative objects already carry the reach. The frame adds nothing.
- **Pixels.** The frame helps, which is the H7 pattern.

So "does a corrected frame fix state?" has no perception-free answer.

4.5 State dropout (D)

Paired differences for the pixel policies, 3 seeds:

Shift	D − N	D − R	D − C
ID	+4.7 [+0.7, +9.0] (+3/+2/+9)	+2.7 [0.0, +6.0] (+3/+2/+3)	−1.7 [−5.7, +2.3]
robot_init 0.1	+3.3 [−0.7, +7.3] (−1/0/+11)	−0.7 [−5.0, +3.7] (+1/−1/−2)	**−9.0** [−15.0, −3.3] (−8/−13/−6)
layout 6 cm	0.0 [−9.7, +8.5]	−3.8 [−8.9, +1.1] (−5/−1/−6)	**−7.6** [−13.0, −1.1] (−6/−7/−10)

- **Under both shifts, D is indistinguishable from raw state**, and it is worse than relative state in all three seeds.
- **In distribution, dropout adds a small, consistent gain over raw state** (+2.7 pp; +2 to +3 in every seed).
- **Reliance probe.** Replacing D's state with its training mean changes its actions by only 0.07 action-SD, about a fifth of R's 0.38–0.41. Dropout taught the policy to treat state as optional, and it then mostly ignored it.
    - Caveat: the probe's replacement value is exactly the dropout value D saw in training. The probe therefore measures how far D's actions depend on the state when it is present, not how it reacts to an unusual state.
- **The question dropout was meant to answer remains open.** Does reducing reliance on state give up state's camera-fault benefit (the pilot's trade-off T)? It cannot be answered here, because every pixel arm, D included, scores 0% under scene blackout (§4.6).

4.6 Camera faults: not reproduced

With the scene camera blacked out, all pixel policies scored 0%. The oracle analogue also scored 0%. Our small CNNs cannot do these tasks from the wrist camera alone, so this study says nothing about state's camera-fault benefit. That half of the evidence still rests on one paper (Blackout vs Freeze).

5. Interpretation

1. **The sign of a state ablation is a comparison between two policies.** It depends as much on what the no-state policy can perceive as on what state adds.

    - Under a robot-start shift, state hurts exactly when the no-state policy's perception is invariant to the arm's configuration: blind replay or gripper-frame objects.
    - When the cameras show the arm, which is the standard third-person LIBERO setup, both policies inherit the same sensitivity, and raw state is neutral.
2. **Object shifts behave differently from robot shifts.** With perception good enough to locate the object, state did not hurt at 6 cm; in the ego oracle it helped. A state *harm* under object shift in a real VLA would therefore indicate **vision suppression**: imperfect visual grounding that state lets the policy skip.
3. **Regularising state away is not the same as correcting its frame.**
    - With pixels, dropout made the policy largely ignore state and left its robustness unchanged (D ≈ R ≈ N under both shifts).
    - Re-expressing state relative to the episode start kept state in use and improved robustness to the robot-start shift (C − R +8 pp).
    - So when state does help, it helps through a frame that stays in distribution, not through less reliance.
4. **For the planner's H6 vs H7.** The pixel results favour an encoding effect under robot-start shift: C beats both R and N. But they do not reproduce the premise that raw state hurts under configuration shift. The pooled "configuration shift" cell should be split, and H6/H7 re-scored per sub-type, before any decision.

6. Testable predictions for large VLAs

Setup. One backbone (π0.5 from `pi05_base`), fine-tuned on LIBERO three times with identical data, cameras and recipe:

- **N:** stock `pi05_libero`, `discrete_state_input=False`;
- **R:** raw state as discretised prompt tokens;
- **C:** episode-relative state, renormalised;
- **D:** raw state with 30% state-token dropout during training (the pilot's planned extension arm).

The protocol, configs and scripts are in `research/pilot_3x2_liberoplus.md` and `pilot/stage0/`. Instances are paired, and the shifts are those of §3, at LIBERO-Plus magnitudes.

