ROBOT MEMORY IN CONTEXT: SEPARATING STORED CONTENT FROM CONDITIONS OF USE

Anonymous authors
Paper under double-blind review

ABSTRACT

After a cue disappears, a robot must still choose the object it indicated. We ask when a stored representation that can guide this choice remains useful as the scene and control phase change. In a frozen recurrent robot policy, we replace saved states and edit cue components estimated from other source scenes, holding the current test scene fixed within each comparison. On independently sampled scenes, replacing one cue component with another redirects 53/64 choices toward the specified new object, versus 0/64 for token-magnitude-matched random changes: an 82.8-percentage-point gain in targeted redirection. The same edited tensors yield 11/64 matches when inserted two actions later, with the same remaining action budget. What the robot did in between matters: at that later query, an unchanged saved state guides 14/64 choices after the ordinary action prefix and 49/64 when the two preceding arm motions are paused. On jointly new source and test scenes, calibration reduces Brier probability error from .1717 to .0583, a 66.0% reduction against a same-budget rate baseline. A memory can hold the right content yet fail to guide action, so robot memory should be evaluated both for what stored content can direct and for when the controller can use it.

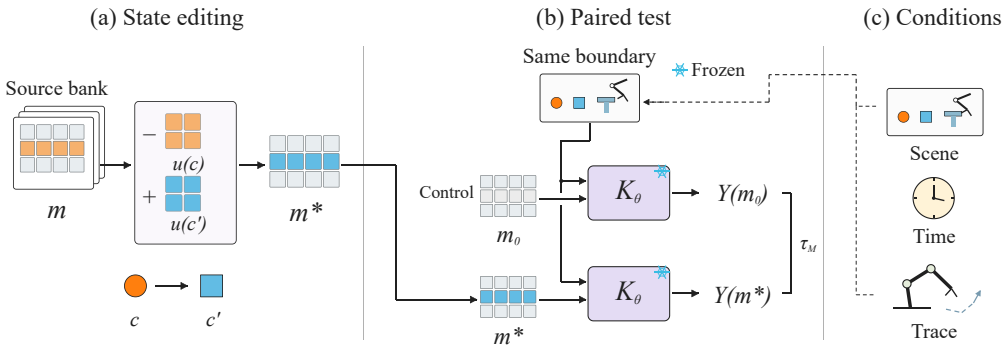

Figure 1: Memory content and conditions of use. (a) Cue conversion replaces $u(c)$ with $u(c')$ in a saved state: $m^* = m - u(c) + u(c')$. Here $u(c)$ denotes the source-excluded component $u_{-s,c}$ (Equation 2); colored blocks depict schematic state components. (b) The no-write state m_0 and inserted state m^* share the test boundary, then resume native recurrent updates under the same frozen policy K_θ. τ_M is the paired cue-alignment effect (Equation 1). (c) Scene, insertion time, and prior action traces vary across comparisons; trace includes physical consequences and fresh inputs. Unchanged states, deletion, add-back, and tokenwise norm-matched random edits use the same paired protocol.

1 INTRODUCTION

Robots often act on cues that have disappeared: selecting an object after its target cue is hidden, or choosing the next action based on completed steps. These tasks require policies to retain and use information beyond the current observation. Recent robot policies address this demand through recurrent states, memory banks, and history retrieval (Cherepanov et al., 2026; Shi et al., 2026a; Gao

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

et al., 2026; Sridhar et al., 2026). MemoryVLA stores perceptual details and high-level semantics, while MemER selects task-relevant historical keyframes for subsequent control.

Understanding memory requires connecting internal representations to behavior. Probes identify encoded information, while representation interventions test its influence on outputs (Geiger et al., 2021; Zhang & Nanda, 2024). Othello-GPT edits internal board representations to change predictions (Li et al., 2023); work on blind navigation reveals spatial maps in recurrent memory and their relationship to navigation behavior (Wijmans et al., 2023). We build on these approaches to examine memory's *conditions of use*: once a state can guide a correct choice, how does it act in another scene or later in the same task?

Each action changes the robot's position, its relation to objects, and the observations it receives next. The controller therefore uses memory within a changing interaction with its environment (Déletang et al., 2021). We vary the receiving scene, insertion time, and physical consequences of preceding actions while holding stored states fixed.

We study the recurrent μVLA policy through paired state replacement and shared-cue editing (Figure 1). After the cue disappears, we insert a saved state or an edited version into a test trajectory. Paired branches share current observations, physical scene, and executed actions, then continue under the original controller with fixed parameters. We estimate shared cue components from other source scenes and delete, add back, or exchange them to test their relationship to target choice, using random perturbations of matched magnitude as controls.

Shared cue components can direct the robot toward a specified new target: cue conversion increases targeted redirection by 82.8 percentage points over matched random edits. Yet the identical edited states lose much of this effect when inserted two actions later, with equal remaining action budgets. At the same late query, pausing the two preceding arm motions restores the effect of an unchanged native state. An independently collected source bank and tests covering all nine shape–color identities reproduce the editing and recovery results.

We independently set internal state and the physical consequences of a preceding action, then measure how they jointly affect choice. The gain from replacing internal state varies with these physical conditions. The receiving scene also provides information for predicting memory's effect. With both source and test scenes new, calibration from state-replacement trials reduces Brier prediction error by 66.0% relative to a same-budget rate baseline. On further new scenes, local cue use also predicts whether an incoming state will guide choice. All forecasts precede their corresponding test runs.

Our contributions are threefold. (1) **Linking memory content to target choice.** Cross-scene state replacement and shared-cue deletion, restoration, and conversion test how stored content guides the original controller (Figure 2). (2) **Identifying conditions of memory use.** Holding states fixed reveals effects of scene, insertion time, and preceding actions; action interventions restore cue-guided choice (Figures 3 and 4; Table 2). (3) **Predicting behavioral outcomes.** Calibration and local cue-use controls support prospective forecasts on new scenes, providing an experimental basis for evaluating both stored content and its conditions of use (Figure 5).

2 Testing Memory through State Replacement

The *source scene* supplies a saved state; the *test scene* receives it. We copy the complete recurrent tensor, or an edited version of it, after the cue disappears and before the next action is computed. At this *intervention time*, both test branches share current inputs, simulator state, clock, executed actions, and random state. Policy weights are fixed, and the controller resumes with its normal later updates.

Controls and behavioral effects. The *no-write* control sees and processes its cue but does not retain the cue-query recurrent updates; delay and later updates proceed normally. The *local-write* control retains cue updates in the test scene itself. Let m_c denote the source state associated with cue c, and m_0 the test-scene control state. A choice is *cue-aligned* if the selected object matches the cue's shape and color; after transfer, this means the *source* cue. Write $Y_{tc}^{b,K}(m)$ for the object chosen in test scene t after inserting m at time b and resuming controller K. Over fixed scene–cue units $\mathcal{U}$,

Table 1: Released robot policies, memory inputs, and behavioral endpoints. K2 and K8 denote separately fine-tuned μVLA checkpoints.

Policy	Task / memory input	Primary endpoint and role
μVLA K2	Shape×Color / recurrent tokens	Stationary object reach within 5 cm, 25 steps; transfer, boundaries, factorial, prediction
μVLA K8	ShellGamePush / recurrent tokens	Native cup reach, up to 30 steps; state–trace factorial
Long-hist	Match Color / visual history	Released-stable bin placement; content and dose controls
GMP	Placement / joint history	Bin placement, checked for release and stability; transport
MemoryVLA	ShellGameTouch / banks	Stationary cup reach within 8 cm after a conflicting prefix

the paired effect is

$$\tau_M(b,K) = \frac{1}{|\mathcal{U}|} \sum_{(t,c)\in\mathcal{U}} \left[\mathbf{1}\{Y_{tc}^{b,K}(m_c) = c\} - \mathbf{1}\{Y_{tc}^{b,K}(m_0) = c\} \right]. \tag{1}$$

This measures the change in cue-matching choices under those conditions. Resampling keeps dependent cue cells and shared controls together.

A changed choice shows influence; a choice that follows a specified cue ties that influence to task content. Reinserting an unchanged state reproduces every original action and geometry trace (Appendix A.1). Matched perturbations and cue conversion test content specificity with all other inputs present.

Policies and choice measurements. One released K2 checkpoint is the primary system for studying scene and time dependence; Table 1 lists the other policies and their roles. History interventions copy selected past inputs and leave current and unselected inputs, and their normal aging, unchanged. Recurrent-state interventions carry no pending action queue. Section 5 separately varies an earlier action's physical consequences.

All policies use released weights. Choices are scored by an observer that does not see the target, and wrong and unresolved choices stay in every denominator. Evaluation seeds are disjoint from earlier experiments. Fixed-phase cohorts present the cue for four queries and a blank for three and save state before query 7. Early runs last 25 actions; query-9 runs can extend to 27 so that both have 18 actions after insertion. Earlier cohorts use their original cue and delay schedules. Development and confirmation cohorts are separate (Appendices A and B).

3 Shared Cue Content Can Direct the Original Controller

Does a transferred state carry an object's identity, or only a movement toward its old position? And can that content be edited within the state?

Identity transfers across layouts. Across 768 transfers (eight K2 source scenes, 24 test scenes, four cues), 583 end in the correct stationary choice, compared with 12/96 for no-write and 84/96 for local-write controls: a paired gain of 63.4 percentage points (pp), with crossed source–test-scene interval $[55.1, 71.7]$ pp. The cued object usually sits elsewhere in the test scene (median displacement 20.3 cm). In the 671 transfers where cue identity and the old source position disagree, the robot reaches the cued object 509 times and the old-position object only 16 times; 103 choices go to another object and 43 are unresolved (Figure 2a; Appendix C.1).

