LET ME CHANGE MY MIND: DECISION-RELEVANT EVIDENCE REUSE FOR BRAIN–AI INTERFACES

Anonymous authors
Paper under double-blind review

ABSTRACT

An adaptive interface can mistake a changed goal for a changed recording channel and preserve evidence about an action the user no longer wants. We study which distinctions must be diagnosed for the current action and which pre-change evidence remains valid. In finite, stable diagnostic episodes, we characterize action identifiability and additional diagnostic cost under a transported belief, and give a policy attaining the leading cost coefficient. An exact noisy-channel example shows when retained recording information helps and when inexpensive goal confirmation removes that advantage. We evaluate Decision-Aware Reuse and Probing (DARP) in exact finite models and a closed-loop categorical simulator. An updated DARP aggregate reports 95.97% coverage, 3.25% executed-command error, 3.43 simulated diagnostic seconds, and 9.53 correct commands per minute. The supplied earlier Goal-IG run reports 94.82%, 3.28%, 4.23 seconds, and 8.64 correct commands per minute; this cross-run comparison is descriptive. The source package reproduces an earlier DARP-KL run, whose paired and stress results are kept separate from the updated aggregate. We also report completed human-interface experiments as a separate aggregate comparison.

1 INTRODUCTION

A user who changes their mind has not necessarily become harder to decode. Yet an adaptive interface sees observations, not a label identifying whether the user's goal or the recording mechanism changed. If the interface explains a new goal by modifying its observation model, it can make the old action easier to predict and the new action harder to express. Conversely, failing to update a genuinely shifted recording model can obscure an unchanged intention. These two errors require different interventions.

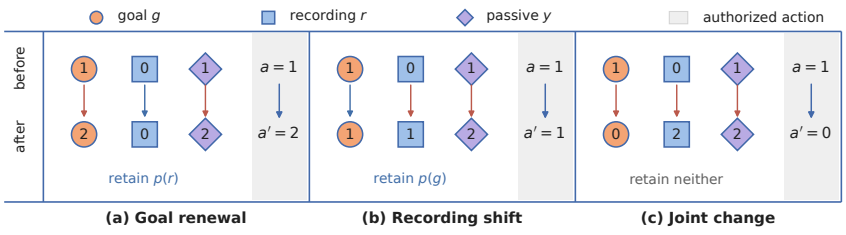

Figure 1: One observation need not identify which latent component changed. For $k = 4$, $u = 0$, $y = (g - u + r) \bmod 4$ and $a = g$, all three post-change states yield $y' = 2$, yet authorized actions and retainable belief marginals differ. This categorical illustration is not EEG or human data.

In neural interfaces, user and machine adaptation occur together (Abu-Rmileh et al., 2019), while AI assistance combines uncertain decoded signals with task-level decisions (Lee et al., 2025). Shared autonomy has long represented uncertain goals (Javdani et al., 2015), including the possibility that a user changes goals after the assistant becomes confident (Jeon et al., 2020). The unresolved issue

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.

here is how a controller should allocate new measurements when goal change and recording change explain the same observations, but only some previous measurements remain informative.

We separate three operations. A *transition model* transports valid information across a change. A *diagnostic policy* resolves distinctions that could alter the current action. An *execution rule* acts only when the evidence and authorization permit it. Reidentifying every latent variable wastes queries; carrying old goal evidence through a goal revision can support an obsolete action. An operational policy should therefore be compared at its attained execution risk, coverage, and interaction time, including the time spent on actions it withholds.

Decision identification can require less information than hypothesis identification (Golovin et al., 2010; Javdani et al., 2014). We characterize how a valid transported belief changes the *additional* cost of that decision after a latent change. A finite noisy-channel construction separates selective transport from full reset and shows that cheap, independent goal confirmation closes the gap. The comparison identifies when memory has decision value rather than treating retention itself as beneficial.

Decision-Aware Reuse and Probing (DARP) instantiates mechanism-specific belief transitions and residual-information allocation. We prove a leading-order cost result for a pilot–allocate–verify policy in stable, correctly specified episodes, then test a practical rolling variant in exact finite models and a completed longitudinal closed-loop simulation. The updated DARP aggregate is numerically competitive with Goal-IG, while the separately executable earlier DARP-KL run shows a risk–time tradeoff against global resetting. Their run-level records are not linked in the supplied package, so we do not transfer paired intervals or mechanism conclusions across versions. Completed human-interface experiments are reported separately.

2 Related Work

Decision-directed acquisition identifies equivalence classes or decision regions under noisy tests (Golovin et al., 2010; Javdani et al., 2014; Chen et al., 2017). Controlled sensing with unequal query costs and standard change-of-measure arguments inform our cost analysis (Nitinawarat et al., 2012; Nitinawarat and Veeravalli, 2013; Kaufmann et al., 2016; Garivier and Kaufmann, 2016); Bayesian changepoint methods motivate uncertainty over change time (Adams and MacKay, 2007). Our question is which components of a *post-change* joint belief remain informative for the current action.

Test-time adaptation updates predictors under distribution shift (Wang et al., 2021; 2022), whereas co-adaptive BCIs, AI copilots, and shared autonomy motivate the goal-revision setting (Abu-Rmileh et al., 2019; Lee et al., 2025; Jeon et al., 2020). Our categorical online-pseudo control is not a reproduction of these neural methods. We report executed risk together with coverage, as in selective prediction (Geifman and El-Yaniv, 2019).

3 Model: Actions, Diagnostic Access, and Valid Memory

Let the finite latent state be $h_t = (g_t, r_t) \in \mathcal{H} = \mathcal{G} \times \mathcal{R}$, with current goal g_t and recording state r_t. The observable task context is u_t. A goal is a task objective or context-dependent command rule, not an unrestricted psychological state. For a specified loss with a unique minimizer,

$$a^\star(h,u) = \arg\min_{a \in \mathcal{A}} \ell(a, g; u), \qquad \mathrm{Alt}_u(h) = \{h' : a^\star(h',u) \neq a^\star(h,u)\}. \tag{1}$$

Recording state changes the observation law, not the definition of the authorized action. States may be piecewise stationary, but (h,u) is fixed during each theoretical diagnostic episode. A user cancellation or goal revision ends the episode; the controller must not continue certifying an obsolete command.

Protocols, not free labels. A finite library Σ contains diagnostic protocols with positive bounded costs c_σ. A protocol may combine observations, goal confirmation, and instructed calibration.

For $\sigma = (q_1, \ldots, q_m)$,

$$P_h^\sigma(y_{1:m}) = \prod_{j=1}^{m} P_h(y_j \mid q_{1:j}, y_{1:j-1}), \qquad d_\sigma(h,h') = D_{\mathrm{KL}}(P_h^\sigma \| P_{h'}^\sigma). \tag{2}$$

For the theorems, outcomes have finite alphabets and common support, log likelihood ratios are bounded, and protocols return to a standardized starting condition. Repetitions are conditionally independent. Order can matter *within* a protocol; uncontrolled carryover *between* protocols is excluded. Costs include instructions and recovery. Optimality is relative to this fixed library.

Selective transition of a joint belief. Let $\pi^-(h)$ be the posterior before a modeled change and $K_J(h' \mid h)$ the transition for change mask $J \subseteq \{G, R\}$. Valid reuse is the predictive transport

$$\pi_J^+(h') = \sum_h K_J(h' \mid h)\pi^-(h). \tag{3}$$

For a goal renewal independent of the old goal, with recording unchanged,

$$K_G((g', r') \mid (g, r)) = q_G(g')\mathbf{1}\{r' = r\} \quad \Longrightarrow \quad \pi_G^+(g', r') = q_G(g') \sum_g \pi^-(g, r'). \tag{4}$$

Thus the recording marginal survives, including information originally obtained jointly with the old goal. This is more precise than deleting every observation that mentions a changed factor. A dependent goal transition must instead use its actual K_G; equation (4) cannot be assumed for an arbitrary change of mind.

When the mask and change time are unknown, the controller filters their joint distribution, with transitions followed by a single likelihood update. Fixed prior weights on the change masks still allow observations to update the posterior probability of each explanation; they do not make the change mechanism known. The implementation does not propagate uncertainty over its fitted observation accuracies or transition rates. It does not detect a change from an observation and then multiply the same observation in again as independent confirmation. Factor lineage is useful for implementation, but a retained factor is not valid merely because it agrees with the preferred hypothesis. Appendix D states the conditioning requirements.

Figure 2 previews the rolling procedure used in the exact-model and archived DARP-KL evaluations; Section 5 specifies its query allocation and stopping rule.

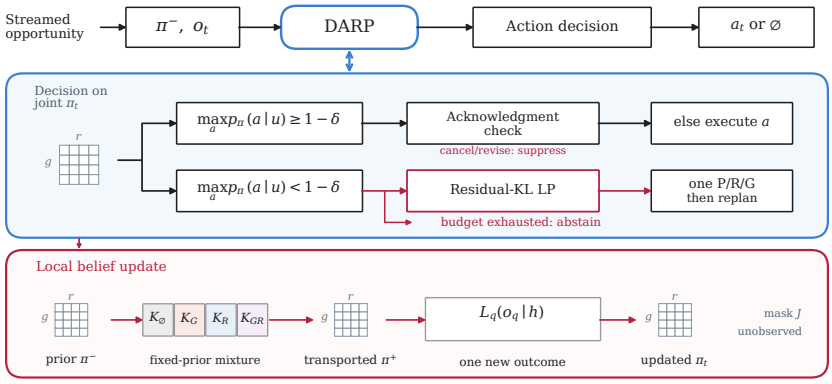

Figure 2: Rolling residual-information controller used in the exact-model and archived DARP-KL evaluations. The middle tier branches on action support: above $1 - \delta$, an acknowledged cancellation or revision check precedes execution; otherwise the residual-KL program (Eq. (11)) selects one feasible query and replans, or abstains if the budget is exhausted. The lower tier expands fixed-prior transport over four unobserved change masks and one likelihood update per acquired outcome. Grids are symbolic; the gate does not certify actuation-time risk.

4 What Must Be Identified, and at What Cost?

Define $h \equiv_\Sigma h'$ when $P_h^\sigma = P_{h'}^\sigma$ for all σ. Only states equivalent in this *observational* sense may be merged without changing diagnostic planning. Merely sharing the current action is not sufficient: such states can assign different values to the next query.

Proposition 1 (Action identifiability)**.** *In the finite stable model, action selection with arbitrarily small error and finite expected cost, uniformly over h, is possible if and only if $h \equiv_\Sigma h' \Rightarrow a^\star(h,u) = a^\star(h',u)$.*

The necessity follows from identical adaptive-history laws; sufficiency follows by repeatedly testing the finitely many separating protocols. This is a decision-class specialization of standard identification reasoning, not a claim to have originated decision classes. The complete proof is in Appendix C.2.

4.1 A CONDITIONAL LOWER BOUND

Let $E = e$ be valid retained evidence and π_e its correctly specified current-state posterior. A policy is conditionally Bayes-δ-correct when it eventually returns an action and

$$\sum_h \pi_e(h) \Pr_h(\widehat{a} \neq a^\star(h,u) \mid e) \leq \delta. \tag{5}$$

This is not a uniform statewise guarantee after every retained history. Write N_σ for additional protocol counts and $C = \sum_\sigma c_\sigma N_\sigma$. Define $\epsilon_e(h) = \min\{1, \delta/\pi_e(h)\}$, taking $\epsilon_e(h) = 1$ for zero prior mass. Let

$$k_e(h,h') = \begin{cases} \mathrm{kl}(1-\epsilon_e(h), \epsilon_e(h')), & \epsilon_e(h) + \epsilon_e(h') < 1, \\ 0, & \text{otherwise.} \end{cases} \tag{6}$$

Here kl is binary relative entropy, and states with zero posterior mass carry no conditional Bayesian guarantee.

Theorem 1 (Additional diagnostic cost)**.** *Every conditionally Bayes-δ-correct policy obeys*

$$\mathbb{E}_h[C \mid e] \;\geq\; \min_{n \geq 0} \left\{ \sum_\sigma c_\sigma n_\sigma : \sum_\sigma n_\sigma d_\sigma(h,h') \geq k_e(h,h') \quad \forall h' \in \mathrm{Alt}_u(h) \right\}. \tag{7}$$

The stopped-history change of measure gives each constraint; minimizing expected cost gives the bound. Reuse enters through a valid conditional posterior. Subtracting a realized old log likelihood ratio from a uniform frequentist bound would be a different, generally unjustified argument.