#	Prediction	Refuted if
P1	**In distribution:** $\|R - N\|$ and $\|C - N\| \leq 5$ pp.	Any ID state effect above 10 pp.
P2	**Robot start 0.1:** $R - N \approx 0$, within ±8 pp, because π0.5's scene camera sees the arm (§4.3).	$R - N \leq -10$ pp with a CI excluding 0. That would mean the harm reported for π0-class models is caused by state itself, not by the no-state arm's invariance.
P3	**Robot start 0.1:** $C - R \geq 0$, expected +5 to +15 pp.	$C - R \leq -5$ pp.
P4	**Robot start, mechanism:** the N arm's own drop from ID is at least as large as R's (pixel pattern), not about half (ego pattern).	N drops less than half as much as R.
P5	**Layout 6 cm:** $R - N \geq 0$, and N, R and C are all well above the blind floor of ~4%.	$R - N \leq -10$ pp. That would directly measure **vision suppression** by state in π0.5.
P6	**Reliance:** in the input-ablation probe (§4.3) on π0.5, removing state changes actions less than removing the agent-view image. Policies where state outweighs perception should be the ones that show P2 harm.	State reliance ≥ agent-view reliance while P2 holds; or the reverse pattern.
P7	**Wrist camera as the only view of the arm:** with the agent view removed or cropped to exclude the arm, the robot-start harm of state reappears (the ego pattern, $R - N < 0$).	$R - N$ stays ≈ 0 when the arm is invisible to the no-state policy.
P8	**State dropout:** under the robot-start and layout shifts, $D - R \approx 0$ (within ±5 pp) and $C - D > 0$. In distribution $D \geq R$.	D beats C under the robot-start shift, which would mean regularisation rather than the frame is what removes state harm at VLA scale.

P2, P5 and P7 are the decisive ones. P7 is a controlled version of our explanation, and it can be run with the same checkpoints by changing only the camera setup at fine-tuning.

What would most change our conclusions. If P2 fails, state harms the policy in a way our small CNNs do not capture, such as a stronger state shortcut at VLA scale or with 256-bin state tokens. That would make the pilot's §3a decision table apply unchanged.

7. Limitations

- **Small policies.** At most 1.7M parameters, trained from scratch on 2 of 4 suites (Spatial, Object), at 64×64. π0.5 has a pretrained 2B vision-language backbone, state as discrete tokens, and different failure modes. The simulation bounds mechanisms; it does not predict π0.5's numbers.
- **One robot-start magnitude** with pixels (0.1). The ego oracle also covered 0.2 and showed the same sign, with 1 seed.
- **Camera faults** could not be tested (§4.6). As a result, the trade-off that state dropout was meant to probe (robustness bought by giving up the camera-fault benefit) is untested.
- **One dropout rate** (30%) was tried, and dropout was applied to the raw-state arm only.
- **The allocentric oracle** has 1 seed.
- **Blind seed-0 layout rows** were re-run after a reset fix. The seed-0 ID and robot-start rows predate it but are not affected by it.

- **Meta-analysis coding** was done by one reader (Claude) without a second human coder. H6 and H7 are post hoc.
- **Novelty.** We searched arXiv up to mid-October 2026 (`research/novelty_check.md`). The blind-LIBERO result extends known memorisation findings. We found no prior report of:
    - the robot-start versus object-shift reversal;
    - the perception-dependence of the frame effect.

Reproducing

What	Where
Meta-analysis	`evidence_map/scripts/` (records, single-variable test, planner)
Simulation	`pilot/sim/`, launch via `job.sh` / `run_queue2.sh` / `run_queue3.sh` (arm D) (CPU only, ~1 day on 4 cores)
Pixel pipeline	`pix_extract.py` → `pix_align.py` → `pix_train.py` → `pix_eval.py`
Analysis	`analyze_seeds.py`; probes `pix_reliance.py`, `ego_reliance.py`
Per-episode results	`pilot/sim/results/{blind,ego,oracle,pix,seeds}/`