Editing content shared across source scenes. We acquire a new library from eight sources, four cues per source, all captured before query 7. Let m_{sc} be a full state and $\bar{m}_{s\cdot}$ its source's mean across cues. A candidate cue component averages the centered states from the other sources:

$$u_{-s,c} = \frac{1}{S-1} \sum_{r\neq s} (m_{rc} - \bar{m}_{r\cdot}), \qquad S = 8. \tag{2}$$

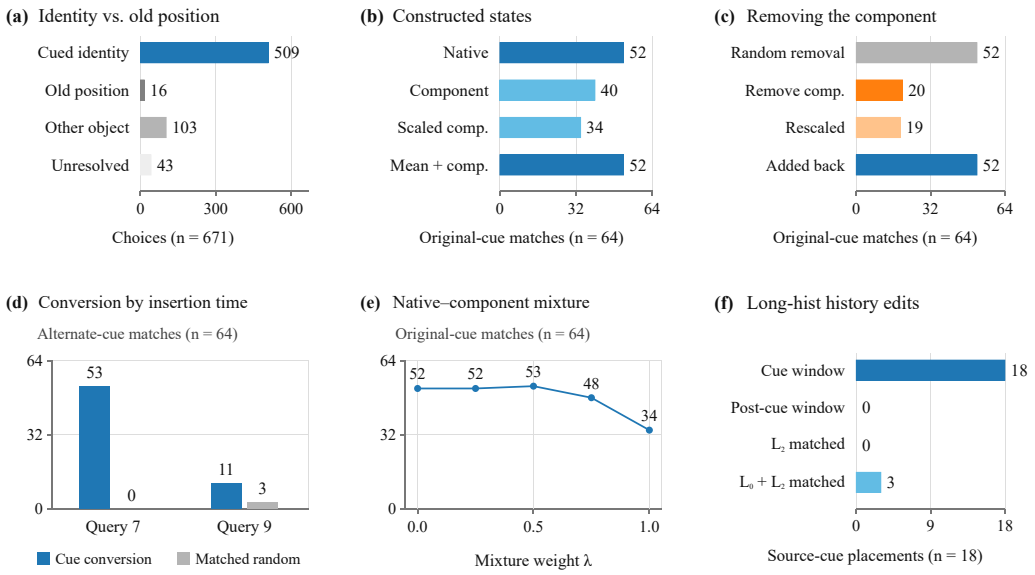

Figure 2: Shared cue edits redirect choice; perturbation size alone does not. Blue: native or edited states; gray: random or competing controls; orange: cue component removed. (a) K2 Shape×Color transfers in which cue identity and old source position conflict (eight sources, 24 test scenes, four cues). (b–e) Independent K2 confirmation with eight new sources, 16 new test scenes, and four cues; 64 units per arm. Choice is the first stationary reach within 5 cm; early insertion is at query 7 and runs end at step 25. (b) The component alone, the component scaled to the native norm, and the source mean plus component at the native norm. (c) Random removal matches each token's change magnitude; rescaling the deleted state does not restore its original cue choices. (d) Conversion, scored against the specified alternate cue. Late insertion places identical tensors at query 9 and runs to step 27, so both times allow 18 further actions. (e) Norm-matched mixtures from native state ($\lambda = 0$) to shared component ($\lambda = 1$). (f) Long-hist Match Color, 18 directed cue conflicts, first released and stable bin placement: the native cue window, the post-cue window, and cue windows matched to the post-cue edit in L_2 or $L_0 + L_2$.

We subtract this component, add it back, or replace it with the component for a predefined alternate cue c'; cue conversion, for example, inserts $m_{sc} - u_{-s,c} + u_{-s,c'}$. Random-direction controls match the perturbation's L_2 magnitude separately for each recurrent token. All candidates and tests are fixed before execution on 16 new test scenes with four cues each (Appendix B).

Cue conversion directs choice toward a specified object. Exchanging the component for that of the alternate cue sends 53/64 choices to the specified *alternate* object; matched random changes do so in 0/64 cases, a gain of 82.8 pp ($[71.9, 92.2]$). Removing the component leaves 20/64 original-cue matches, whereas matched random removal leaves 52/64, a specificity contrast of 50.0 pp ($[37.5, 62.5]$). A state built from the source mean plus the shared component, scaled to the native norm, gives 52/64 original-cue matches against 7/64 for no-write (70.3 pp, $[57.8, 81.3]$). All three primary contrasts have Holm-adjusted two-sided $p = 9.16 \times 10^{-5}$; intervals resample whole test scenes. Rescaling the deleted state to its original norm gives 19/64, whereas exact add-back reproduces every native action and geometry trace, including the 52/64 matches.

The surrounding state supports shared content. The component alone guides 40/64 choices, and 34/64 when scaled to the native norm; adding back the source mean restores 52/64. Total magnitude therefore cannot stand in for the rest of the state (Figure 2e traces norm-matched mixtures). A real state chosen by source-excluded centered cosine also gives 52/64, matching the native-state count (Appendix B).

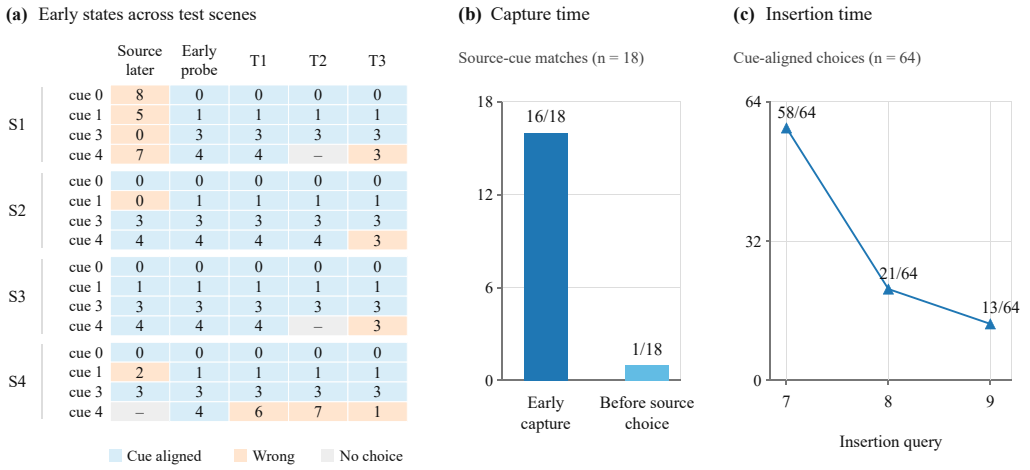

Figure 3: Scene, capture time, and insertion time change K2 state usefulness. All panels: Shape×Color, first stationary reach within 5 cm; a match selects the source-cued object. Panels (a–b) use horizon 25; panel (c) shows horizon-25 results and reports the horizon-27 control. (a) Each of 16 early states inserted unchanged into test scenes T1–T3. The first two columns give the source episode's later choice and the early probe's classification of the saved state's cue. Object codes 0/1/2 are red/green/blue cubes, 3/4/5 spheres, and 6/7/8 T-shapes in the same color order; a dash means no choice. (b) The six later-error states give 16/18 matches when saved early and 1/18 when saved just before the source choice; test insertion is early in both. (c) An independent 16-scene cohort with four cue and three blank queries inserts the same query-7 states at queries 7, 8, or 9. Extending query-9 trials to step 27 equalizes the post-insertion budget and still gives 13/64.

Past-input memories. For Long-hist, copying the cue-visible image window redirects 18/18 stable placements. Matching the smaller post-cue edit's magnitude gives 0/18 redirections, and matching both magnitude and changed-value count gives 3/18 (Figure 2f; Appendix F.1). In GMP's fixed-bin placement task, joint history gives 64/64 source-cue placements versus 4/16 balanced no-history matches.

4 A State's Usefulness Depends on Where and When It Is Used

We first hold a saved state fixed and change the test scene. In the full K2 grid of 32 states and 24 test scenes, every state produces at least one correct and one wrong choice. The nine states whose own source episodes chose wrongly still yield 139/216 correct transfers, against 29/216 for matched no-write controls. A source episode's outcome does not determine whether its saved state can guide another scene.

Useful states can precede a later error. A prospective cohort captures four cue states from each of four new sources before the first cue-hidden action query and tests them in three new scenes. A separately fitted classifier identifies all 16 early cues; the source continuations later make nine correct choices, six wrong choices, and one unresolved choice. The six states *saved before later errors* give 16/18 correct transfers versus 0/18 no-write choices, and all states together give 40/48 (Figure 3). Saving state immediately before the source's later choice instead gives 9/48 transfers, including 1/18 in the same error group, with test insertion still early. Late capture varies both the stored representation and the source phase (Appendix C.1).

An unchanged state loses cue-specific control when inserted later. On 16 new, unselected scenes with four cues each and a fixed cue/delay schedule, we save each local cue-written state before query 7 and insert exactly that tensor into the paired no-write trajectory at query 7, 8, or 9. Cue-matching choices fall from 58/64 to 21/64 and 13/64. The early-minus-delayed effect is 70.3 pp (whole-scene

Table 2: New-source replication and recovery at a fixed insertion time. K2 Shape×Color: two 16-scene cohorts, four cues, 64 units per arm. First stationary reach within 5 cm scores the new cue for conversion, the original cue for recovery. Query 7/9 insertion allows 18 further actions each. At query 9, pause zeros two prior arm commands; update recomputes state along the ordinary executed prefix. Differences: percentage points with paired-scene 95% intervals.

Comparison (left minus right)	Matches / 64	Difference [95% CI]
New bank: conversion vs. matched noise	51 / 0	79.7 [64.1,92.2]
New bank: same edited state, early vs. late	51 / 13	59.4 [46.9,70.3]
Recovery: paused vs. ordinary prefix	49 / 14	54.7 [37.5,71.9]
Recovery: updated vs. original state	13 / 14	-1.6 [-4.7,0.0]

95% interval $[59.4, 81.3]$; Holm-adjusted two-sided $p = 9.16 \times 10^{-5}$), and every scene declines. Insertion happens 22.7–34.7 cm from the nearest object, before any contact.