4.2 LEADING-ORDER ACHIEVABILITY AND ITS QUALIFICATIONS

Set $L = \log(1/\delta)$. After merging observationally equivalent hypotheses, assume full-support retained posteriors with $\pi_{e_\delta}(h) \geq p_0 > 0$ at the state under analysis, and, for every remaining j,

$$-\log \pi_{e_\delta}(j)/L \longrightarrow \kappa(j) < \infty. \tag{8}$$

Define the residual-information complexity

$$V_h(\kappa) = \min_{x \geq 0} \left\{ \sum_\sigma c_\sigma x_\sigma : \sum_\sigma x_\sigma d_\sigma(h,j) \geq [1-\kappa(j)]_+ \quad \forall j \in \mathrm{Alt}_u(h) \right\}. \tag{9}$$

Theorem 2 (Stable-episode optimality)**.** *Under the finite-support, bounded-likelihood, repeatability, and separation assumptions above, every conditionally Bayes-δ-correct policy has* $\liminf_{\delta \downarrow 0} \mathbb{E}_h[C \mid e_\delta]/L \geq V_h(\kappa)$*. A pilot–allocate–verify version of* DARP *is conditionally Bayes-δ-correct and attains* $\mathbb{E}_h[C \mid e_\delta] = LV_h(\kappa) + o(L)$*.*

The construction uses $O(L^{2/3})$ pilot repetitions, an information allocation with $L^{3/4}$ slack, and a posterior action-class stopping rule, with a separating fallback (Appendix C.4). At $\delta = 0.025$, the three-query pilot alone needs nine queries, exceeding the evaluated eight-query budget; this asymptotic theorem does not certify the rolling heuristic.

For *fixed* full-support evidence, all $\kappa(j) = 0$: reuse does not improve the leading asymptotic coefficient. Finite-confidence savings remain possible. Improving the coefficient requires evidence strength that scales with L. Earlier acquisition of that evidence is not free; the comparison concerns *additional* cost after change. Detection delay and finite segment length are separate costs, not hidden in the theorem.

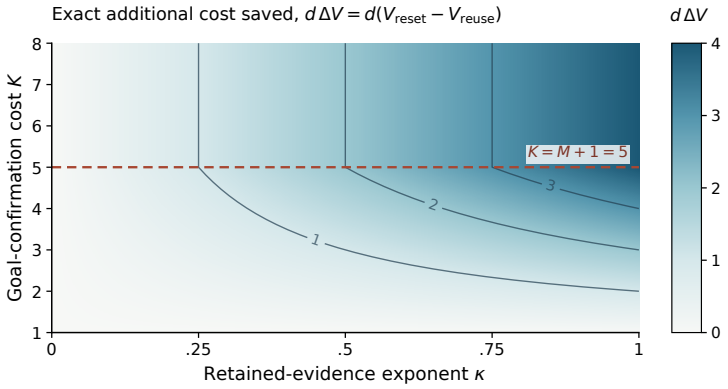

Figure 3: Analytic diagnostic-cost saving for $M = 4$. Color and contours show $d(V_{\text{reset}} - V_{\text{reuse}}) = \kappa \min\{4, K - 1\}$, where d is binary-test KL divergence. Saving vanishes at $\kappa = 0$ or $K = 1$ and saturates at $K = 5$ (dashed line). The leading-order coefficient assumes valid retained evidence and excludes its acquisition cost.

4.3 An exact reuse benefit—and a cheap-query ceiling

Consider $g, r \in \{0, 1\}$, action $a = g$, and three binary-symmetric tests with independent noise probability $\nu \in (0, 1/2)$: a parity observation of $g \oplus r$ costing 1, a recording calibration of r costing $M > 0$, and independent goal confirmation of g costing $K \geq 1$. Let $d = (1 - 2\nu)\log((1 - \nu)/\nu)$. After a goal renewal, take g uniform and $\pi_e(r = 1) = \delta^\kappa/(1 + \delta^\kappa)$, with true state $(0, 0)$.

Theorem 3 (Reset separation with confirmation access)**.** *For this model the leading additional-cost coefficient is exactly*

$$V_{\text{reuse}}(\kappa) = \frac{1 + [1 - \kappa]_+ \min\{M, K - 1\}}{d}. \tag{10}$$

A global-reset test certified under its uniform reset prior has coefficient $V_{\text{reset}} = \min\{M + 1, K\}/d$. *For* $\kappa \geq 1$*, their ratio is* $\min\{M + 1, K\}$.

This is a concrete separation, not a universal effect size. If $K = 1$, the ratio is one: cheap goal confirmation makes preserving recording information unnecessary for this decision. Large savings require both costly recalibration and the absence of an equally cheap goal-only shortcut. The independent confirmation channel must therefore be a measured resource, not an oracle hidden in the method. Appendix D.2 derives the formula by a two-constraint linear program.

The model also has complementary observations: with independent uniform g, r, neither a parity nor a recording query alone informs g, but at $\nu = 0.15$ their pair provides $1 - H_2(0.255) = 0.181$ bits, where H_2 is binary entropy. A one-step goal-information score can miss this pair; costed superiority over direct confirmation does not follow.

Figure 3 maps where valid retained evidence lowers additional diagnostic cost and where cheap confirmation removes that advantage; the plotted values are analytic coefficients rather than empirical performance.

5 Decision-Aware Reuse and Probing

DARP maintains a joint belief over goals, recording hypotheses, change masks, and (when needed) run lengths. A neural implementation could derive a finite observation model from an EEGNet encoder (Lawhern et al., 2018); no such encoder is trained or evaluated here. Development data determine likelihood tables and protocol costs. The exact-model theory treats these as known, whereas the virtual-agent implementation estimates query accuracy from finite calibration.

For current posterior π, candidate $\widehat{h}$, and $\beta = \log((|\mathcal{H}| - 1)/\delta)$, the planner computes

$$\min_{n \geq 0} \sum_{\sigma} c_\sigma n_\sigma \quad \text{s.t.} \quad \sum_{\sigma} n_\sigma d_\sigma(\widehat{h}, j) \geq \left[\beta - \log \frac{\pi(\widehat{h})}{\pi(j)} \right]_+, \quad j \in \mathrm{Alt}_u(\widehat{h}). \tag{11}$$

The idealized theorem implementation adds its specified slack and executes the allocated batch. A rolling implementation may recompute after every protocol; its empirical efficiency is an evaluation question, not a consequence of Theorem 2. Posterior mass is aggregated only for stopping:

$$p_\pi(a \mid u) = \sum_{h : a^\star(h,u) = a} \pi(h), \qquad \max_a p_\pi(a \mid u) \geq 1 - \delta. \tag{12}$$

A finite interaction budget introduces abstention. Under a correctly specified model, the posterior threshold controls error for the goal state represented by that posterior at query termination. Execution-time risk additionally depends on unobserved changes during commitment; a final authorization check suppresses acknowledged cancellations but cannot detect every silent revision. The unrestricted stopping-time cost theorem does not describe the budgeted policy's coverage.

Appendix B gives executable-step pseudocode for the tested cancellation gate, not an execution-time posterior-risk certificate. Appendix L.10 gives a separate propagation calculation; incorporating its hazard-adjusted threshold into a controller would require a new implementation and evaluation.

No automatic benefit from sequencing. If primitive outcomes are independent, repeatable, and have additive costs, the KL divergence of a sequence is just the sum of its primitive divergences. Adding such sequences cannot improve the allocation optimum. An ordering advantage requires an actual order-dependent response law or cost, estimated without evaluator-label leakage. Appendix D.3 proves this negative result. Likewise, a current action-class posterior is not generally a sufficient statistic for choosing probes.

Robustness is not automatic. If a budgeted interaction's true and modeled joint trajectory laws differ by at most Δ in total variation, and true execution coverage is $c > 0$, then its executed-command risk is at most $\min\{1, \delta + \Delta/c\}$ (Appendix E). This quantifies why a nominal posterior threshold cannot certify an uncalibrated neural interface. It does not provide a measured value of Δ.

6 Finite-Model and Closed-Loop Evaluation

The executable evaluation has two components: exact finite-model computations and closed-loop virtual-agent simulations. Author-reported human-interface aggregates appear separately in Appendix K; the prospective protocol in Appendix I is a distinct design. We report attained execution risk, coverage, diagnostic time, and throughput together. The operational target is lower total interaction time at a prespecified executed-risk limit and minimum coverage; the tested virtual policies do not all reach a common feasible operating point, so their observed tradeoffs remain separate.

6.1 Exact finite-model evidence

We retain 71 exact evaluations of the four-state noisy-channel model, enumerating all reachable outcomes within eight queries. With noise $\nu = 0.15$, costs $(1, 4, 8)$, valid recording probability 0.99, and posterior risk threshold 0.05, rolling DARP, State-IG, and an exact belief-state planning reference tie at 99.58% coverage, 3.96% executed-command error, and 2.67 cost units. The reference minimizes query cost plus 40 times abstention probability within its stopping constraint; it is not a fixed-coverage cost optimum. The reset ablation uses 13.98 units at 75.33% coverage; these are not matched-coverage results. A cheap direct confirmation query removes the advantage. An overconfident retained probability of 0.99 when the true probability is 0.80 raises DARP's error to 21.81%. These examples establish conditional benefits and failure cases, not a general ordering of controllers (Appendix H).

6.2 A closed-loop virtual-agent generator

Three independent seed panels (173, 947, 2027) each contain 96 main, 24 shifted-sensor, and 32 alternative-sensor virtual agents. After calibration, five evaluated visits each have two categorical

Table 1: Completed closed-loop experiment aggregates. The updated DARP row was supplied by the study team; the remaining rows come from the archived executable run with three panels of 96 agents. Coverage is attained, not matched.

Controller	Coverage (%)	Error (%)	Diag. sec.	Correct/min
Global-reset-KL	91.84	1.16	5.41	7.26
DARP (updated)	95.97	3.25	3.43	9.53
DARP-KL (archived)	92.93	4.06	4.13	8.53
State-IG	94.05	9.44	3.32	9.16
Goal-IG	94.82	3.28	4.23	8.64
Always-confirm	93.77	0.56	5.91	6.98

tasks ($k = 2, 4$) and 48 opportunities per task. Visit indices are simulated, not elapsed weeks; no EEG or gaze samples enter the generator.

The hidden state is a goal g and recording offset r. Passive symbols follow $(g - u + r) \bmod k$; noisy recording and goal queries are calibrated from finite samples. Six-opportunity segments randomize sham, goal-only, recording-only, and joint changes. A contextual suggestion is correct with probability 0.50 although the controller assigns it 0.70 prior mass. Policies cannot observe intervention labels or ground truth. Actions, correction time, fatigue, skill, silent revision, and acknowledged cancellation affect later interactions, so policies induce different histories. The ten stress families vary access, change rate, dependence, priors, feedback, visits, and sensor quality. Across eight controllers, core and stress runs contain 2,119,680 opportunities, not independent human observations. Parameters were fixed before outcome inspection, but not preregistered or fitted to people; Appendix L specifies the generator.

6.3 Comparators, operating points, and uncertainty

DARP-KL is the practical rolling residual-information heuristic, not the pilot–allocate–verify policy of Theorem 2. Global-reset-KL uses the same query planner but discards both mechanisms when its renewal mixture activates. State-IG and Goal-IG change only the query objective, using joint-state or goal entropy reduction per cost. Always-confirm uses the same noisy confirmation channel and requires it before executing. Frozen, Online-pseudo, and No-reset-KL are additional diagnostic controls, not modern neural-adaptation implementations. All policies receive the same initial calibration, permitted observations, cost schedule, and cancellation gate. None receives a true goal, shift label, fatigue state, or future random outcome.