Time left to act does not explain the decline: extending query-9 trials to 27 actions, so that both conditions have 18 actions after insertion, leaves the count at 13/64, and the first 25 actions reproduce exactly. The ordering holds under all 30 prespecified radius, dwell, and stationarity definitions of choice. The cue also stays readable: a post-confirmation leave-one-scene-out readout identifies 63/64 of the same saved states from raw or scene-centered features. The decline is specific to cue alignment: delayed states still change 36/64 choices relative to no-write (Appendix C.2).

Edited content also depends on the receiving situation. The component-edit cohort repeats the test with constructed states and independently sampled sources and test scenes. Cue conversion gives 53/64 alternate-cue choices at query 7 and 11/64 at query 9, with identical tensors and 18 post-insertion actions in both conditions (Figure 2d). The predeclared secondary decline is 65.6 pp ($[54.7, 76.6]$); delayed conversion still changes 32/64 baseline choices. Native and source-mean-plus-component states both fall from 52/64 to 11/64.

A new source bank replicates the effect; changing prior actions restores use. Eight newly collected source scenes reproduce both the conversion effect and its timing dependence on 16 further test scenes (Table 2). A separate 16-scene cohort asks what changes between query 7 and query 9. We fix the saved state and its query-9 insertion but zero the two preceding arm-motion commands, keeping gripper commands, physics, and policy queries. Cue matches rise from 14/64 after the ordinary prefix to 49/64 after the paused one, whereas recomputing the state along the ordinary executed prefix gives 13/64. The same state at the same query becomes usable again when the preceding actions differ. With each prefix's own no-write control (10/64 ordinary, 4/64 paused), the state adds 6.25 versus 70.31 pp: a secondary state-by-prefix difference of 64.06 pp ($[48.44, 76.56]$; Appendix C.3). Extending the new bank to all nine shape–color identities on 16 further scenes repeats both results: color- and shape-targeted conversions give 106/144 and 111/144 choices versus 2/144 and 5/144 for matched noise, and pausing raises late matches from 27/144 to 100/144 (Appendix B.2). Across the three four-cue timing cohorts, early versus late matches are 58/64 versus 13/64 for native states, 53/64 versus 11/64 for converted states, and 51/64 versus 13/64 with the new bank.

5 Conditional Effects of State and Action Consequences

An action changes more than recurrent state: it also changes the physical scene and hence the next camera and proprioceptive inputs. We therefore vary the internal state independently of these action consequences at the next query. The *embodied trace* denotes the resulting simulator state and fresh inputs together.

For K2, we restore the same no-write simulator state and inputs at query 7. From this shared boundary, we execute one action from either the no-write control or the cue-retaining reference, then set the internal state before query 8. Reference construction is specified in Appendix A.1. This gives four combinations: neither change (baseline), the internal-state change alone (state-only), the action's consequences alone (trace-only), and both (joint). Let M_0 and M_C be the complete recurrent states from the no-write and cue-local references, and E_0 and E_C the physical consequences and fresh inputs generated by their respective actions at the shared boundary. With native control resumed, the

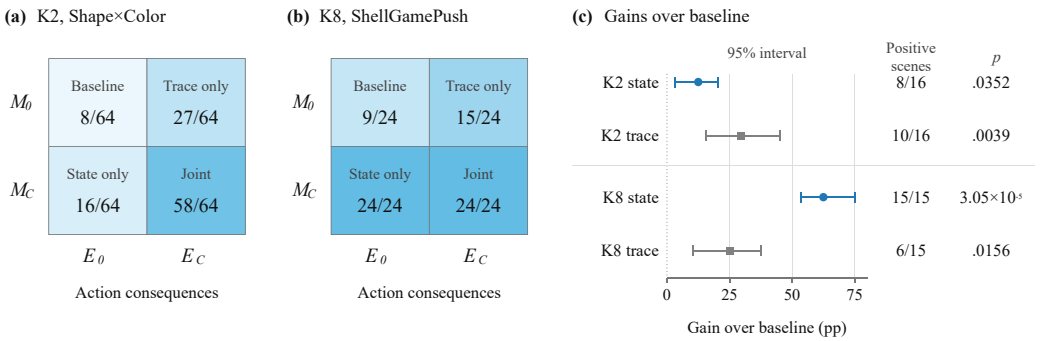

Figure 4: Internal state and one action's consequences have conditional effects. M_0/M_C are states without/with retained cue updates; E_0/E_C are the physical consequences and fresh inputs from their respective actions on the shared boundary. K2: 16 unselected Shape×Color scenes, four cues, first stationary reach within 5 cm by step 25; 64 units per cell. K8: 15 ShellGamePush scenes selected by local cue use, two cues, native cup reach by step 30; 24 units per cell. (c) Gains over M_0E_0 with 95% whole-scene intervals, the number of scenes that improve, and p values (K2: two-sided, Holm-corrected; K8: the original one-sided tests).

paired contrasts are

$$\Delta_M = \overline{Y}(M_C, E_0) - \overline{Y}(M_0, E_0), \qquad \Delta_E = \overline{Y}(M_0, E_C) - \overline{Y}(M_0, E_0), \tag{3}$$

where $\overline{Y}$ is the cue-aligned fraction over fixed scene–cue units. The two factors are set independently at insertion; afterwards, observations, actions, and state updates evolve together under the original controller.

K2: state acts conditionally on the physical situation. At query 8 the four combinations give $(8, 16, 27, 58)/64$ cue matches. Action consequences alone add 29.7 pp ($[15.6, 45.3]$; Holm-adjusted $p = .0039$). State alone adds 12.5 pp ($[3.1, 20.3]$; $p = .0352$) (Appendix D). The state effect is strongly conditional: replacing the state adds 8 matches under baseline consequences but 31 under cue-related ones, a descriptive post-confirmation interaction of 35.9 pp ($[20.3, 51.6]$).

K8: state and action consequences in ShellGamePush. We evaluate K8 on 24 cue units across 15 ShellGamePush scenes, selected by local cue use (Appendix D). The cells give $(9, 24, 15, 24)/24$: state and action consequences add 15 and six matches with no losses. Both scene intervals exclude zero; the one-sided intersection–union test gives $p = .015625$.

Combining physical state and current inputs. With internal state fixed, crossing full simulator state P with fresh input O gives $(1, 4, 4, 13)/24$ for baseline, input-only, physical-only, and joint: cue use rises most under joint replacement (Appendix D). In MemoryVLA, a conflicting executed prefix with empty banks yields 22/24 prefix-following choices, and enabling the banks changes one (Appendix F.2).

6 Context Dependence Supports Prospective Prediction

If usefulness depends on the receiving scene, observations of that scene might predict it. We forecast whether an incoming state will produce a source-cue match, using either calibration transfer outcomes or local cue-use controls, and fix every forecast before the corresponding test runs. Forecasts target source-cue matches at the original insertion time and are scored by Brier score, the mean squared error of the predicted match probability; lower is better.

Both source and test scenes are new. Eight new source and eight new test scenes, with four cues, give 64 calibration transfers and 192 held-out transfers. Six models share these labels and fixed settings: cue rate, source-only, scene-only, additive source–scene logistic regression, rank-one interaction, and additive regression with full-state cosine features (Appendix E.1).

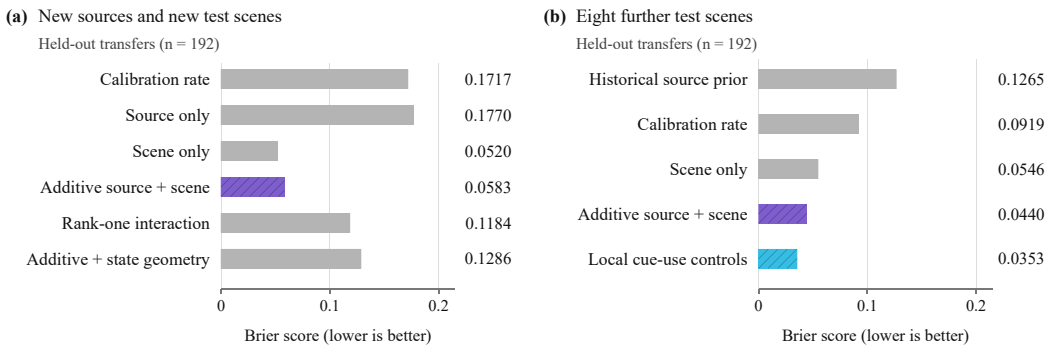

Figure 5: Transfer outcomes are predictable from the conditions of use. K2 Shape×Color: first stationary reach within 5 cm by step 25; wrong and unresolved outcomes count as nonmatches. (a) Eight new source and eight new test scenes, four cues: all six models receive the same 64 calibration outcomes, with settings fixed beforehand, and then predict 192 unrun transfers. (b) Eight further new test scenes: forecasts from 32 local cue-use controls, using probabilities fixed on the previous cohort, precede all 256 transfers; bars compare every method on the 192 transfers withheld from transfer calibration. The local forecast uses 32 local-control labels and the calibrated models 64 transfer labels.

The additive model reduces Brier error from .1717 (cue rate) to .0583, a 66.0% reduction (crossed source–scene interval for the improvement $[.0624, .1649]$). Most of the gain comes from the test scene: scene-only gives .0520, and the additive–scene interval includes zero, whereas source-only (.1770), rank-one (.1184), and state-geometry (.1286) models do worse. A scene-held-out readout identifies all 64 captured cues, yet these geometric features do not improve prediction. Prospective state selection gives 23/32, versus 25/32 for a single global source.

Local cue use predicts incoming-state utility. A cheaper signal is whether the test scene can use its own cue-written state. Having examined this diagnostic retrospectively, we fixed a two-probability rule and tested it on eight further new scenes with the existing source bank: 32 local controls ran, and forecasts for all 256 transfers were saved before any transfer ran. Brier error is .0379 versus .1289 for the historical source–cue prior, a 70.6% reduction (scene-paired interval $[.0542, .1292]$; exact two-sided $p = .0078125$). Local and incoming states agree on cue matching in 246/256 cases.

On the 192 transfers withheld from transfer calibration, errors are .0353 for local controls, .0440 additive, .0546 scene-only, and .0919 rate. The local–additive crossed interval includes zero, and the comparison uses 32 local controls versus 64 calibration transfers. Earlier separate new-source and new-test forecasts, with their retrospective logistic comparison, are in Appendix E.3.

7 Related Work

Memory in robot policies. Memory Gym, POPGym, RoboMME, and RoboMME-Interference evaluate memory-demanding behavior (Pleines et al., 2023; Morad et al., 2023; Dai et al., 2026; Rathi, 2026). Robot policies store the past in recurrent tokens, perceptual and cognitive memory banks, gated histories, retrieved experience, control-indexed or event-driven memories, and temporally conditioned memory fusion (Cherepanov et al., 2026; Shi et al., 2026a; Gao et al., 2026; Sridhar et al., 2026; Guo et al., 2026; Yang et al., 2026; Liang et al., 2026). These methods improve what is stored and retrieved; we hold stored content fixed and vary the situation in which it is used.

Measuring memory use. Related analyses quantify or study memory utilization (Parnichkun et al., 2025; Kwon et al., 2026), trace VLA representations to behavior (Shi et al., 2026b), and measure the value of sensory information to a robot (Krishna et al., 2025).

Interventions on representations. Interchange interventions and activation patching test whether a representation mediates behavior (Geiger et al., 2021; 2022; Zhang & Nanda, 2024). Related work

studies subspace steering (Makelov et al., 2024) and representation shifts under interventions (Grant et al., 2026). Our matched random edits test whether perturbation alone explains an effect. Steering methods apply related interventions to VLA policies (Häon et al., 2025; Khayatan et al., 2026). We examine how these intervention effects depend on the conditions of use.

8 Discussion

Why the same state stops working. Othello-GPT uses representation edits to test decoded board information (Li et al., 2023), and blind-navigation work probes a trained agent for spatial memory (Wijmans et al., 2023); both ask what a representation encodes. Here the original controller receives the same tensor at different times and uses it differently. A plausible explanation is that the controller reads memory jointly with the current geometry and task progress. The decline persists with equal action budgets, while pausing the preceding actions restores use of the same saved state.

Content and its surroundings. Adding the source mean supports more cue-guided choices than scaling the component to the native norm, implicating the rest of the state rather than total magnitude.

Actions as memory. Because actions change subsequent inputs, an agent can use its environment to carry information (Déletang et al., 2021). State and action-consequence effects therefore connect memory use to the robot's ongoing interaction with its environment.

Implications for evaluation. MemoryVLA and MemER improve access to useful history (Shi et al., 2026a; Sridhar et al., 2026); our results concern how that history is used. Scene-only prediction and local-control forecasts suggest that much of the predictable variation reflects a scene's capacity to use cues. We suggest a two-step evaluation: first establish that stored content can direct choice, then vary the conditions under which it is used, for which local cue-use controls are a cheap diagnostic.

9 Conclusion

Shared cue content can redirect the original controller toward a specified object. Whether it does so depends on the scene, on when the content is inserted, and on the physical consequences of earlier actions, and part of this dependence can be predicted from calibration or local cue-use runs. Robot memory should be evaluated on both counts: what stored content can direct, and the conditions under which the controller uses it.

Reproducibility Statement

The state-replacement and cue-editing methods are described in Sections 2 and 3. Appendix A specifies the experimental protocols, runtime versions, state definitions, and evaluation criteria. Detailed experimental settings, controls, sample sizes, prediction models, and statistical procedures are provided in Appendices B–F. Table 1 identifies the released policies and tasks. Appendix G provides a worked example of state capture, replacement, behavioral measurement, and replay verification.

Ethics Statement

The study uses released robot policies and simulated manipulation tasks.

LLM USE

LLMs were used to assist paper editing and to write the code for plotting experiments.

References

Egor Cherepanov, Nikita Kachaev, Daniil Zelezetsky, Aydar Bulatov, Artem Pshenitsyn, Yuri Kuratov, Alexey Skrynnik, Aleksandr I. Panov, and Alexey K. Kovalev. μVLA: On recurrent memory for partially observable manipulation in VLA models. *arXiv preprint arXiv:2606.12497*, 2026.

Yinpei Dai, Hongze Fu, Jayjun Lee, Yuejiang Liu, Haoran Zhang, Jianing Yang, Chelsea Finn, Nima Fazeli, and Joyce Chai. RoboMME: Benchmarking and understanding memory for robotic generalist policies. In *International Conference on Machine Learning*, 2026.

Grégoire Déletang, Jordi Grau-Moya, Miljan Martic, Tim Genewein, Tom McGrath, Vladimir Mikulik, Markus Kunesch, Shane Legg, and Pedro A. Ortega. Causal analysis of agent behavior for AI safety. *arXiv preprint arXiv:2103.03938*, 2021.

Yihuai Gao, Jeff Jinyun Liu, Shuang Li, and Shuran Song. Gated memory policy: In-context memorization and adaptation. *arXiv preprint arXiv:2604.18933*, 2026.

Atticus Geiger, Hanson Lu, Thomas Icard, and Christopher Potts. Causal abstractions of neural networks. In *Advances in Neural Information Processing Systems*, volume 34, pp. 9574–9586, 2021.

Atticus Geiger, Zhengxuan Wu, Hanson Lu, Josh Rozner, Elisa Kreiss, Thomas Icard, Noah Goodman, and Christopher Potts. Inducing causal structure for interpretable neural networks. In *International Conference on Machine Learning*, pp. 7324–7338, 2022.

Satchel Grant, Simon Jerome Han, Alexa R. Tartaglini, and Christopher Potts. Addressing divergent representations from causal interventions on neural networks. In *International Conference on Learning Representations*, 2026.

Xinying Guo, Chenxi Jiang, Hyun Bin Kim, Yuhang Han, Ying Sun, Yang Xiao, and Jianfei Yang. Chameleon: Control-indexed prospective memory for visuomotor manipulation. *arXiv preprint arXiv:2603.24576*, 2026.

Bear Häon, Kaylene Caswell Stocking, Ian Chuang, and Claire Tomlin. Mechanistic interpretability for steering vision-language-action models. In *Conference on Robot Learning*, 2025.

Pegah Khayatan, Sara Meziane, Jayneel Parekh, and Matthieu Cord. DiMaS: Distribution matching for steering vision-language-action models. *arXiv preprint arXiv:2607.14280*, 2026.

Arjun Krishna, Edward S. Hu, and Dinesh Jayaraman. The value of sensory information to a robot. In *International Conference on Learning Representations*, 2025.

Taeyoon Kwon, Dongwook Choi, Hyojun Kim, Sunghwan Kim, Seungjun Moon, Beong-woo Kwak, Kuan-Hao Huang, and Jinyoung Yeo. Embodied agents meet personalization: Investigating challenges and solutions through the lens of memory utilization. In *International Conference on Learning Representations*, 2026.

Kenneth Li, Aspen K. Hopkins, David Bau, Fernanda Viégas, Hanspeter Pfister, and Martin Wattenberg. Emergent world representations: Exploring a sequence model trained on a synthetic task. In *International Conference on Learning Representations*, 2023.

Yushen Liang, Yue Peng, Baosheng Jin, Tianluo Zhang, Xinyu Zhang, Shuyi Zhou, Zhuoran Chen, Xinqi Liu, and Shenji Wan. TFP: Temporally conditioned memory-fusion policies for visuomotor learning. *arXiv preprint arXiv:2607.08283*, 2026.

Aleksandar Makelov, Georg Lange, Atticus Geiger, and Neel Nanda. Is this the subspace you are looking for? An interpretability illusion for subspace activation patching. In *International Conference on Learning Representations*, 2024.

Steven Morad, Ryan Kortvelesy, Matteo Bettini, Stephan Liwicki, and Amanda Prorok. POPGym: Benchmarking partially observable reinforcement learning. In *International Conference on Learning Representations*, 2023.

Rom Parnichkun, Neehal Tumma, Armin W. Thomas, Alessandro Moro, Qi An, Taiji Suzuki, Atsushi Yamashita, Michael Poli, and Stefano Massaroli. Quantifying memory utilization with effective state-size. In *International Conference on Machine Learning*, 2025.

Marco Pleines, Matthias Pallasch, Frank Zimmer, and Mike Preuss. Memory Gym: Partially observable challenges to memory-based agents. In *International Conference on Learning Representations*, 2023.

Soumil Rathi. RoboMME-Interference: Benchmarking robot memory under interference. *arXiv preprint arXiv:2606.22338*, 2026.

Hao Shi, Bin Xie, Yingfei Liu, Lin Sun, Fengrong Liu, Tiancai Wang, Erjin Zhou, Haoqiang Fan, Xiangyu Zhang, and Gao Huang. MemoryVLA: Perceptual-cognitive memory in vision-language-action models for robotic manipulation. In *International Conference on Learning Representations*, 2026a.

Haoyuan Shi, Xiancong Ren, Yingji Zhang, Qinfan Zhang, Jiayu Hu, Haozhe Shan, Han Dong, Jinpeng Lu, Yinda Chen, Yi Zhang, Yong Dai, and Xiaozhu Ju. VLA-Trace: Diagnosing vision-language-action models through representation and behavior tracing. *arXiv preprint arXiv:2605.30117*, 2026b.