In the archived executable run, 24 separate development agents are used to select among five risk thresholds, choosing the strictest one reaching 90% coverage when available. Thresholds are then locked. The online-pseudo control cannot reach that target; its low coverage is reported as observed. Test outcomes use *attained* coverage, and a nominal posterior threshold is not treated as a measured risk guarantee. Equal-weight means of per-agent ratios and 2,000 paired agent-bootstrap replicates are available for this archived run, stratified by independent seed panel and virtual site. Its intervals quantify variation among virtual-agent trajectories under the stipulated generator; they do not quantify uncertainty for the separately supplied updated DARP aggregate.

Risk is error per executed command; coverage is executions per scheduled opportunity. Diagnostic time includes abstentions. Throughput includes amortized initial calibration, queries, execution, intertrial intervals, and corrections. Outdated-action rate additionally uses all eligible opportunities after a changed authorized action:

$$R_{\text{old}} = \Pr(S = 1, \widehat{a} = a_{\text{old}} \mid a_{\text{new}} \neq a_{\text{old}}). \tag{13}$$

Correct completion by a fixed deadline counts nonexecution and cancellation as noncompletion. An acknowledged cancellation must suppress execution; this is checked as a software invariant, not interpreted as a behavioral finding.

6.4 Nominal risk, coverage, and time

Study-team DARP aggregates are 95.97% coverage and 3.25% executed error (Table 1). Reported differences to rounded Goal-IG imply 3.43 diagnostic seconds and 9.53 correct commands per minute. The differences are $+1.15$ coverage points, -0.03 error points, -0.80 seconds (18.91%), and $+0.89$ correct commands per minute (10.30%); no paired interval or updated run-level linkage was supplied.

The archived executable DARP-KL run is separate: its 1.280-second saving against global resetting accompanies 2.894 points more error, and Goal-IG has a paired 0.772-point error advantage. Its paired intervals, stress outcomes, and 2.5% stopping threshold apply only to that run (Appendix L). The canceled-commit bookkeeping correction cannot explain the updated coverage or error; without the updated event and policy records, the cause of the change is not identified.

The archived nominal seed-173 panel regenerates 368,640 event records at locked thresholds; executed, error, and query totals match the grouped CSV. Figure 4 aligns opportunities around programmed changes.

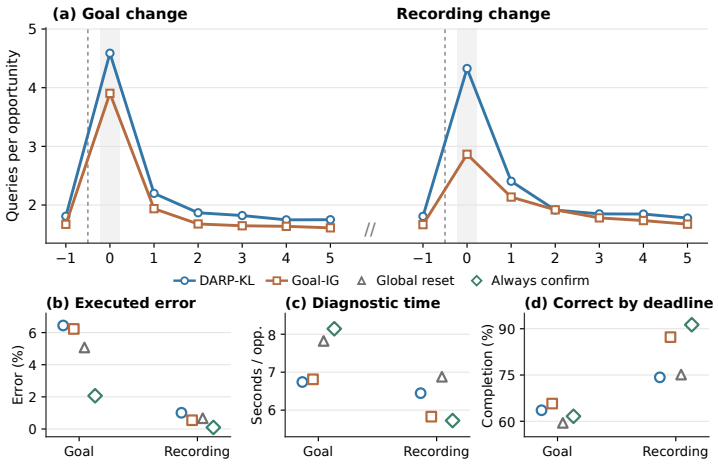

Figure 4: Completed archived experiment: response to programmed changes. (a) Seed-173 queries per opportunity for DARP-KL and Goal-IG. Goal and recording blocks are separate (-1 before, 0–5 after; 1,708 and 1,715 opportunities per position). (b–d) First-two-postchange equal-agent means over three archived seed panels: executed error, diagnostic time, and correct deadline completion.

6.5 Reversal, mechanism checks, and transfer boundaries

The proposed adaptation–intent reversal requires recording-shift improvement *and* goal-shift deterioration, not just an interaction

$$\mathcal{I} = (R_{\text{online},G} - R_{\text{frozen},G}) - (R_{\text{online},R} - R_{\text{frozen},R}). \tag{14}$$

The simulator does *not* reproduce this reversal. Online-pseudo has lower conditional risk than Frozen in both change strata but much lower overall coverage (38.58% versus 90.85%). A matched-coverage benefit is untested. Severe symbolic recording permutations and the limited passive-only comparator preclude broader EEG-adaptation conclusions.

Paired-memory clones fix a valid recording mechanism and retain or erase initial evidence; retention reduces diagnostic cost from 4.42 to 3.71 seconds. This tests known-valid reuse, not how deployment identifies it. With matched primitive content, order has no effect under independent stability; when goals change between primitives, a final goal query lowers error. This stipulated-model recency test establishes neither an optimal planner nor a human order effect.

Archived shifted- and alternative-sensor virtual tests show no uniform advantage over Goal-IG or global resetting; they are separate from human EEG and gaze aggregates. Against an exploratory matched-horizon reference, removing confirmation gives DARP-KL 76.91% coverage; rapid revision

yields 80.14% coverage and 6.26% error, and correlated lapses 87.13% coverage and 6.75% error. Figure 5 shows these appendix operating points; none were rerun for updated DARP. Appendix L gives thresholds, stress results, order tests, missing visits, and baselines.

6.6 Separate human-interface aggregate report

The completed human-interface experiments report a 90% coverage setting (Table 5). Their summary gives DARP 5.8% goal-shift error versus 11.3% for DCT-FR, 8.1% joint-shift error versus 15.4%, and 9.4 versus 8.1 correct commands per minute. DCT-FR is the change-detection/full-reset comparator. Participant records, execution denominators, and analysis scripts are outside the supplied package, so uncertainty and mechanism attribution are not recalculated here.

7 Discussion and Limitations

Correct retention reduces diagnosis in the exact model and paired intervention. Updated DARP improves coverage and time against archived Goal-IG, but its 0.03-point error difference lacks paired uncertainty. Updated old-action and deadline outcomes are unavailable; archived DARP-KL trails Goal-IG (7.31% versus 6.17% old actions; 84.18% versus 86.46% completion). The update does not establish better goal-shift handling.

The theorem assumes valid transport and excludes its validation cost. The archived controller uses 32-outcome plug-in calibration, hidden changes, and a 0.70 contextual prior despite actual suggestion accuracy of 0.50. Silent post-query changes also risk error (Appendix L.10). The categorical simulator is not fitted to the human cohort; recording permutations, fatigue, learning, and latency are stipulated. Virtual-agent precision does not establish biological realism; sensor variants differ from the EEG and gaze comparison. EEGNet, CBraMod, LaBraM, and public-data tests remain prospective (Lawhern et al., 2018; Wang et al., 2025; Jiang et al., 2024; Jayaram and Barachant, 2018).

8 Conclusion

Valid transport of pre-change evidence can reduce information needed for a current action; the exact model specifies when. Updated DARP is numerically competitive with Goal-IG, while the executable earlier run shows that lower diagnostic cost can accompany more errors under hidden change and model mismatch. Goal-directed information gain remains a strong comparator. Author-reported human aggregates favor DARP in a separate setting, but record-level linkage is needed to test whether the mechanism explains them.

Ethics Statement

For the human experiment described by the authors, participants were healthy, BCI-naive, and had no history of neurological disorders. Before participation, each participant received detailed information about the protocol and provided written informed consent in accordance with the Declaration of Helsinki. The authors report that an institutional review board reviewed and approved the experimental environment and protocols. The supplied materials do not identify the board, approval reference, or which aggregate cohorts in Appendix K were covered by that protocol. The virtual-agent simulations use synthetic evaluator labels hidden from the controller; cancellation suppression in the simulator is a software invariant rather than a human behavioral finding.

Reproducibility Statement

Appendices C–E provide assumptions and proofs, and Appendix H documents 71 exact evaluations. Appendix L describes the archived virtual-agent study, which reproduces the earlier DARP-KL results; updated event-level records for Table 1 were unavailable, so its uncertainty and stress analyses were not independently regenerated. Appendix K reports separate author-provided human aggregates without participant-level records or analysis scripts, while Appendix I gives a distinct prospective validation protocol.

AI USE STATEMENT

An AI assistant helped draft and revise the text, inspect literature, develop and check finite-model arguments, implement and run the exact and virtual-agent computations, inspect their outputs, and prepare the LaTeX source. The computational experiments use synthetic states and observations; the separately supplied human aggregate values were not generated or reconstructed by these programs. Numerical checks are not formal proof verification or independent expert review. The authors are responsible for the manuscript's claims and for disclosing this use in the submission form.

REFERENCES

Abbas Abu-Rmileh, Elad Zakkay, Lior Shmuelof, and Oren Shriki. Co-adaptive Training Improves Efficacy of a Multi-Day EEG-Based Motor Imagery BCI Training. *Frontiers in Human Neuroscience, 13:362*, 2019. Publisher record.

Ryan P. Adams and David J. C. MacKay. Bayesian Online Changepoint Detection. *arXiv preprint arXiv:0710.3742*, 2007. arXiv:0710.3742.

Yuxin Chen, Hamed Hassani, and Andreas Krause. Near-optimal Bayesian Active Learning with Correlated and Noisy Tests. *Proceedings of AISTATS, PMLR 54:223–231*, 2017. Publisher record.

Aurélien Garivier and Emilie Kaufmann. Optimal Best Arm Identification with Fixed Confidence. *Proceedings of COLT, PMLR 49:998–1027*, 2016. Publisher record.

Yonatan Geifman and Ran El-Yaniv. SelectiveNet: A Deep Neural Network with an Integrated Reject Option. *Proceedings of ICML, PMLR 97:2151–2159*, 2019. Publisher record.

Daniel Golovin, Andreas Krause, and Debajyoti Ray. Near-Optimal Bayesian Active Learning with Noisy Observations. *Advances in Neural Information Processing Systems, 23*, 2010. Publisher record.

Shervin Javdani, Yuxin Chen, Amin Karbasi, Andreas Krause, Drew Bagnell, and Siddhartha Srinivasa. Near Optimal Bayesian Active Learning for Decision Making. *Proceedings of AISTATS, PMLR 33:430–438*, 2014. Publisher record.

Shervin Javdani, Siddhartha Srinivasa, and Andrew Bagnell. Shared Autonomy via Hindsight Optimization. *Proceedings of Robotics: Science and Systems*, 2015. Publisher record.

Vinay Jayaram and Alexandre Barachant. MOABB: Trustworthy algorithm benchmarking for BCIs. *Journal of Neural Engineering, 15:066011*, 2018. arXiv:1805.06427.

Hong Jun Jeon, Dylan P. Losey, and Dorsa Sadigh. Shared Autonomy with Learned Latent Actions. *Proceedings of Robotics: Science and Systems*, 2020. Publisher record.

Wei-Bang Jiang, Li-Ming Zhao, and Bao-Liang Lu. Large Brain Model for Learning Generic Representations with Tremendous EEG Data in BCI. *International Conference on Learning Representations*, 2024. arXiv:2405.18765.

Emilie Kaufmann, Olivier Cappé, and Aurélien Garivier. On the Complexity of Best-Arm Identification in Multi-Armed Bandit Models. *Journal of Machine Learning Research, 17(1):1–42*, 2016. Publisher record.

Vernon J. Lawhern, Amelia J. Solon, Nicholas R. Waytowich, Stephen M. Gordon, Chou P. Hung, and Brent J. Lance. EEGNet: A Compact Convolutional Network for EEG-based Brain-Computer Interfaces. *Journal of Neural Engineering, 15:056013*, 2018. arXiv:1611.08024.

Johannes Y. Lee, Sangjoon Lee, Abhishek Mishra, Xu Yan, Brandon McMahan, Brent Gaisford, Charles Kobashigawa, Mike Qu, Chang Xie, and Jonathan C. Kao. Brain–computer interface control with artificial intelligence copilots. *Nature Machine Intelligence, 7:1510–1523*, 2025. Publisher record.

Sirin Nitinawarat, George Atia, and Venugopal V. Veeravalli. Controlled Sensing for Multihypothesis Testing. *arXiv preprint arXiv:1205.0858*, 2012. arXiv:1205.0858.

Sirin Nitinawarat and Venugopal V. Veeravalli. Controlled Sensing for Sequential Multihypothesis Testing with Controlled Markovian Observations and Non-Uniform Control Cost. *arXiv preprint arXiv:1310.1844*, 2013. arXiv:1310.1844.