Ajay Sridhar, Jennifer Pan, Satvik Sharma, and Chelsea Finn. MemER: Scaling up memory for robot control via experience retrieval. In *International Conference on Learning Representations*, 2026.

Erik Wijmans, Manolis Savva, Irfan Essa, Stefan Lee, Ari S. Morcos, and Dhruv Batra. Emergence of maps in the memories of blind navigation agents. In *International Conference on Learning Representations*, 2023.

Ganlin Yang, Zhangzheng Tu, Yuqiang Yang, Sitong Mao, Junyi Dong, Tianxing Chen, Jiaqi Peng, Jing Xiong, Jiafei Cao, Jifeng Dai, Wengang Zhou, Yao Mu, and Tai Wang. EventVLA: Event-driven visual evidence memory for long-horizon vision-language-action policies. *arXiv preprint arXiv:2606.20092*, 2026.

Fred Zhang and Neel Nanda. Towards best practices of activation patching in language models: Metrics and methods. In *International Conference on Learning Representations*, 2024.

A EXPERIMENTAL PROTOCOL

All studies evaluate released policies with frozen weights in simulation. New scenes are disjoint from earlier evaluations. Table 3 summarizes the cohorts.

Table 3: Cohorts supporting the main conclusions. Units are scene–cue cells with shared controls. Historical/development studies are distinguished from tests fixed before new outcomes.

Question	Evaluated sample	Role and scope
Cross-layout transfer	8 sources, 24 test scenes, 4 cues; 768 transfers	Historical grid; scene dependence and prediction development
Useful state before source errors	4 sources, 3 test scenes; 48 transfers	Confirmation; source-error group fixed before test execution
Same-state timing and factorial	16 scenes, 4 cues; 64 units per arm	Fixed-phase confirmation; one K2 checkpoint
Shared-cue editing	8 sources, 16 test scenes, 4 cues	Separate confirmation; source-excluded components
New-bank editing / prefix recovery	8 new sources; two disjoint 16-scene cohorts	Independent source bank, shared between these two cohorts
Nine-identity extension	Same 8 sources extended to 9 cues; 16 further scenes	Confirmation over 144 units per arm
Joint source–scene prediction	8 new sources, 8 new test scenes, 4 cues	64 calibration and 192 held-out transfers
Local-control prediction	Existing sources, 8 further test scenes	32 local controls precede 256 transfer forecasts
K8 state–trace extension	24 cue units across 15 of 40 candidate scenes	Local cue use determines sample eligibility
Long-hist / MemoryVLA / GMP	18 frame conflicts / 24 prefix conflicts / 64 transports	Supporting tests with policy-specific endpoints

A.1 INTERVENTIONS AND RUNTIME

At a replacement boundary, paired branches match current inputs, the restored simulator state, executed prefix, clock, and random state. Only the specified memory or action consequence is changed; subsequent control and state updates are native. No-write processes the cue but restores the preceding recurrent state after each cue query. Delay and later updates are committed normally. Equality checks cover the state inventory in Table 4.

Fixed-phase K2 cohorts use four cue queries, three blank queries, and capture before query 7. The policy predicts eight actions, executes the first, and replans without a carried action queue. Complete states have shape $1 \times 64 \times 4096$ and bfloat16 precision. Query-7 and query-9 insertions use horizons 25 and 27, respectively, when matching the 18-action continuation budget. The no-write reference uses cue 8; ordinary four-cue local references use $(0, 1, 3, 4)$. Local references retain their cue updates, receive the no-write reference's inputs at queries 4–7, and restore its simulator state at query 7. Their query-7 observations, physical state, and RNG match, while their earlier executed actions can differ. After query 7, each branch uses its own fresh inputs and native control. State-replacement branches follow the no-write trajectory up to insertion. Earlier cohorts use their original cue and delay schedules. The fixed-phase runtime uses PyTorch 2.2.1, the memory-enabled Transformers 4.40.1 fork, SAPIEN 3.0.0b1, and ManiSkill 3.0.0b15.

We compare unmodified-state shams with reference actions and geometry, consistent factorial corners with their specified references, and mixed branches with their designated states and action consequences. History edits preserve dtype, slot order, and normal aging, while leaving current and unselected inputs in place. MemoryVLA logs proposed and executed actions separately.

A.2 ENDPOINTS AND INFERENCE

The primary K2 choice is the first post-action event with a unique nearest object within 5 cm and `robot.is_static(.2)`. Exact ties and missing events are unresolved. The target-blind observer scores object identity. Wrong and unresolved outcomes remain in every denominator. Alternative

Table 4: State inventory matched at the insertion boundary. RNG denotes random state; KV denotes key–value cache. Checkpoints and normalization parameters remain fixed.

Policy	Captured or intervened state	Matched boundary variables
μVLA	Complete runner-local `mem_state` tensor	Observation, restored simulator state, clock, write schedule, RNG; no carried action queue and KV cache disabled
Long-hist/GMP	Image-history tensors and proposed-action histories with chronology, shapes, dtypes, and diffusion levels	Recipient current input, query index, episode state, executed prefix, native gate, and RNG
MemoryVLA	`cur_timestep`, cognitive bank, perceptual bank, bank dictionaries, and `eid_stream`	Native query schedule, executed prefix, current image, physical record, queue state, and continuation RNG

radii, dwell times, and stationarity rules are applied to the stored trajectories. K8 instead uses first native ShellGamePush cup reach by step 30, without stationarity; its saved `stationary_choice` field retains that native definition and native tie ordering. Long-hist/GMP use bin placement, checked for release and stability. MemoryVLA uses stationary cup reach within 8 cm from step 9 through the 90-step horizon.

Inference keeps all cues and shared controls within the same scene block. Fixed-phase K2 factorial, timing, editing, and recovery tests average paired differences within each of 16 test scenes. The prespecified relevance threshold is 20 percentage points (pp), with exact two-sided scene sign-flip $p < .05$ and Holm correction within each declared family. Sign-flip inference assumes symmetric paired differences under the null. The 20,000-draw percentile bootstrap gives pointwise 95% intervals. Editing and recovery use bootstrap seeds 9202026 and 9222026. Crossed source–test analyses resample both axes while retaining cue blocks and the evaluation mask.

Confirmation samples, eligibility, and stopping rules were fixed before outcomes. Four development scenes informed the 16-scene editing design: an approximate 80%-power calculation gave detectable effects of $(.234, .303, .165)$ at level $.05/3$.

B Shared-Cue Editing and Source Replication

Eight fresh source scenes provide four cues each, without selection by later behavior. Sixteen new test scenes each receive all four cues; test index modulo eight assigns the source, reusing each source twice. The cohort runs in four execution partitions. Component estimates (Equation 2) exclude the assigned source and never use test-scene tensors or outcomes. The alternate cue is the next entry in $(0, 1, 3, 4)$, with wrap-around.

Write $x = m_{sc}$, $u = u_{-s,c}$, $v = u_{-s,c'}$, and $b_s = \bar{m}_{s\cdot}$. Norm matching uses

$$N_x(z) = z\,\frac{\|x\|_F}{\|z\|_F}. \tag{4}$$

Gaussian random directions ϵ_u and ϵ_{v-u} are rescaled within each token to match u and $v - u$, using fixed scene/cue seeds. Arithmetic is performed in float64 before bfloat16 casting.

Table 5: Complete early component-edit results on 64 K2 scene–cue units, query 7 and horizon 25. Original and alternate are distinct cue identities; other resolved choices fill the remaining denominator. Wrong-cue add-back denotes the conversion arm.

State	Definition	Original	Alternate	Unresolved
Native	x	52	0	2
Component	u	40	0	5
Norm-matched component	$N_x(u)$	34	1	5
Source mean + component	$b_s + u$	52	0	3
Norm-matched mean + component	$N_x(b_s + u)$	52	0	3
Deletion	$x - u$	20	14	2
Norm-matched deletion	$N_x(x - u)$	19	16	1
Matched random deletion	$x - \epsilon_u$	52	0	3
Correct add-back	$(x - u) + u$	52	0	2
Cue conversion	$(x - u) + v$	1	53	1
Matched random conversion	$x + \epsilon_{v-u}$	52	0	2
Mixture, $\lambda = .25$	$N_x(.75x + .25u)$	52	0	3
Mixture, $\lambda = .50$	$N_x(.50x + .50u)$	53	0	2
Mixture, $\lambda = .75$	$N_x(.25x + .75u)$	48	0	5
Centered representative	Real state; see text	52	0	2
Raw-cosine representative	Real state; see text	52	0	2

The centered representative selects the real state maximizing $\cos(m_{sc} - \bar{m}_{s\cdot}, u_{-s,c})$ across sources, with each prototype excluding the candidate source. The raw comparator maximizes mean pairwise same-cue cosine. Ties select the first source. Both rules select source index 2 for every cue here.

Table 6: Prespecified early editing contrasts. Counts are matches out of 64 in each arm; conversion targets the alternate cue. Intervals resample 16 paired scenes.

Contrast	Counts	Gain [95% CI], pp	Holm p
Norm-matched mean + component versus no-write	52 / 7	70.3 [57.8,81.3]	9.16×10^{-5}
Matched random versus component deletion	52 / 20	50.0 [37.5,62.5]	9.16×10^{-5}
Cue conversion versus matched random change	53 / 0	82.8 [71.9,92.2]	9.16×10^{-5}

The first and third contrasts improve in all 16 scenes; deletion specificity improves in 15 with one tie. All 64 correct add-backs exactly reproduce native actions and geometry. Norm-matched mean-plus-component exceeds norm-matched component by 28.1 pp ($[20.3, 35.9]$). Representative minus native is 0 pp ($[-4.7, 4.7]$).