Dequan Wang, Evan Shelhamer, Shaoteng Liu, Bruno Olshausen, and Trevor Darrell. Tent: Fully Test-Time Adaptation by Entropy Minimization. *International Conference on Learning Representations*, 2021. arXiv:2006.10726.

Jiquan Wang, Sha Zhao, Zhiling Luo, Yangxuan Zhou, Haiteng Jiang, Shijian Li, Tao Li, and Gang Pan. CBraMod: A Criss-Cross Brain Foundation Model for EEG Decoding. *International Conference on Learning Representations*, 2025. arXiv:2412.07236.

Qin Wang, Olga Fink, Luc Van Gool, and Dengxin Dai. Continual Test-Time Domain Adaptation. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022. arXiv:2203.13591.

A Archived Virtual-Agent Stress Operating Points

Figure 5 shows attained coverage and conditional error for four archived stress settings on separate aligned axes. The matched-horizon reference was added as an exploratory control after the primary run; no policy is interpolated to another policy's coverage.

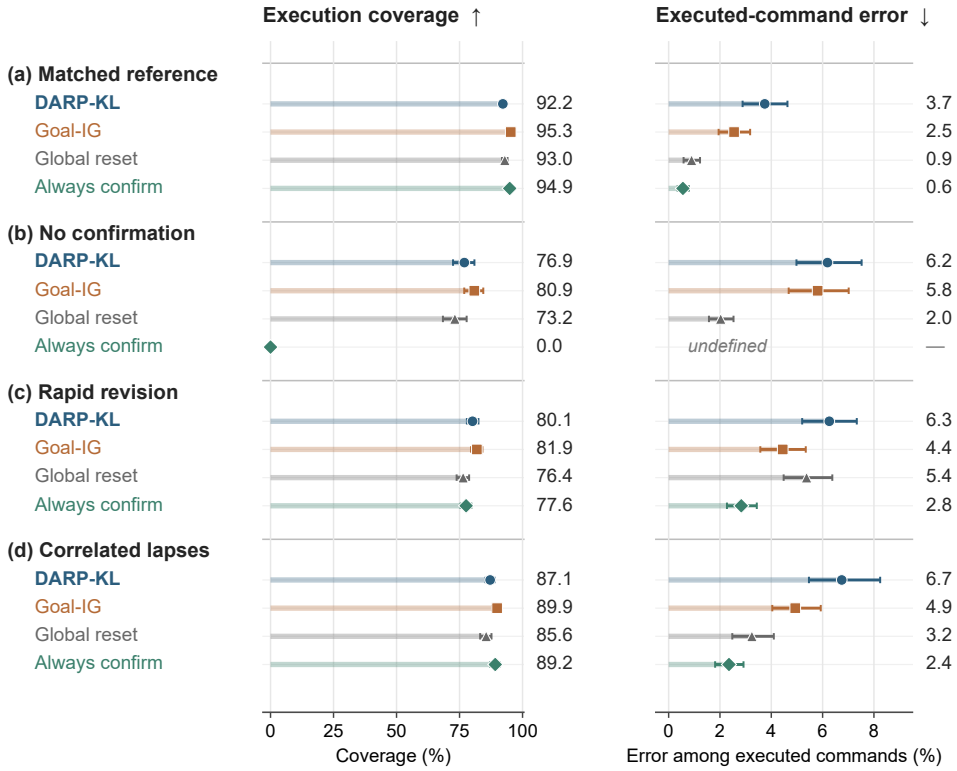

Figure 5: Completed archived closed-loop stress experiments across four settings. Each row is one locked-threshold policy; pale stems end at attained means, and dark whiskers show separate marginal 95% agent-bootstrap intervals for coverage and error among executed commands (32 agent clusters per cell). Coverage spans 0–100%: under no confirmation, Always-confirm has 0% coverage and undefined conditional error because no command was executed.

B Rolling Controller Pseudocode

Algorithm 1 DARP: transition, diagnose, then execute

Require: Joint belief, change kernels, protocol laws and costs, budget, risk threshold

Apply the predictive mixture of mechanism-specific kernels (3).
Update with each newly acquired observation exactly once.
while no action satisfies (12) **do**
    if the user cancels or revises the objective **then**
        return cancellation; discard the pending action.
    else if the interaction budget is exhausted **then**
        return abstention.
    end if
    Choose the next protocol from the allocation (11), with exploration.
    Observe its outcome, charge its full cost, and update the joint belief.
end while
if an acknowledged cancellation or revision arrives before or during commitment **then**
    return cancellation; discard the pending action.
end if
Execute the supported action after commitment; record feedback and total time.

C Assumptions and Complete Proofs

C.1 The finite experiment and the scope of the probability statements

Fix a context u for an identification episode. Let $H = |\mathcal{H}|$ and $S = |\Sigma|$ be finite. For each protocol σ, outcomes belong to a finite set $\mathcal{Y}_\sigma$, and $P_h^\sigma(y) > 0$ for every h and every outcome in that protocol's common support. Since the sets are finite and the model is fixed, a constant $B < \infty$ bounds the absolute log likelihood ratios. Costs satisfy $0 < c_{\min} \le c_\sigma \le c_{\max} < \infty$. All logarithms are natural.

A protocol is an atomic experiment for this analysis: it may contain internally dependent observations, but its completed outcome is drawn from P_h^σ independently of earlier completed protocols conditional on h. The protocol-selection rule can depend arbitrarily on the observed history and on internal randomization. It cannot depend on the hidden state except through that history. A protocol's price is fixed, not an unmodeled state-dependent random duration.

An idealized episode has no latent transition while diagnosis runs. Applying its result to a finite stationary segment requires a separate assessment of whether another change occurs before stopping. The manuscript does not condition on an unobserved future event of stability and then claim an unconditional guarantee. If a goal revision is explicitly communicated, the controller closes the pending episode.

For asymptotic achievability, observationally equivalent hypotheses are merged. By Proposition 1, each merged class has one action in the fixed context. Their posterior masses are added, and their protocol laws are identical. Distinct resulting classes are separated by at least one protocol, including classes that happen to share the action. The posterior scaling condition (8) is imposed on *all* resulting hypotheses, not only action-changing alternatives. In particular, all log inverse prior masses are $O(L)$. This is used to bound costs after an erroneous pilot estimate.

For $H = 1$, or when every supported hypothesis prescribes the same action, the policy returns that action immediately. Expressions involving $\log(H - 1)$ are used only in the nontrivial case. The finite lower bound allows zero prior mass with the convention in the main text; the asymptotic theorem assumes full support.

C.2 Proof of Proposition 1

Suppose $h \equiv_\Sigma h'$. Conditional on any common history, the same randomized policy has the same distribution over its next protocol under the two states. The next outcome also has the same law. Induction proves equality of the complete adaptive-history distributions, including the stopping time

and final action. If $a^\star(h,u) \neq a^\star(h',u)$, the two correct-action events are disjoint under this common distribution. Their probabilities sum to at most one. At least one state therefore has error at least $1/2$.

Conversely, cycle through every protocol a fixed number m of times. For each pair of distinct observational classes, the vector of protocol outcome laws differs. Its log likelihood ratio, summed over a complete cycle, has positive mean under the first class and is bounded. A concentration bound therefore makes the probability of selecting an incorrect observational class at most $C_1 \exp(-C_2 m)$ for constants determined by the finite model. A finite union bound over pairs yields a finite m for every target error $\delta > 0$. Selecting that class's action is then uniformly δ-correct. The expected cost is finite. Identifying every observational class here proves existence; it is not asserted to be the cheapest way to identify an action. □

C.3 Proof of Theorem 1

Let $A_h = \{\widehat{a} = a^\star(h,u)\}$. Conditional Bayes correctness implies

$$\Pr_h(A_h^c \mid e) \leq \epsilon_e(h).$$

For every $j \in \mathrm{Alt}_u(h)$, returning $a^\star(h,u)$ is an error under j, hence $\Pr_j(A_h \mid e) \leq \epsilon_e(j)$. When $\epsilon_e(h) + \epsilon_e(j) < 1$, data processing through the indicator of A_h gives

$$D_{\mathrm{KL}}(\mathbb{P}_h^{\mathrm{hist}} \| \mathbb{P}_j^{\mathrm{hist}} \mid e) \geq \mathrm{kl}(1-\epsilon_e(h), \epsilon_e(j)). \tag{15}$$

The minimum over the permitted binary probabilities occurs at these boundary values: the first probability is above the second, binary relative entropy increases in its first argument and decreases in its second on that ordered region. If the regions overlap, zero is the valid general lower bound used in (6).

For a finite expected stopping cost, bounded log likelihood ratios permit summation over stopped observations. Alternatively, first truncate at a deterministic horizon and then pass to the limit. Adaptive selection kernels cancel in the likelihood ratio because the same policy is used under both states conditional on a history. The chain rule gives

$$D_{\mathrm{KL}}(\mathbb{P}_h^{\mathrm{hist}} \| \mathbb{P}_j^{\mathrm{hist}} \mid e) = \sum_{\sigma \in \Sigma} \mathbb{E}_h[N_\sigma \mid e]\, d_\sigma(h,j). \tag{16}$$

If expected cost is infinite the desired inequality is automatic. Otherwise, combining (15) and (16) makes the expected count vector feasible in (7). Its objective value is $\mathbb{E}_h[C \mid e]$, proving the bound. This is the standard adaptive change-of-measure technique specialized to action-changing alternatives and conditional Bayesian risk (Kaufmann et al., 2016). □

Why the conditioning matters. If one instead requires $\Pr_h(\widehat{a} \neq a^\star(h,u) \mid e) \leq \delta$ for every state after *every* admissible e, the corresponding pairwise target is $\mathrm{kl}(1-\delta, \delta)$. Strong prior odds do not remove that statewise requirement. Theorem 1 uses (5), which can allocate more error to low-posterior states. These are distinct decision guarantees, and neither should be described as the other.

C.4 Proof of Theorem 2: the lower bound

Write $\pi_\delta = \pi_{e_\delta}$. Since $\pi_\delta(h) \geq p_0$, $\epsilon_e(h) = O(\delta)$. If $\kappa(j) < 1$, then

$$\epsilon_e(j) = \exp\{-[1-\kappa(j)]L + o(L)\}, \qquad k_e(h,j)/L \longrightarrow 1 - \kappa(j).$$

If $\kappa(j) > 1$, the constraint is eventually vacuous. At the boundary $\kappa(j) = 1$, the normalized binary divergence tends to zero, whether $\pi_\delta(j)/\delta$ tends to zero, a finite constant, or infinity subexponentially. Thus

$$k_e(h,j)/L \longrightarrow [1-\kappa(j)]_+. \tag{17}$$

The finite allocation problem is continuous in its nonnegative right-hand side. To see this directly, separation provides a strictly positive common direction: assigning equal positive mass to every protocol gives strictly positive divergence against every relevant alternative. Adding a sufficiently small multiple of this direction compensates for a sufficiently small change in all constraints. Positive costs also bound minimizers on compact sets for bounded right-hand sides. Scaling (7) by L and using (17) proves the lower limit in Theorem 2.

C.5 Proof of Theorem 2: a constructive upper bound

Set $m = \lceil L^{2/3} \rceil$ and $s = L^{3/4}$. The following construction uses exact model likelihoods. It is provided to establish attainability rather than to claim optimal finite-sample constants.

Pilot stage. First check whether the posterior action threshold already holds. Otherwise execute m full cycles through the protocol library, and select the maximum-likelihood observational class $\widehat{h}$ using only the new likelihoods, not the prior. Distinct classes are separated; bounded increments and a finite union bound give

$$\Pr_h(\widehat{h} \neq h \mid e_\delta) \leq C_1 e^{-C_2 m}. \tag{18}$$

The pilot costs $O(L^{2/3}) = o(L)$. Its outcomes are then included once in a posterior π_{p}. They change each log posterior odds by at most $O(m)$, uniformly over pilot histories.