After casting, maximum relative norm error is .00001992 (tolerance .0005); tokenwise perturbation errors are .00032474 for deletion and .00027342 for conversion (tolerance .005). Readouts classify all 64 native/component assignments correctly and all 64 conversions as the alternate cue; deletion retains 2/64 original-cue labels, or 4/64 after norm matching. These 64 assignments reuse 32 source states. Across all 35 balanced source splits, full-state cue-contrast cosines range .5672–.8742; tokenwise values range .5033–.9059. Cue, source, and residual interaction account for .4129, .1768, and .4102 of the algebraic sum of squares.

B.1 Replication with a new source bank

Eight newly sampled sources and 16 new, unselected test scenes repeat the protocol. Conversion yields 51/64 alternate-cue matches versus 0/64 for noise; deletion yields 16/64 original-cue matches versus 51/64 for random deletion. Native and exact add-back both yield 50/64, with exact replay. The same converted tensors yield 13/64 when inserted at query 9. The three effects and their inference are retained together with the recovery cohort in Table 12.

B.2 ALL NINE CUE IDENTITIES

The new bank is extended to all nine shape–color identities, reproducing the earlier 32 captures bit for bit. Sixteen further unselected test scenes each receive all nine cues. Two fixed conversions target the next color with shape held fixed, or the next shape with color held fixed. Noise controls match per-token perturbation magnitude. Recovery uses the same query-9 paused-prefix intervention as Appendix C.3. The three primary contrasts use the same 20 pp rule, scene-level inference, and within-family Holm adjustment.

Table 7: Nine-identity confirmation: 144 units per arm across 16 test scenes. Conversion scores the specified alternate identity; recovery scores the original cue. Intervals are paired-scene differences in pp.

Comparison (left minus right)	Matches / 144	Difference [95% CI]	Holm p
Color conversion vs. matched noise	106 / 2	72.2 [61.8,81.9]	9.16×10^{-5}
Shape conversion vs. matched noise	111 / 5	73.6 [61.8,84.7]	9.16×10^{-5}
Paused vs. ordinary prefix, query 9	100 / 27	50.7 [42.4,59.0]	9.16×10^{-5}

Local-write and early native states match 110/144 and 107/144 cues. For the five newly included identities, the descriptive counts are 51/80 versus 0/80 for color conversion, 66/80 versus 3/80 for shape conversion, and 52/80 versus 13/80 for paused versus ordinary prefixes. Early and paused choices agree in 108/144 units, including errors. These whole-identity edits cover 18 directed conversions.

C SCENE, INSERTION TIME, AND PRIOR ACTIONS

C.1 CROSS-LAYOUT TRANSFER AND SOURCE ERRORS

The historical grid contains eight source scenes, 24 test scenes, and four cues. Each of the 96 no-write and 96 local-write control executions is shared across eight source comparisons. All eight source-level and all 24 test-scene-level paired stationary-choice effects are positive; crossed source–scene resampling gives the 63.4 pp gain and $[55.1, 71.7]$ interval reported in the main text. Of 671 identity/old-position conflicts, 509 follow cue identity, 16 the old position, 103 another object, and 43 are unresolved.

Table 8: Cross-layout K2 outcomes under three readouts. Controls have 96 executions per arm, shared across eight sources; transfers have 768. No-write suppresses cue updates only. Wrong and unresolved choices are retained.

Readout	Memory	Correct	Wrong	Unresolved	Total
Proximity	Source	574	149	45	768
	No-write	12	84	0	96
	Local	83	11	2	96
Stationary	Source	583	137	48	768
	No-write	12	84	0	96
	Local	84	10	2	96
Final success	Source	601	167	0	768
	No-write	12	84	0	96
	Local	84	12	0	96

Every one of the 32 states has both a correct and an observed-wrong test outcome. Transfers from source-correct, source-error, and source-unresolved states yield 422/528, 139/216, and 22/24 matches, against balanced no-write counts 63/528, 29/216, and 4/24. Grouping uses observed source outcomes.

Prospective early-state retention. Four new sources supply four early states each to three new test scenes, giving 48 transfers within 86 total trajectories. Capture queries for sources S1–S4 are 9, 3, 5, and 6; their prefix lengths differ. Source continuations yield nine correct, six wrong, and one unresolved choice. A Euclidean nearest-centroid cue readout fitted on four separate sources labels all

16 early states correctly. The six-state error group is fixed under the primary readout before transfer outcomes. Five of these states, from three sources, improve cue matching in at least two test scenes each, with no paired decrease.

Table 9: Early and event-aligned late source states in the same three test scenes. Match/wrong/unresolved counts sum to 48; error-group matches are out of 18. No-write matches are zero throughout. The late follow-up uses the earlier experiment as its reference.

Readout	Early M/W/U	Late M/W/U	Early error group	Late error group
Stationary 5 cm	40/6/2	9/37/2	16	1
Stationary 4 cm	36/4/8	7/23/18	14	1
Stationary 6 cm	42/6/0	9/39/0	17	1
Stationary 5 cm, dwell 2	41/5/2	10/35/3	16	1
Final stationary 5 cm	38/2/8	9/21/18	14	1

Late capture occurs immediately before each source's observed choice; the unresolved source uses its final query. All 16 states are transferred with fresh controls and shams (66 trajectories). Saved source states, action traces, and all 15 test controls match the earlier run. The early probe labels 7/15 choice-query states and 4/16 final states correctly; late transfer yields 9/48 matches overall and 1/18 in the source-error group.

C.2 Fixed-state timing confirmation

Sixteen new scenes and four cues test the exact same query-7 native states at queries 7, 8, and 9. The cohort has 592 trajectories including references, shams, timing, factorial, and extended-horizon branches. The extension reproduces the first 25 actions exactly. All 30 prespecified combinations of 4–6 cm radii, dwell 1–3, and stationarity inclusion retain the early-over-late ordering. Query-8 and query-9 insertion still change 52/64 and 36/64 choices relative to no-write, including unresolved transitions. Cue readout is 63/64 under leave-one-scene-out raw or centered features; that diagnostic was performed after confirmation.

Table 10: Fixed-state timing on 64 K2 units. The same query-7 tensor is inserted at each time; horizon 27 equalizes the continuation budget.

Insertion	Horizon	Match	Wrong	Unresolved
Query 7	25	58	5	1
Query 8	25	21	29	14
Query 9	25	13	42	9
Query 9, equal action budget	27	13	43	8

The primary early-minus-query-9 effect is 70.3 pp ($[59.4, 81.3]$), Holm-adjusted $p = .0000916$. Its correction family also contains the two fixed-phase factorial contrasts (Appendix D). Independent component edits show the same dependence (Table 11). Early-minus-late differences are 64.1 pp ($[53.1, 75.0]$) for native, mean-plus-component, and representative states; 35.9 pp ($[26.6, 46.9]$) for norm-matched components; and 65.6 pp ($[54.7, 76.6]$) for conversion. These comparisons were prespecified secondary analyses.

Timing controls. Spatially defined 25/20/15 cm boundaries produced zero cue-aligned choices out of 12 units for every tested early or phase-local insertion, despite 9/12 correct local continuations. Action-indexed development instead gave 9/12, 2/12, and 0/12 for insertion at the cut, one action later, and two actions later. The subsequent factorial uses the one-action boundary. An exploratory update-rollback study on four exposed sources gave 12/16 matches versus 9/16 with normal updates (no-write controls: 1/16 versus 2/16).

C.3 Recovery at a fixed late query

A second 16-scene cohort uses the same new bank as the editing replication, with 512 test trajectories. Ordinary and paused prefixes share their first seven actions. At queries 7 and 8, the pause zeros six

Table 11: Identical edited tensors inserted early (query 7/horizon 25) or late (query 9/horizon 27). All rows retain 64 units and 18 subsequent actions. Conversion and random-conversion matches target the alternate cue; other rows target the original. Late changed: choices differing from no-write, including unresolved transitions.

State	Early matches	Late matches	Late unresolved	Late changed
Native	52	11	13	31
Norm-matched component	34	11	12	29
Norm-matched mean + component	52	11	12	31
Cue conversion	53	11	12	32
Matched random conversion	0	3	12	33
Centered representative	52	11	11	30

arm-motion increments but retains the ordinary gripper command. Physics, queries, and recurrent updates continue. Both branches receive the identical saved query-7 native state at query 9, with 18 further actions.

The update control inserts that source state at query 7, forces the ordinary prefix's next two executed actions, and retains its recurrent updates. Inserting the resulting state at the ordinary query-9 boundary reproduces the update reference. The pause changes physical configuration and current inputs jointly.

Table 12: New-bank and prefix-recovery primary contrasts. Each block uses its own 16-scene cohort, 64 units per arm, and three-test Holm family. Intervals are scene-paired differences in pp.

Comparison (left minus right)	Matches / 64	Difference [95% CI]	Holm p
Random deletion vs. cue deletion	51 / 16	54.7 [39.1,70.3]	1.83×10^{-4}
Conversion vs. matched noise	51 / 0	79.7 [64.1,92.2]	1.83×10^{-4}
Same converted tensor, early vs. late	51 / 13	59.4 [46.9,70.3]	1.83×10^{-4}
Paused vs. ordinary prefix	49 / 14	54.7 [37.5,71.9]	3.66×10^{-4}
Updated vs. original state	13 / 14	-1.6 [-4.7,0.0]	1
Paused conversion vs. matched noise	50 / 2	75.0 [60.9,89.1]	9.16×10^{-5}

The new-bank editing and prefix-recovery cohorts have separate three-test Holm families. The secondary state-by-prefix comparison uses each prefix's own no-write control: 10/64 ordinary and 4/64 paused. Corresponding state effects are 6.25 pp ($[0, 14.06]$) and 70.31 pp ($[59.38, 81.25]$), with difference 64.06 pp ($[48.44, 76.56]$). Neither reference selects an object before query 9. Early native matching is 52/64, and early/paused choices agree in 57/64, including errors. A transductive source-excluded template readout labels all 64 cues before and after the update, using four-state batch centering.