Allocation stage. For every $j \in \mathrm{Alt}_u(\widehat{h})$, set

$$b_j = \left[\beta - \log \frac{\pi_{\mathrm{p}}(\widehat{h})}{\pi_{\mathrm{p}}(j)} \right]_+ + s, \qquad \beta = \log((H-1)/\delta).$$

Solve the nonnegative linear program minimizing $\sum_\sigma c_\sigma n_\sigma$ subject to $\sum_\sigma n_\sigma d_\sigma(\widehat{h}, j) \geq b_j$. Round each count up to the next integer and execute that allocation. Because all prior log masses are $O(L)$ and pilot increments are $o(L)$, every target is $O(L)$, even on incorrect-pilot histories. Separation gives a common feasible allocation of $O(L)$ cost; hence the chosen allocation also has $O(L)$ observations. Rounding adds at most $\sum_\sigma c_\sigma$.

On $\widehat{h} = h$, uniformly over pilot histories,

$$b_j / L \longrightarrow [1 - \kappa(j)]_+.$$

The same continuity argument as above gives allocated cost $LV_h(\kappa) + o(L)$. Conditional on the pilot, the future observations are independent. Under h, the mean final log posterior odds against each relevant j exceeds β by at least s. With $O(L)$ bounded increments, Hoeffding's inequality yields a failure probability at most

$$C_3 \exp(-C_4 s^2 / L) = C_3 \exp(-C_4 L^{1/2}) \tag{19}$$

for the event that some relevant log odds remain below β. If all exceed β, then

$$\sum_{j \in \mathrm{Alt}_u(h)} \pi(j) \leq (H-1) e^{-\beta} \pi(h) \leq \delta,$$

so the posterior action threshold holds.

Verify and fall back. If the threshold does not hold, cycle through all protocols until it does. At the start of fallback, every log posterior odds is bounded below by $-C_5 L$, uniformly over the pilot/allocation histories: both their length and their increments are bounded as established above. For any true class h, each complete cycle has positive mean log likelihood ratio against every other class. After $t \geq C_6 L$ cycles, bounded-increment concentration bounds the probability that the threshold still fails by $C_7 e^{-C_8 t}$ after increasing C_6 if necessary. Summing this tail proves a uniform $O(L)$ conditional expected fallback cost.

Under the state h with $\pi_\delta(h) \geq p_0$, fallback is entered with probability bounded by (18) plus (19). Its contribution to expected cost is therefore $o(L)$. Incorrect-pilot allocation cost is also $O(L) e^{-\Omega(L^{2/3})} = o(L)$. Together with the pilot and correct-pilot allocation this proves

$$\mathbb{E}_h[C \mid e_\delta] \leq LV_h(\kappa) + o(L).$$

At every terminal history the selected action has posterior error at most δ. The fallback stops almost surely under every supported state, so averaging the terminal posterior errors gives conditional Bayes correctness. Combining with the lower bound proves equality to leading order. □

What is and is not asymptotically free. The pilot spends subleading cost, not zero cost, identifying a useful observational class. Historical calibration can have cost comparable to L or larger. That cost is a common sunk cost only when both policies genuinely received it earlier. The theorem cannot justify comparing a heavily pretrained method with an uncalibrated baseline while omitting the former's initial expenditure.

D Selective Transport and Exact Diagnostic Separations

D.1 Why transport preserves the correct information

Let $\mathcal{F}^-$ denote all modeled pre-change information. If a known change mask obeys

$$\Pr(H^+ = h' \mid H^- = h, \mathcal{F}^-, J) = K_J(h' \mid h),$$

then the law of total probability gives (3). No assumption that the old goal and recording state are posterior-independent is needed. With the independent renewal kernel in (4), summing the old joint posterior over goals preserves exactly its recording marginal.

For example, let the old joint posterior, with rows indexed by g and columns by r, be

$$\pi^- = \begin{pmatrix} 0.45 & 0.05 \\ 0.15 & 0.35 \end{pmatrix}.$$

A uniform independent goal renewal gives

$$\pi^+_G = \begin{pmatrix} 0.30 & 0.20 \\ 0.30 & 0.20 \end{pmatrix},$$

not a uniform joint posterior and not the old conditional distribution of recording state given the most likely goal. The recording marginal remains $(0.60, 0.40)$. These are arithmetic examples, not measured biological probabilities.

With an uncertain mask, allow its hazard to depend on the old state and history. The predictive joint law is

$$\bar{\pi}(h', J) = \sum_h \lambda_J(h, \mathcal{F}^-) K_J(h' \mid h, \mathcal{F}^-) \pi^-(h), \tag{20}$$

where $\sum_J \lambda_J(h, \mathcal{F}^-) = 1$. The new observation y then produces

$$\pi(h', J \mid y) \propto P^\sigma_{h'}(y) \bar{\pi}(h', J). \tag{21}$$

Run length can be included in the state, as in Bayesian changepoint inference (Adams and MacKay, 2007). A detector that uses y to generate an updated mask distribution must not subsequently treat this updated distribution as a fresh prior and multiply the same evidence independently again. Equations (20)–(21) give the joint calculation.

Model dependence of reuse. A known-command EEG observation need not be physiologically independent of the private goal. Instruction compliance, fatigue, or attention can depend on both. Such dependencies belong in P^σ_h or in an expanded state. If a goal change also changes the effective recording/encoding mechanism, the transition is joint rather than goal-only. The kernel identity remains valid only for the correct state and transition description. It is not evidence that a chosen goal-only kernel is true.

D.2 Proof of Theorem 3

For the binary-symmetric tests, the KL divergence between opposite underlying bits is

$$d = (1-\nu)\log\frac{1-\nu}{\nu} + \nu\log\frac{\nu}{1-\nu} = (1-2\nu)\log\frac{1-\nu}{\nu} > 0.$$

Under true state $(0,0)$, only $(1,0)$ and $(1,1)$ require another action. The divergence matrix, with columns for parity, recording calibration, and goal confirmation, is

$$D_{\text{action}} = \begin{pmatrix} d & 0 & d \\ 0 & d & d \end{pmatrix}. \tag{22}$$

The corresponding posterior exponents are 0 and κ. Let $b = [1-\kappa]_+ \in [0,1]$ and set $u = dx_P$, $v = dx_R$, and $z = dx_G$. Equation (9) becomes

$$dV = \min_{u,v,z\geq 0}\{u + Mv + Kz : u + z \geq 1,\ v + z \geq b\}.$$

For fixed z, the smallest feasible values are $u = (1-z)_+$ and $v = (b-z)_+$. Because $K \geq 1$, minimizing

$$f(z) = (1-z)_+ + M(b-z)_+ + Kz$$

over $z \geq 0$ reduces to its breakpoints $0, b, 1$. The values at 0 and b are $1 + Mb$ and $1 + (K-1)b$, respectively; the value K at 1 is no smaller than the latter. Therefore

$$dV = 1 + b\min\{M, K-1\},$$

which proves (10). A global-reset controller that requires Bayes-δ correctness under its uniform reset prior sets $b = 1$. The formula and the ratio follow. The comparison concerns two retained-information states and their respective Bayesian certification contracts. It is not a lower bound for every policy under the unchanged, informative posterior, nor does discarding evidence change the actual physical state. Theorem 2 supplies an attaining sequence of tests, so this is a sharp leading-cost characterization, not only a lower bound. □

Exact numerical check. For $\nu = 0.2$, $d = 0.6\log 4 \approx 0.8317766$. Table 2 reports *normalized analytical coefficients* dV, not observed diagnostic times. The included script constructs the test distributions, computes their KL matrix, solves the allocation program, and checks this formula over a parameter grid.

Table 2: Exact model values of dV with $M = 4$. These follow algebraically from Theorem 3.

Confirmation cost	$\kappa = 0$	$\kappa = 0.25$	$\kappa = 0.5$	$\kappa = 1$
$K = 1$	1	1	1	1
$K = 2$	2	1.75	1.5	1
$K = 8$	5	4	3	1

Full-state information can be unnecessary. A separate elementary construction observes g and r through independent binary-symmetric tests with divergences d_G and d_R and unit costs. With a fixed full-support prior, action identification $a = g$ has coefficient $1/d_G$. Full-state identification requires coefficients $1/d_G + 1/d_R$: the alternatives changing only g and only r impose the two separate constraints. As the recording test becomes uninformative, $d_R \downarrow 0$, their ratio diverges. This is an illustrative consequence of decision-class identification, not a new general active-learning theorem.

D.3 When order cannot help

Proposition 2 (No sequence advantage with additive independent primitives)**.** *Suppose every primitive q is independently repeatable from the same standard state. Each sequence σ consists of these primitives with $P_h^\sigma = \bigotimes_{q\in\sigma} P_h^q$ and $c_\sigma = \sum_{q\in\sigma} c_q$, counting multiplicity. If all primitives are already available, adding such sequences leaves the allocation values in* (7) *and* (9) *unchanged.*

Proof. Product additivity gives $d_\sigma(h,j) = \sum_{q\in\sigma} d_q(h,j)$. Replace each unit of sequence allocation with its primitive counts. The replacement has exactly the same cost and exactly the same divergence against every alternative, so it is feasible in the primitive problem. Conversely, the augmented library contains the primitive library. Both inequalities between their optimal values follow. □

An order-dependent protocol can still be useful, but its benefit must come from a change in its joint law or effective cost. Merely grouping queries, calling them sequence-aware, or treating a noisy response as a noiseless label does not establish such a benefit.

D.4 Analytic diagnostic-cost phase diagram

The exact phase relationship is shown in Figure 3 in the main text.

E Selective Execution under Model Error

Lemma 1 (Trajectory perturbation bound)**.** *Consider a common budgeted policy and authorized-action definition under a modeled joint law Q and a true joint law P over the latent state, history, and terminal decision. Suppose $\mathrm{TV}(P, Q) \le \Delta$. Under Q, the policy executes only at histories with conditional action error at most δ. If $P(S = 1) = c > 0$, then*

$$P(\widehat{a} \neq a^\star \mid S = 1) \le \min\{1, \delta + \Delta/c\}.$$

Proof. Define $f = \mathbf{1}\{S = 1, \widehat{a} \neq a^\star\} - \delta \mathbf{1}\{S = 1\}$. Its range lies in $[-\delta, 1 - \delta]$, an interval of length one. The posterior stopping property gives $\mathbb{E}_Q f \le 0$. The variational characterization of total variation gives $\mathbb{E}_P f \le \mathbb{E}_Q f + \Delta \le \Delta$. Rearranging and dividing by c proves the result. □

If initial-state/posterior error is at most η in total variation, the conditional outcome-kernel error is at most ε at every matched history, and at most $N_{\max}$ protocols are used, a sequential coupling gives $\Delta \le \eta + N_{\max}\varepsilon$. An additional probability ξ of an unmodeled latent change can be added, yielding $\Delta \le \min\{1, \eta + N_{\max}\varepsilon + \xi\}$. The virtual-agent results do not supply empirical bounds on these discrepancies.

This bound also explains a coverage caveat: a small joint-model discrepancy can have a larger conditional effect when execution is rare. The result concerns the full joint law, not merely agreement between two marginal feature histograms.

F Proposed Human-Interface Validation Protocol

F.1 Study population, workload, and timing

The design specifies 96 healthy adults, equally allocated to two sites, with visits at weeks 0, 1, 2, 4, 6, and 8. The first visit supplies task familiarization and initial calibration; the remaining visits evaluate adaptation. These are design choices, not a record of completed recruitment. No age distribution, demographic balance, exclusion count, attrition rate, equipment manufacturer, institution, or approval identifier is invented.

The primary crossover compares five policies on both tasks. A candidate schedule is ten policy–task blocks per evaluation visit, each containing 24 command opportunities, with breaks and a hard session-duration limit set by the eventual ethics protocol. Factor assignments are balanced across participants and visits rather than requiring every factorial cell to be exhausted inside each short block. Blocks contain stable segments and explicit intervention opportunities; minimum dwell length must exceed the allowed diagnostic interaction budget. A pilot is needed to determine comfortable epoch duration, number of repetitions, and recovery intervals. A design that exceeds the fatigue or duration budget must reduce block length or distribute conditions across visits, not silently remove slow participants.

The primary policies, including always-confirm, are crossed within participants. More extensive ablations can be allocated through a randomized balanced incomplete-block design, with each ablation paired with DARP within the same assigned participants. This avoids an implausible requirement that every participant complete every possible baseline, prior stratum, task, and diagnostic variant in a single visit. Such ablation cohorts must report their own sample sizes; the full $n = 96$ cannot automatically be attached to them.

Use 32-channel EEG sampled at 250 Hz as an explicit engineering specification, with hardware-dependent antialiasing and reference choices documented before collection. Candidate task-specific preprocessing is causal filtering within an approved acquisition pipeline, motor-imagery spectral content, and ERP epochs linked to stimulus onsets. Filter settings, artifact thresholds, baseline correction, and epoch lengths must be fixed from development data. Their exact physiological adequacy is not established here. EEGNet is a defensible common encoder architecture (Lawhern et al., 2018), not a completed trained model supplied with this paper.

F.2 Factorial interventions and noncircular intent labels

Private goals are recorded by a separate evaluator channel before each action window or stable goal segment. The controller does not receive that record. A goal-change manipulation offers the participant a new task objective; execution correctness is measured against their most recent authorized objective, not against the old instruction or the AI's recommendation. Goal recording itself can alter behavior, and its timing and interaction cost belong in the protocol.

Recording perturbations are applied digitally after acquisition: bounded channel gains, controlled mixing, and predefined masking. Perturbation families are held out at the appropriate evaluation level to test whether a method recognizes new recording regimes rather than a short list of memorized transforms. These interventions isolate changes to recorded observations but do not establish robustness to every physical electrode shift, contact change, or neurophysiological adaptation.

The AI prior is a contextual suggestion distribution with full support. A candidate main setting assigns probability 0.7 to either the registered objective or an alternative, distributing the remaining mass across other goals. Additional strengths can be tested as sensitivity conditions. These values are proposed manipulations, not measured priors. Alignment is randomized relative to the *pre-suggestion* registered goal. If the participant subsequently changes their goal, the new authorization is recorded. The intention-to-treat alignment label remains the randomized assignment; actual later alignment is a separate descriptive variable, not a basis for silently reassigning conditions.

All policies can use the same paid confirmation channel. In healthy adults, a separate manual response is experimentally convenient, but it is not an acceptable unstated assumption about a person with severe motor impairment. Always-confirm control and confirmation-unavailable conditions are therefore essential. Confirmation prompts should not display the AI's preferred answer as a default, and nonresponse is not agreement.

F.3 Observation models and evidence accounting

A common encoder maps each epoch to features. Development-only quantization or a finite mixture yields protocol-specific likelihood tables over (g, r), with positive smoothing to avoid assigning zero probability to unobserved outcomes. Model choices must be fixed before held-out outcomes are inspected. The finite model is an approximation; calibration curves, goodness-of-fit diagnostics, and sensitivity to state-space resolution are needed.

Initial calibration data may be shared identically across policies to isolate the controller comparison. Its equivalent per-policy acquisition cost is charged to each policy's end-to-end accounting. Later observations are policy-specific and cannot be transferred from a condition in which a competing controller paid for them. Policies maintain separate adaptive states across their blocks. A source encoder can remain fixed during a diagnostic phase while goal/recording beliefs adapt; changing the encoder requires a new likelihood model or explicit recomputation, not reuse of likelihood ratios defined under obsolete features.

The predictive kernels must permit no change, goal change, recording change, and joint change. A perfectly informed change mask is an oracle ablation, not the default method. A method-specific reset event is not a ground-truth transition label. The same distinction applies to run-length truncation: discarding old branches for computation is an approximation that needs an error or sensitivity analysis.

F.4 Strong baselines and isolation of mechanisms

The frozen policy leaves its initial observation model unchanged. The online-adaptation policy uses a documented confidence-filtered pseudo-label/entropy update. Candidate learning rates, update masks, buffer lengths, and confidence thresholds are tuned only on development data. Calling this a Tent-inspired comparator does not claim faithful reproduction of Tent's original vision experiment (Wang et al., 2021).

DCT-FR has change detection, action-class beliefs, the complete diagnostic protocol library, and a cost-sensitive query score selected on validation participants. Its candidate scores can include expected reduction in cross-action posterior disagreement and worst-alternative KL gain per unit cost.

Following its detected change, it performs a global reset and full recalibration. Because DCT-FR differs in both planning and resets, its contrast with DARP is not the isolated causal effect of reuse.

The *global-reset* DARP variant retains the same planner and stopping rule and changes only the transition/reset behavior. A *full-state* variant retains selective transport but values complete latent identification. A *single-interaction* variant restricts the protocol library without changing evidence access or primitive prices. Matched-attention and matched-label-budget comparisons test whether improvements arise only because one policy asks more questions. The always-confirm policy tests the cheap-query ceiling of Theorem 3 directly.

Required negative controls include incorrect goal-only resets under joint changes, shortened stable segments, noisy/nonindependent confirmations, and a recording state outside the fitted hypothesis set. No outcomes are asserted for these controls. A result table populated in their absence would not establish the proposed mechanism.

F.5 PARTICIPANT-DISJOINT ESTIMATION AND LIVE-POLICY EVALUATION

Six outer folds each contain 16 held-out participants balanced across sites. For a fold, the other 80 participants supply visit-1 data only: 64 for fitting and 16 for validation. Visit 1 includes a common scripted acquisition and diagnostic protocol, not a confirmatory policy comparison. This permits fitting observation laws, measuring candidate interaction costs, and choosing protocol and stopping settings without access to visits 2–6. All fold-specific population models and hyperparameters must be frozen after the visit-1 acquisition phase and before any confirmatory visit-2 interaction. Enrollment and scheduling must respect that ordering. Within a held-out fold, only the allowed, time-charged individual calibration subset is available from its participants; other visit-1 labels are not used to tune that fold's policy.

The evaluation phase consists of counterbalanced live blocks during visits 2–6. No setting may be selected using the later held-out trajectories. A participant may supply development data for another fold, but never for their own population model; no evaluation-phase data feed population fitting in any fold. The shared development data create dependence between fitted fold-specific controllers. The participant bootstrap is therefore an evaluation-variability summary conditional on these fitted controllers, not a full estimate of retraining uncertainty. A full training-and-evaluation uncertainty analysis would need a nested refitting design. This phase-separated specification avoids claiming that a model fitted after a live interaction could have generated that interaction.

Counterbalance policy and task order. Reset task contexts, not the user, between blocks; allow recovery from the preceding condition. Statistical adjustment cannot guarantee removal of all learning/fatigue carryover. Recorded replay can test an observation model, but cannot by itself estimate the throughput of a different closed-loop policy whose queries and actions change subsequent inputs.

F.6 COVERAGE, TIME, AND PAIRED EFFECTS

Let M_i be command opportunities, S_{ij} execution, and Z_{ij} an erroneous execution for participant i. Compute

$$\widehat{c}_i = \frac{\sum_j S_{ij}}{M_i}, \qquad \widehat{R}_i = \frac{\sum_j Z_{ij}}{\sum_j S_{ij}}, \qquad \widehat{T}_i = \frac{N_{i,\text{completed correct}}}{T_{i,\text{all}}}.$$

If no action is executed, conditional risk is undefined, not zero; that participant's zero coverage remains reported. Corrected attempts are counted once upon completion, with the earlier error and all elapsed time retained. Deliberate cancellations and timeouts remain in the opportunity denominator. Report performance by participant, task, change condition, and prior assignment before standardizing with prespecified weights.

Operating thresholds are locked using validation data to *target* 90% coverage. Actual test coverage must be reported even when it differs from 90%. It is not legitimate to adjust thresholds on held-out correctness outcomes or retrospectively delete executions from a closed-loop trajectory and call the result a live policy. A descriptive risk–coverage curve is a separate analysis.

If randomized mixtures of two live operating regimes are used, their coverage and risk obey

$$c_\lambda = \lambda c_1 + (1-\lambda)c_2, \qquad R_\lambda = \frac{\lambda c_1 R_1 + (1-\lambda)c_2 R_2}{c_\lambda}.$$

Conditional risks do not interpolate linearly in general. Mixture weights must be specified from development data for a prospective policy. Matching aggregate coverage also does not ensure equal coverage in each shift stratum; those counts and risks must be shown.

The proposed primary effect is a participant-paired goal-shift risk difference, DCT-FR minus DARP, standardized over task and prior strata. A future protocol must fix its endpoint, strata, and analysis plan before data collection. A site-stratified participant bootstrap would resample whole participants with their sessions, conditions, and paired policy outcomes. Secondary interactions and multiple endpoints require separately reported uncertainty and a declared multiplicity policy.

The paired uncertainty depends on within-participant correlation and repeated trials. Means and participant counts alone cannot determine it.

F.7 External validation and transfer

The external EEG specification has 24 participants at a third site; the separate gaze specification has 32. The query planner, state-transition logic, and selection procedure are fixed before collecting those cohorts. Interface-specific observation models can be calibrated using only permitted resources, with their cost reported. Goals and recording hypotheses need an explicit correspondence across modalities; they are not assumed to share feature distributions.

These designs could test transfer of the procedure, but a single additional site would not establish broad site generalization. A gaze cohort would test interface behavior rather than the physiological assumptions of an EEG observation model.

G Evidence Layers and Executable Supplement

The manuscript uses existing literature for its background, analytical proofs for finite-model statements, and executable computations for its numerical results. The exact benchmark evaluates specified finite observation laws; the longitudinal evaluation uses virtual agents under stipulated categorical sensing and feedback. The human-interface protocol above is prospective and has no outcome data.

Executed checks. The supplement's `verify_model.py` constructs finite Bernoulli observation laws, computes the action-changing KL matrix, and solves the residual allocation programs independently of the closed-form expression. It tests the selective-transition marginal identity, product-KL sequence equivalence, asymptotic conditional thresholds on a grid, and the finite-outcome selective-risk perturbation inequality. `verification_report.json` records the actual outcomes. Passing finite checks is useful for detecting implementation/algebra errors; it is not a proof assistant or a substitute for the analytical proofs.

What is not supplied. The package contains no fitted EEG or gaze decoder, raw neural recording, human event log, participant bootstrap, or completed cross-site study. Testing the proposed human protocol requires suitable data with goal and authorization labels.

Novelty boundaries. Decision equivalence, controlled sensing, Bayesian filtering, and selective prediction are the foundations of the analysis. The contribution is the conditional-cost formulation tied to valid mechanism transport, the explicit reset/confirmation separation, and the closed-loop test of a rolling controller under hidden change.

H Exact Computational Benchmark

H.1 Generative model and state separation

States are ordered $(0,0),(0,1),(1,0),(1,1)$. Test P measures $g \oplus r$, test R measures r, and test G measures g. Each binary answer is flipped independently with probability $\nu = 0.15$. The actual

initial distribution is

$$p_q(g,r) = \tfrac{1}{2}\big(q\mathbf{1}\{r=0\} + (1-q)\mathbf{1}\{r=1\}\big).$$

The controller instead initializes with $p_{\widehat{q}}$. Valid memory means $q = \widehat{q} = 0.99$, no information means $q = \widehat{q} = 0.5$, and retention misspecification holds $\widehat{q} = 0.99$ while varying q. Full reset sets its own $\widehat{q}$ to 0.5; the evaluator does not change the physical distribution to make the baseline easier or harder. The prior stores information acquired before this episode; all computational costs below are *additional* costs. Human throughput must also account for initial acquisition.

The theoretical passive ambiguity is checked directly: states $(1,0)$ and $(0,1)$ give exactly the same Bernoulli law under P and different authorized goals. This does not make independent confirmation free; all three tests have the declared costs $(1, M, K)$.

H.2 Exact belief-state reference

For integer signed counts $z = (z_P, z_R, z_G)$, where z_j is the number of ones minus zeros from test j, the posterior is

$$\pi_z(g,r) \;\propto\; p_{\widehat{q}}(g,r)\exp\{\eta[z_P(g \oplus r) + z_R r + z_G g]\}, \qquad \eta = \log\frac{1-\nu}{\nu}. \tag{23}$$

Query-specific factors independent of (g,r) cancel. Thus signed counts and remaining horizon are sufficient for exact dynamic programming; no binning of the posterior is needed.