D State and Action Consequences

For fixed-phase K2, references share the restored query-7 boundary specified in Appendix A.1. A no-write or cue-local action advances the simulator and produces fresh inputs; before query 8, the runner independently supplies the no-write or cue-local recurrent state. Thus E contains the action's physical consequences and observations, and M the state at insertion. Native control then resumes. The K8 construction is specified below.

Table 13: State–trace factorial cohorts, retaining all cue units. K2 uses stationary reach on Shape×Color; K8 uses native cup reach on ShellGamePush, with sample selection described below.

Cohort	M_0E_0	M_CE_0	M_0E_C	M_CE_C
K2 development, 3 scenes	0/12	1/12	6/12	9/12
K2 historical confirmation, 6 scenes	1/24	7/24	13/24	18/24
K2 fixed-phase confirmation, 16 scenes	8/64	16/64	27/64	58/64
K8 confirmation, 15 scenes	9/24	24/24	15/24	24/24

K2 factorial results. In the earlier six-scene cohort, state-only matching is 7/24. A retrospective scene-level analysis gives state and trace gains of .25 ($[.083, .417]$, one-sided $p = .0625$) and .50 ($[.333, .708]$, $p = .015625$); direct trace-minus-state $p = .125$.

The fixed-phase cohort is the main-text result. Baseline, state-only, trace-only, and joint have wrong/unresolved counts $(52, 4)$, $(44, 4)$, $(29, 8)$, and $(5, 1)$, respectively. State and trace gains are 12.5 pp ($[3.1, 20.3]$, Holm $p = .0351563$) and 29.7 pp ($[15.6, 45.3]$, $p = .0039063$). State's effect under cue-related consequences is 48.4 pp ($[35.9, 62.5]$); the 35.9 pp interaction ($[20.3, 51.6]$) is a post-confirmation descriptive contrast.

What is inside the trace. The six-scene diagnostic fixes M_0 and crosses full simulator state P with freshly prepared camera/proprioceptive inputs O. Consistent corners replay exactly. Mixed arms each give three scene gains and three ties: mean .125, interval $[.042, .208]$, one-sided $p = .125$. Matching is 4/24 in each mixed arm and 13/24 in the consistent compound.

Table 14: Crossing physical state P and fresh inputs O with internal state fixed to M_0. Each row retains 24 units from six K2 scenes. Consistent corners replay exactly.

Cell	Cue-aligned	Wrong	Unresolved	Improved scenes vs. P_0O_0
P_0O_0 (baseline)	1/24	15	8	—
P_0O_C (fresh input only)	4/24	16	4	3/6
P_CO_0 (physical state only)	4/24	16	4	3/6
P_CO_C (consistent compound)	13/24	10	1	6/6

K8 ShellGamePush protocol. No-write and local controls were evaluated on 40 candidate scenes. Local controls matched cues 1 and 2 in 19 and 17 scenes, respectively. The first 12 matching scenes per cue supplied 24 units across 15 scenes; these local controls reproduced before the factorial evaluation. Applying the designated cue-local action to the exact no-write precursor defines E_C. All 109 serialized simulator entries, current observations, and geometry match their intended references; maximum closure error is zero. State-only has 15 gains and no losses across 15 positive scenes (one-sided $p = 3.05 \times 10^{-5}$, interval $[.536, .750]$); trace-only has six gains and no losses across six scenes ($p = .015625$, $[.105, .375]$). The intersection–union conjunction is $p = .015625$.

Additional K8 evaluations. Local cue matching on Shape×Color was 6/24 under the earlier schedule and 0/16 at fixed phase. A ShellGamePush development cohort gave $(11, 22, 16, 24)/24$, with five trace gains across five scenes. A separate two-unit run was excluded for mismatched boundary prefixes.

E Prospective Prediction and Baselines

Forecasts predict whether a transferred state produces a source-cue match. Wrong and unresolved both map to zero. Prospective means the relevant model and predicted probabilities were fixed before its test outcomes; retrospective comparisons below are identified separately.

E.1 Jointly new source and test scenes

Eight sources and eight test scenes are new to evaluation, with four cues. For each cue, calibration uses source index equal to the test index or its successor modulo eight. This connected cycle gives 64 calibration and 192 held-out cells. Eighty source/test references and eight shams accompany the transfers (344 trajectories). Local controls yield 28/32 source and 25/32 test matches; test no-write gives 4/32. All sampled scenes are retained, and the six prediction fits use transfer-calibration labels only.

For each cue, logistic models use an intercept and source offsets, test-scene offsets, or both. Rank-one adds a product of source and scene latent scalars. Quadratic penalties are 3, .3, .1, and 1 for source, scene, additive, and rank-one models, selected on three historical eight-scene blocks under the same cycle mask. The intercept penalty is 10^{-6}. Rank-one uses four deterministic initializations and selects penalized calibration likelihood. The geometry model adds cosine between complete within-scene

cue-centered source/local-test states, standardized over unlabeled pairs per cue, with penalty fixed at 1. It uses additional captured tensors with the same behavioral labels and policy calls. The rate baseline uses Jeffreys smoothing.

Held-out probabilities and original selection rules were fixed before testing. The prespecified primary comparison uses an additive-minus-rate Brier improvement of at least .02 with a positive crossed source–scene 95% interval. The observed improvement is .1134 ($[.0624, .1649]$). Twenty thousand bootstrap draws resample sources and test scenes independently while retaining cue blocks and the held-out mask.

Table 15: Jointly new source and test scenes: six prespecified predictors on the same 192 held-out transfers after 64 calibration labels. Lower Brier scores and log loss are better; binary forecasts use a .5 threshold.

Predictor	Brier	Log loss	Correct binary forecasts
Calibration rate	.1717	.5252	149/192
Source only	.1770	.5415	149/192
Scene only	.0520	.2155	185/192
Additive source–scene	.0583	.2287	184/192
Rank-one interaction	.1184	.3807	157/192
Additive with state geometry	.1286	.4088	153/192

The additive improvement over scene-only is $-.0064$ with interval $[-.0195, .0033]$. These secondary comparisons are descriptive. For each of 32 test-scene–cue decisions, selection is restricted to the same six non-calibration sources. Global source quality and the geometric medoid each give 25/32 matches; rank-one and geometry-augmented selection each give 23/32. Uniform selection's expected matching fraction over this fixed library is .7760. Context selection minus global-source performance is $-.0625$ ($[-.15625, 0]$, two-sided $p = .5$). An additional additive selector also gives 23/32; its rule was registered after the original probabilities were fixed, during the first held-out run and before its outcome.

E.2 Independent local-control forecast

A retrospective analysis of the joint cohort found local/transfer agreement in 187/192 cases. A subsequent rule was fixed at probabilities .0581395 when the local cue-written control fails and .9768212 when it matches, using Jeffreys-smoothed conditional rates from the previous 192 outcomes. Historical source–cue priors use all 256 previous transfers.

Eight further test scenes use the existing bank. Eight no-write references, 32 local controls, and eight shams precede all 256 transfer forecasts. After the 64 designated calibration transfers, rate, scene-only, and additive probabilities are fixed before testing the remaining 192 cells. Models retain their original settings. The primary local-versus-historical-source comparison uses all 256 transfers and eight scene clusters, requiring at least .02 Brier improvement and two-sided scene sign-flip $p < .05$. Improvement is .0910 ($[.0542, .1292]$, $p = .0078125$).

Table 16: Independent local-control forecasts. The first block is fixed before all 256 transfers; the second evaluates every method on the same 192 cells excluded from transfer calibration. Local controls use 32 local labels; calibrated models use 64 transfer labels.

Predictor / evaluation set	Brier	Log loss	Binary accuracy
Local controls / all 256	.0379	.1704	246/256
Historical source prior / all 256	.1289	.4247	216/256
Historical cue rate / all 256	.1295	.4231	216/256
Local controls / same 192	.0353	.1594	185/192
Historical source prior / same 192	.1265	.4197	163/192
Historical cue rate / same 192	.1270	.4175	163/192
Calibration rate / same 192	.0919	.2859	161/192
Scene only / same 192	.0546	.1983	179/192
Additive / same 192	.0440	.1391	182/192

Local controls produce 28 correct, three wrong, and one unresolved choice; transfers produce 216 correct, 37 wrong, and three unresolved. On the shared 192-cell set, local-minus-additive Brier improvement is .0087 ($[-.0403, .0428]$), and local-minus-scene improvement is .0193 ($[.0042, .0368]$).

E.3 Earlier separate-axis forecasts

The earlier source-order model was developed on the historical $8 \times 24 \times 4$ grid. Per cue it exhausts $8!$ source orders and fits recipient thresholds from -1 (all sources predict a match) to 7 (none do), with Jeffreys-smoothed error and Laplace-one threshold counts. Nested leave-one-recipient-out development selects zero-based anchor ranks 1 and 6 from five candidate pairs.

Six new test scenes receive two fixed anchor transfers per cue, reusing the known source-state captures. Probabilities for all 144 non-anchor transfers are fixed before testing; the run totals 254 trajectories. Outcomes are 95 matches, 46 wrong, and three unresolved. Calibrated order reduces Brier error relative to uncalibrated order and source frequency in all six scenes (one-sided $p = .015625$ for each comparison). Calibrated-minus-baseline scene intervals are $[-.177, -.086]$ and $[-.184, -.066]$.

The reciprocal test captures six new sources and reuses the calibrated test scenes. Two old test scenes per cue update a uniform prior over the nine insertion gaps among eight known source ranks; recipient posteriors remain fixed. All 144 probabilities are fixed before testing (228 total trajectories). Outcomes are 104 matches, 37 wrong, and three unresolved. Calibrated order has lower Brier error than the prospective baselines (Table 17); crossed source–test intervals include zero against the two closest baselines.