Write $e(\pi) = 1 - \max_a \sum_{h:a^\star(h)=a} \pi(h)$, $A = 40$, and let $V_t(\pi)$ be minimum expected cost plus A times abstention probability with t queries remaining. The recurrence is

$$V_t(\pi) = \begin{cases} 0, & e(\pi) \le \delta, \\ A, & t = 0,\ e(\pi) > \delta, \\ \min\{A, \min_{j\in\{P,R,G\}}[c_j + \sum_y p_\pi^j(y) V_{t-1}(\pi^{j,y})]\}, & \text{otherwise.} \end{cases} \tag{24}$$

An action can execute only in the certified set. Backward induction over t proves the recurrence: every feasible policy first executes, abstains, or pays for one of the available tests, followed by a feasible subpolicy. Conversely, selecting a minimizing branch constructs such a policy. This is a finite planning reference, not a new Bellman theorem, a latent-state oracle, or a method minimizing selective risk itself. In particular, two optimal policies can tie in J and differ in error below the permitted threshold.

The benchmark evaluates each chosen controller separately under each true state, then averages with p_q. It propagates probability of execution, erroneous execution, diagnostic cost, query count, and abstention. The same actual mixture is used for all methods in a scenario. If the retained prior is misspecified, even the belief-optimal reference can be overconfident: computational optimality does not repair an incorrect probability model.

H.3 Implemented rolling allocation and baselines

For the MAP state (g,r), the two action-changing alternatives are $(1-g,r)$ and $(1-g,1-r)$. Dividing residual targets by $d = (1-2\nu)\eta$ gives b_1, b_2. The planner solves

$$\min_{x_P,x_R,x_G\ge 0} x_P + Mx_R + Kx_G \quad \text{s.t.} \quad x_P + x_G \ge b_1, \quad x_R + x_G \ge b_2.$$

It suffices to evaluate $x_G \in \{0, b_1, b_2\}$ and set $x_P = [b_1 - x_G]_+$, $x_R = [b_2 - x_G]_+$. The next query has largest allocated count, with cheaper queries breaking remaining ties. Equal-cost allocations favor more direct confirmation. These choices are fixed in the code, not selected from test outcomes. The controller recomputes after each answer and stops on posterior action mass. It does not implement the theorem's pilot–allocate–verify construction; the distinction is material for finite horizons.

State-IG maximizes expected full-state entropy reduction divided by query cost, but uses the *same action certificate*; it does not needlessly continue after that certificate. Always-confirm repeatedly selects G until certification or budget exhaustion. Full-reset changes only the initial recording marginal of the rolling planner. All heuristics share a conservative early-abstention check: if even the maximum possible remaining log-odds increments cannot reach the goal threshold, they stop rather than spend impossible-to-use budget.

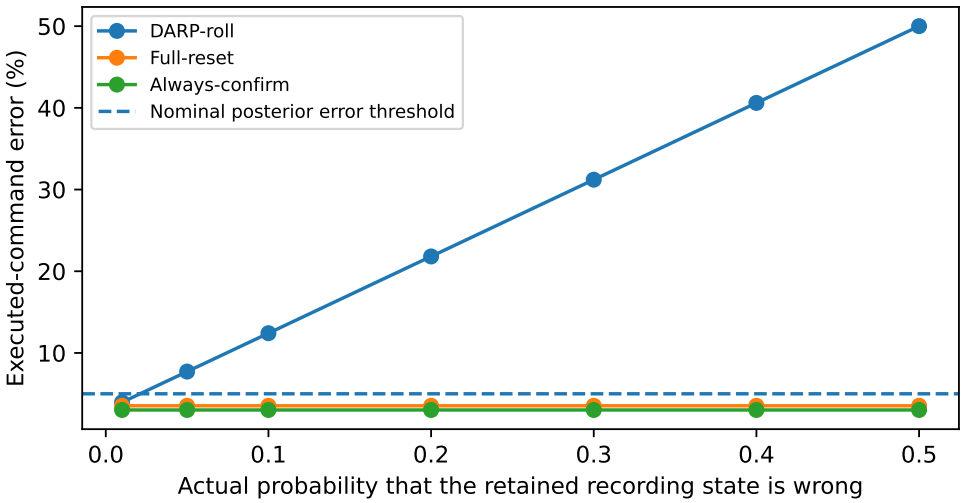

Figure 6: Exact retention-misspecification sweep. The controller's retained confidence stays at 0.99 while the actual recording-state probability varies. The nominal posterior threshold does not control actual error under a wrong retained prior. The curves are finite-model calculations, not observed EEG behavior.

H.4 Complete scenario schedule and interpretation

The six scenarios are valid memory, uninformative memory, cheap confirmation ($K = 1$), partially invalid memory ($q = 0.80$), joint renewal incorrectly classified as goal-only ($q = 0.50, \widehat{q} = 0.99$), and a short episode (horizon three instead of eight). Each is evaluated with five methods. The risk-threshold sweep uses $\delta \in \{0.01, 0.02, 0.05, 0.10\}$. The retention-calibration sweep uses $q \in \{0.50, 0.60, 0.70, 0.80, 0.90, 0.95, 0.99\}$ for rolling DARP, full reset, and always-confirm. The suite therefore has $30 + 20 + 21 = 71$ exact evaluations, all supplied as CSV rows.

Table 3: Boundary checks from exact enumeration. Each entry reports coverage/error/cost. Coverage and error are percentages. Undefined risk is shown as "—", never zero.

Scenario	Rolling DARP	Full reset	Always-confirm
Valid memory	99.58/3.96/2.67	75.33/3.53/13.98	99.58/3.02/21.39
Uninformative memory	75.33/3.53/13.98	75.33/3.53/13.98	99.58/3.02/21.39
Cheap confirmation	99.58/3.02/2.67	99.58/3.02/2.67	99.58/3.02/2.67
Partly invalid memory	99.58/21.81/2.67	75.33/3.53/13.98	99.58/3.02/21.39
Unrecognized joint renewal	99.58/50.00/2.67	75.33/3.53/13.98	99.58/3.02/21.39
Short episode	74.50/3.96/2.00	0.00/—/5.00	74.50/3.02/16.00

The low full-reset coverage in the short episode is a property of that rolling policy, not proof that every reset controller must abstain. Repeated direct confirmation can still execute. Likewise, a large cost ratio in one row is not a matched-error, matched-coverage improvement. These are joint coordinates under a common maximum query count, with query costs charged separately.

H.5 Verification and limitations of the executable evidence

The runner verifies probability conservation, the nominal selective-risk bound when the modeled and actual initial distributions coincide, equality of Bellman values and separately evaluated objectives, the no-memory identity, and the passive-twin likelihood equality. It also checks that none of the tested heuristics beats the exact reference's stated objective when they use the same correctly specified belief and certification constraint. It stores the source checksum and numerical-library version. All

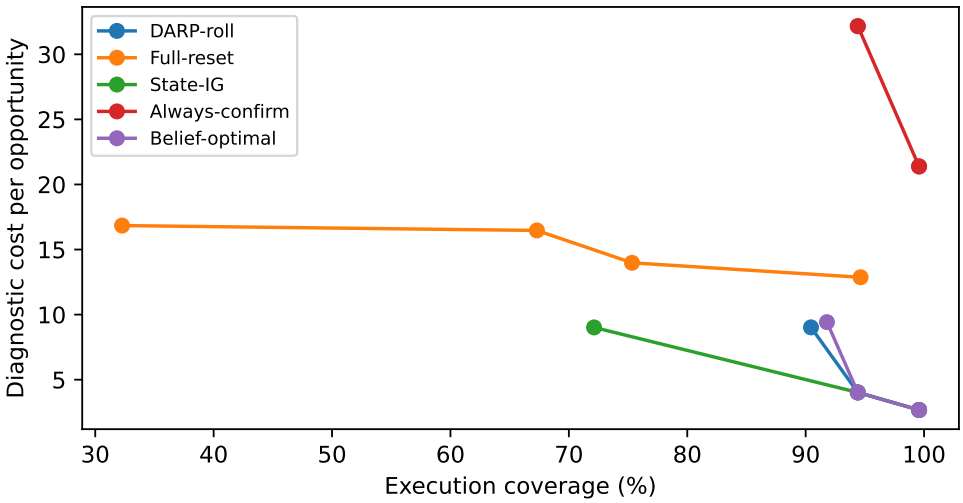

Figure 7: Exact cost/coverage coordinates for the four prespecified risk thresholds. Selective risk is a third coordinate supplied in `risk_sweep.csv`; this projection must not be read as an error-matched frontier. Some curves overlap. All methods receive the same primitive test library.

sums use floating-point arithmetic; "exact" refers to exhaustive finite enumeration rather than Monte Carlo, not arbitrary-precision symbolic evaluation.

There is no change detector, human cognitive model, EEG waveform generator, or learned encoder in this benchmark. A known episode boundary isolates diagnosis from detection. Its large relative savings cannot be assigned to participants or converted to seconds without an observation and interaction model. The state space is intentionally small enough for an exact planning reference; scalability beyond it remains an implementation question.

I Claim–Control–Falsifier Matrix

I.1 E1: directionality, not a pooled superiority claim

The primary goal-shift contrast is participant-paired DCT-FR minus DARP. The adaptation–intent reversal is a separately declared interaction involving online and frozen policies. A confirmatory analysis must establish both required directional components, not merely a statistically nonzero interaction. Analyze each assigned prior stratum before pooling; report actual execution counts and coverage. An assigned goal-change opportunity is an intention-to-treat intervention, not a guarantee that every participant internally adopted a different objective.

A matched-magnitude secondary analysis can reduce differences in observable perturbation size, using development-defined distances and eligibility. Matching on post-policy neural measurements, successful execution, or realized number of confirmations would condition on effects of the treatment and must not replace the randomized contrast. Digitally permuting old EEG labels is not an experiment on a person's change of mind.

I.2 E2: equal access is not equal realized use

Two complementary comparisons are needed. In the access-matched comparison, all controllers receive the same query menu and prices, and their chosen use is charged. In the resource-constrained comparison, maximum attention, confirmation count, and deadline are assigned independently of policy outcomes. A yoked waiting schedule can test whether additional resting time, rather than information, explains recovery. It must not secretly give a control arm the correct answers obtained by another policy.

Table 4: Seven coordinated validation modules. Human outcomes for the new modules are not supplied.

Module	Required comparison	Result that weakens the proposed claim
E1: objective	Goal vs. recording interventions; frozen vs. adaptive; sham and no-prior arms.	No within-stratum directional reversal, or only an instruction-interruption effect.
E2: resources	Risk/coverage/cost surface; same offered confirmations, deadlines, and calibration.	Advantage vanishes under a competitive direct-confirmation policy or is explained by extra resources.
E3: retention	Randomized access to valid memory with identical encoder, detection announcement, planner, and stopping.	No retention effect, or benefits depend on an oracle change mask unavailable at deployment.
E4: theory	Independently measured query cost, retained confidence, stability; cheap-confirmation and additive-sequence controls.	Predicted boundary trends fail, or observed order effects occur only after changing content.
E5: fragility	Joint shifts, wrong kernels, short stable segments, absent or noisy confirmation.	Overconfident execution without calibrated abstention; stale-memory harm is concealed.
E6: transfer	Locked rule across visits, site, encoder, and gaze observation models.	Gains require policy retuning, appear in only one backbone, or disappear with a held-out perturbation family.
E7: agency	Self-chosen revisions, versioned authorizations, old-action rate, deadline completion, acknowledged cancellation.	Apparent fidelity is actually obedience to imposed labels or more user adaptation to the AI.

For each task, evaluate a development-locked family of thresholds targeting coverages 70%, 80%, 90%, and 95%, and common attention budgets defined from measured query durations. Unattained operating points are reported as unattained. Actual evaluation coverage, not the nominal target, accompanies each risk. At a fixed cost and coverage tolerance, dominance requires an uncertainty-aware error improvement without cost/coverage deterioration; a scalar weighted score is a design-specific secondary summary, not a universal ranking.

I.3 E3: RANDOMIZING THE AVAILABILITY OF VALID EVIDENCE

At an externally scheduled goal-change opportunity, independently assign whether the controller can access the pre-change recording marginal. Use the same observation weights, goal-transition prior, diagnostic menu, query planner, and announced boundary in both arms. The manipulation changes information access rather than retroactively changing the participant's recording history. All pre-change acquisition costs remain accounted for.