Table 17: Earlier separate-axis predictions, 144 held-out transfers per cohort. Ordering forecasts are prospective. The same-label logistic comparison was added retrospectively. Lower Brier and log loss are better.

Cohort	Predictor	Brier	Log loss	Correct labels
New test	Calibrated order	.0459	.1738	138/144
	Order, no anchors	.1761	.5341	111/144
	Development source frequency	.1700	.5132	111/144
	Current two-anchor rate	.0972	.3457	129/144
	Current local ability	.2117	.6108	95/144
	Development base rate	.2343	.6660	95/144
	Retrospective additive logistic	.0443	.1667	137/144
New source	Calibrated order	.0702	.2459	132/144
	Uniform insertion gap	.0842	.2881	132/144
	Recipient known-source average	.0892	.3020	132/144
	Two-anchor source rate	.2315	.6613	106/144
	Retrospective additive logistic	.0625	.2103	132/144

Same-budget logistic comparison. The added baseline fits $q_{stc} = \sigma(\alpha_c + u_{sc} + v_{tc})$ with summed cross-entropy and quadratic penalties on source/test effects. Old-grid validation leaves out each axis separately, uses two fixed calibration cells per cue, and equally weights the axes' Brier scores. Candidates $.01, .1, 1, 10, 100$ select $\lambda = .1$; intercept penalty is 10^{-6}. For new scenes or sources, only the corresponding effect is fitted using the original 48 calibration labels; test outcomes neither fit coefficients nor select regularization. Logistic-minus-ordering Brier differences are $-.0015$ and $-.0078$, with crossed intervals $[-.0260, .0145]$ and $[-.0321, .0195]$ (20,000 draws, seed 360914). Two-sided scene sign-flip values are $.875$ and $.0625$.

A separate retrospective anchor-budget analysis exhausts the 1/28/70/28/1 subsets at budgets 0/2/4/6/8. Mean Brier scores are .1761/.1325/.0962/.0674/.0459 on the same 144 cells. Only budgets 0 and 8 reproduce the original prospective forecasts; intermediate budgets are sensitivity analyses.

State readouts. Full-state diagnostics on 64 joint-grid states give 64/64 cue classifications under raw or centered leave-one-scene-out nearest centroids. Centering uses each held-out scene's unlabeled four-state batch, so it is transductive. Timing states give 63/64. These diagnostics and the geometry-augmented forecasts separate cue decodability from predictive utility.

F Supporting Tests in Other Policies

F.1 Long-hist visual history

The released Match Color cross-attention diffusion checkpoint uses 16 images, 121 proprioceptive observations, 16 output actions, and eight DDIM steps. Identities 0, 2, and 3 match all 24 mapped source/test roles in development and are used for confirmation; identity 1 matches 0/8. Six new scenes each use conflicts $0 \leftarrow 2$, $2 \leftarrow 3$, and $3 \leftarrow 0$, giving 18 units.

At query 7, replacement edits past donor images at the same temporal offsets, leaving current input (slot 15) and proprioception unchanged. Cue-window slots are 12–13 in scenes 0/1, 11–13 in 2/5, and 11–12 in 3/4. The post-cue slot is 14, except 13 in scenes 3/4. Selected slot j corresponds to sampled query $j - 8$ and ages to $j - (q - 7)$ at later query q. Boundary physical states, prefixes, normalized inputs, and random states match their references. First placement and released-stable placement agree.

Table 18: Long-hist confirmation and follow-up controls. Donor/recipient means agreement with source/test cue. Six pairs have a third cue frame. Matched windows use the post-cue edit's full-input norm and, for sparse edits, changed-value count.

Condition	Total	Donor	Recipient	Other	Unresolved
Recipient control	18	0	18	0	0
First cue frame / rank 0	18	10	6	1	1
Complete cue window	18	18	0	0	0
First post-cue frame	18	0	18	0	0
Single cue frame, rank 1	18	5	13	0	0
Single cue frame, rank 2	6	0	6	0	0
Dense window, matched L_2	18	0	17	1	0
Sparse window, matched $L_0 + L_2$	18	3	14	1	0

Onset and full-window edits yield 10/18 and 18/18 donor choices, respectively; post-cue edits retain the recipient choice in all 18 pairs. Median full-input RMS differences are .0547/.0800/.0133 for onset/full window/post-cue.

The follow-up covers all 42 single cue frames and both matched-dose arms (78 trajectories); rank 0 reproduces the onset row. Maximum norm mismatch is 1.21×10^{-5} (tolerance 2×10^{-5}); sparse edits also match the changed-value count. Single frames yield 15 donor, 25 recipient, one other, and one unresolved choice; retrospective inspection finds an effective frame in 10/18 pairs. A compatibility rerun retained the design after three observed outcomes.

Earlier four-scene modality development gave 12/12 donor matches for visual history or visual-plus-proprioceptive history, and 0/12 for proprioception alone or recipient control.

F.2 Executed prefixes in MemoryVLA

The ShellGameTouch checkpoint uses 224-pixel RGB, four executed actions per query, guidance 1.5, and ten denoising steps. Cue A is shown at queries 0/4 and covered at step 5; actions 0–7 come from a same-scene no-write run with conflicting cue B. Current proposals are logged but not executed. Perceptual and cognitive consolidation are disabled and bank counts stay zero. At query 8 the queue is empty, and native control resumes through step 90.

After a two-scene pilot, four new scenes supply 24 conflicts within 60 confirmation trajectories. Choices follow the executed-prefix cue in 22/24, the seen cue in 1/24, and the third cup in 1/24, with no unresolved primary outcome. The 22/24 prefix count holds at 7/8/9 cm, dwell two, and last committed choice. Native and no-write controls both give 11/12 primary matches; all 18 own-prefix replays across pilot/confirmation reproduce. Cut geometry, image, and first autonomous action chunk match the prefix reference in all 24 conflicts.

Enabling native bank writes under the same 24 executed prefixes gives 1/24 seen-cue matches, just as no-write, with zero paired gains or losses. Primary choices agree in 23/24 pairs; prefix matches are

21/24 versus 22/24, and unresolved counts are one versus zero. Alternative readouts also show no gain in seen-cue matching.

F.3 GMP HISTORY TRANSPORT

Four sources and four test scenes give 64 joint-history transports, 16 local controls, four once-erased controls, and four sustained-read-off controls. Each erased/read-off execution is evaluated against four balanced cue labels. Joint history yields 64/64 source-cue placements versus 4/16 balanced no-history matches. First and released-stable placement agree. Both 28-trajectory component cohorts match their boundary references and reproduce native-anchor actions. Fixed bin locations couple identity and position.

G AN ILLUSTRATIVE STATE REPLACEMENT

In Figure 6, source S1 (seed `2852617598`) retains the red-sphere cue at queries 0–3 and is captured before source query 9. Test T2 (seed `1264166751`) saw a blue T-shape but did not retain cue-query writes; its paired branches split before query 7. Source and test clocks are separate. Both test branches match inputs, world, prefix, and RNG; replacement copies the complete source tensor. At step 15, control first reaches the blue cube at 4.665 cm, whereas replacement reaches the red sphere at 2.623 cm. S1 itself later reaches the red cube at step 18 (3.983 cm).

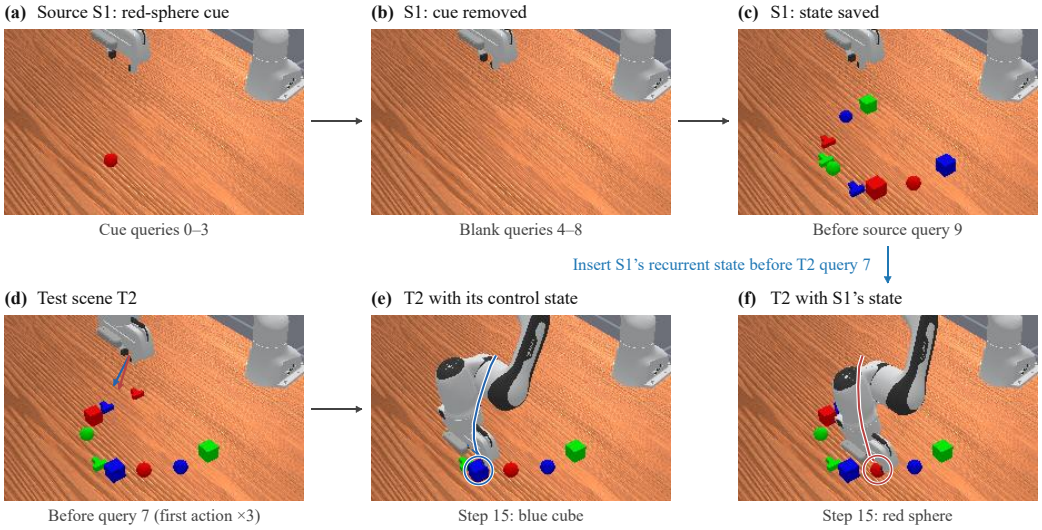

Figure 6: Paired state replacement from source S1 to test T2. (a–c) Capture of the red-sphere state. (d–f) Replacement changes T2's first stationary choice at step 15 from blue cube to red sphere. Arrows show the first displacement at 3× scale; solid paths end at choice and dashed paths continue to step 25. The control's final distance is 5.316 cm against the 5 cm threshold.

Source and test blank inputs and boundary worlds come from their respective same-seed no-write references; source cue writes are retained. Capture/replacement state hashes, boundary observations/worlds, and paired prefix actions match. The paired action replays reproduce the recorded actions, choices, event times, and geometry (error zero at a 10^{-5} m tolerance).