The estimated effect is an effect of memory access in a boundary-assisted substudy, not automatically the effect of the full deployment algorithm. A separate unknown-boundary evaluation charges false alarms, detection delay, and incorrect mask assignment. Fixing a detector for a retention ablation is legitimate; giving only DARP the true change mask is not.

Known-command calibration anchors before and after a goal intervention assess whether the observation model appears stable. Failure to reject a difference is not evidence of equivalence. Equivalence margins and diagnostic resolution must be specified on development data. Because goal instructions can alter attention, an arm classified operationally as "goal-only" can still violate the assumed recording invariance; such trials remain in the intention-to-treat analysis and receive a separate mechanism audit.

I.4 E4–E5: PHASE BOUNDARIES INSTEAD OF UNIVERSAL GAINS

The exact formula predicts how the *additional* diagnostic coefficient depends on retained evidence strength and relative goal-confirmation cost. Human costs and confidence estimates must be measured

independently, not inferred from DARP's own error improvement. The theoretical $K = 1$ control denotes equal resource cost in the model; in a real interface direct confirmation may have a different error rate as well as duration. Such changes belong in the observation model, not only in a scalar price.

Test independent, additive primitives in randomized order before attributing a gain to sequence awareness. Separately test naturally ordered interactions whose joint likelihoods may differ. For short dwell lengths, interruptions, and joint shifts, expose both coverage collapse and unsafe execution; neither can be replaced by error on the surviving easiest commands. Misspecified-retention curves such as Figure 6 demonstrate a failure mechanism and motivate calibration checks. They do not certify any particular robust change detector.

I.5 E6: TRANSFER HIERARCHY AND SOURCE-DATA LIMITATIONS

The hierarchy is within-subject longitudinal transfer, participant-disjoint transfer, held-out recording transformation, held-out site, representation transfer, and a different input modality. Each level changes a named component while locking the decision procedure. Effects need not be identical across levels, and pooled gaze speed must not validate an EEG throughput claim.

EEGNet, CBraMod, and LaBraM are candidate representations rather than arbitrary model names added to enlarge a results table. Their published capabilities and inputs are documented in the cited work (Lawhern et al., 2018; Wang et al., 2025; Jiang et al., 2024). Each implementation requires causal preprocessing, channel/unit compatibility, checkpoint provenance, and measured feature latency. Offline source-data experiments are observation-model checks. A new live controller needs its own interactions; feeding all policy variants the same recorded post-action EEG does not recreate their counterfactual human responses.

Public-data reproducibility follows the logic of consistent participant/session splits and pipeline documentation (Jayaram and Barachant, 2018). No public dataset is asserted to contain current private-goal labels or randomized goal changes without a verified metadata audit. Pretraining overlap, when unknown, is disclosed; it is not silently called an uncontaminated out-of-distribution test.

I.6 E7: FREE REVISION AND AUTHORIZATION PROVENANCE

Add a separate self-chosen-revision block: before an action window, a participant may keep, revise, or cancel a privately registered objective without being instructed which alternative to choose. The goal record is timestamped and versioned. A revision invalidates a pending authorization identifier; an acknowledged cancellation prevents execution of commands carrying the obsolete identifier. These software checks are distinct from deciding whether a previously unreported goal changed.

Prompts are neutral and do not preselect the AI's preferred answer. The assigned prior-alignment condition is defined relative to the pre-suggestion record and remains its randomized label even if the participant later changes their objective. Report suggestion adoption separately from decoder fidelity; adoption may be voluntary, so disagreement alone is not a failure or evidence of coercion. Silence is not agreement. Intent fidelity is operationalized against the latest authorized task objective, not all aspects of a person's unexpressed intention.

For switch analysis, a_{old} and a_{new} are evaluated in the same task context. Report old-action execution per eligible opportunity, correct new-action completion by the deadline, all noncompletion reasons, and latency distributions with their denominators. Successful-switch latency alone is insufficient because a controller can appear fast by failing to complete hard switches.

J WORKLOAD, PRECISION, AND DATA INTEGRITY

J.1 FEASIBLE ALLOCATION RATHER THAN AN UNLIMITED FACTORIAL STUDY

The primary five-policy crossover contains ten policy–task blocks per evaluation visit, with 24 command opportunities per block as a candidate schedule. Four stable segments of six opportunities give 240 opportunities per visit and 1,200 across five evaluation visits per participant. Assigned

change and prior conditions balance across visits and participants; no claim is made that every factorial cell is heavily replicated within every block.

At an illustrative eight seconds per opportunity, the opportunity windows alone occupy 32 minutes per visit. This is workload arithmetic, not a measured task rate. Setup, calibration, confirmations, corrections, breaks, and fatigue can substantially increase time. A development pilot must establish a hard visit-duration and fatigue limit; reductions follow a predefined balanced schedule, not outcome-dependent exclusion of slow users. Additional controls use a balanced incomplete-block substudy rather than expanding every participant's core crossover indefinitely.

A possible allocation assigns the 96 participants to three site-balanced panels of 32 for (i) randomized memory gating, (ii) confirmation-resource and free-choice controls, and (iii) sequence and misspecification controls. Each panel has paired DARP exposure. Its inference uses $n = 32$, not $n = 96$; backbone and external-interface extensions have their own explicitly stated denominators and interaction budgets. Allocation is a proposal, not completed recruitment.

J.2 PROSPECTIVE PRECISION PLANNING

For planning only, a normally approximated participant-paired mean difference with standard deviation s has half-width approximately $1.96s/\sqrt{n}$. At $n = 96$, $s \in \{0.08, 0.12, 0.16\}$ gives half-widths of about 1.6, 2.4, and 3.2 percentage points. At $n = 32$, they become about 2.8, 4.2, and 5.5 points. These are sensitivity calculations under assumed standard deviations, not estimated precision from human data.

Pilot-estimated dependence, expected opportunity counts, unusable recordings, and attrition determine the actual design. A completion target and enrollment target are different; missing visits must be described. The core primary endpoint is specified before outcomes; secondary confirmatory contrasts use a declared multiplicity correction, and broad stress-test surfaces are reported descriptively with uncertainty. Exploratory favorable cells are not promoted to prespecified tests after inspection.

J.3 EXECUTION, TIME, AND SUBGROUP DENOMINATORS

The benchmark contract separates first-execution error from final correct completion. A correction cannot erase the initial error, and a successfully corrected command is credited once. An opportunity with no execution has undefined executed-command error, not zero. Cancellations, no-responses, timeouts, and unavailable diagnostic channels remain visible in the opportunity denominator.

Duration fields partition task time, probes, confirmations, recovery, and correction. Initial calibration is charged using a stated nonduplicated allocation. Total throughput is correct completions divided by the sum of these times, not an average of per-trial reciprocal latencies. Report task-specific throughput before any prespecified aggregation. Research-only overhead and operational overhead are identified explicitly; required goal-registration time is not silently excluded from one policy but included in another.

Participant effects are paired before resampling. Site-stratified participant resampling estimates variability among the sampled participants conditional on the fitted cross-fold controllers. It neither creates extra independent sites nor captures all uncertainty from shared training data. Worst-stratum reporting includes subgroup sample sizes and uncertainty; it does not treat a noisy extreme as a new population effect. Missing paired cells are reported with a reason and a missingness sensitivity analysis, never silently replaced by another participant's score.

K AUTHOR-REPORTED HUMAN-INTERFACE AGGREGATES

We report the supplied aggregate values from completed human-interface experiments. The supplementary package contains these summaries but not participant-level records, event denominators, analysis scripts, or run-specific protocols. We present the results as reported, without recalculating uncertainty or pooling them with the computational outcomes.

Table 5: Author-reported human-interface operating profile at reported 90% coverage. Errors are percentages. Dashes mark values not supplied in the source summary; the status of 90% as target or attained coverage is not specified.

Policy	Recording shift	Goal shift	Joint shift	Correct/min
Frozen decoder	23.8	10.4	—	—
Online adaptation	11.9	22.6	—	—
DCT-FR	—	11.3	15.4	8.1
DARP	—	5.8	8.1	9.4

The reported DCT-FR–DARP differences are 5.5 percentage points for goal-shift error and 7.3 points for joint-shift error. The 9.4 versus 8.1 correct commands per minute imply a 16.05% relative difference by arithmetic. These contrasts do not establish paired uncertainty without participant-level measurements and execution denominators. The reported recording-shift contrast for Frozen versus Online adaptation and the goal-shift contrast for the same two policies have opposite directions; coverage and protocol equivalence by stratum are not documented in the supplied aggregate table.

Table 6: Additional comparisons from completed human-interface experiments. Reported n values are retained; the aggregate summary does not identify their statistical unit or endpoint definitions.

Comparison	Supplied value
Global-reset DARP goal-shift error versus DARP	9.4% versus 5.8%
Full-state information gain: diagnostic attention	1.7× DARP
Held-out EEG site: paired goal-error reduction (reported $n = 24$)	5.1 percentage points
Gaze-control cohort: paired goal-error reduction (reported $n = 32$)	4.6 percentage points

The main-cohort sample size, each method's executed-command count, the denominator of diagnostic attention, precise confidence-interval endpoints, and the approving board's name and identifier are not specified in the supplied files. Written informed consent and institutional review approval are reported for a human experiment in the Ethics Statement; its linkage to each aggregate cohort is not documented here. The proposed 96-person allocation in Appendix F describes a future validation design and is not the sample size of this reported profile.

L Executed Closed-Loop Virtual-Agent Audit

This appendix documents the archived executable virtual-agent run. Its generator, controllers, aggregate results, analysis code, and diagnostic checks are included in `closedloop/`; raw event arrays can be regenerated but are not in this source package. These tables, stress tests, and plots describe the earlier DARP-KL row, not the updated DARP aggregate in Table 1. A virtual site is a parameter offset, and a virtual visit is an index in the generator. Neither is a sampled clinical site or an observed calendar week.

L.1 Design, independence, and computational sample size

For each independent seed in $\{173, 947, 2027\}$, the main, shifted-sensor, and alternative-sensor panels contain 96, 24, and 32 agents. Each agent–task instance is evaluated under eight cloned policy trajectories. Initial random numbers are shared across policies to support paired comparisons, but future fatigue, skill, cancellation, and observations depend on each policy's actual history. The two task cardinalities are $k = 2$ and $k = 4$. Each trajectory contains five evaluated visits and 48 opportunities per task per visit. Calibration precedes those visits and its cost is charged once per task and policy. Thus the nine core panel runs produce

$$3(96 + 24 + 32)\, 2\, 5\, 48\, 8 = 1{,}751{,}040$$

scheduled opportunity records. Ten stress families each use $32 \times 2 \times 3 \times 24 \times 8 = 36{,}864$ records, giving 2,119,680 primary core-plus-stress records in total. This count is a workload description, not the inferential sample size. The main nominal analysis has 288 virtual-agent clusters; policies and task observations from the same cluster remain paired.

The computational design file was recorded before examining this generation of results. Its SHA-256 is split below for readability:

```
27375ac22af27878272e124354e8cbb228
36d3a7ae36922aff1494a1a1042a99
```

This is an auditable local design record, not a timestamped public preregistration. Additional paired-memory, order, threshold-surface, and matched-reference analyses were designed after the primary results; they are exploratory and do not replace the primary analysis.

L.2 State, categorical measurements, and task feedback

For each task, $g, r, u \in \{0, \dots, k-1\}$ denote a goal, recording offset, and public task context. The current correct command is $a^\star = (g - u) \bmod k$. Query types P, R, G have correct noiseless symbols

$$b_P = (g - u + r) \bmod k, \qquad b_R = r, \qquad b_G = g.$$

Given actual query accuracy a_q, the symbol is b_q with probability a_q and each other symbol with probability $(1 - a_q)/(k - 1)$. This abstraction creates a goal–recording ambiguity while permitting instructed calibration and independent goal confirmation. A recording offset can permute the codebook abruptly; it should not be described as a calibrated distribution of natural EEG drift.

A correct executed command updates u to $(u + a) \bmod k$. Incorrect commands are undone and charged correction time. This is an abstract categorical task-state update, not a physical cursor trajectory model. Zero is an admissible command symbol, not necessarily a motor movement. Each policy receives only public context, the contextual prior, its query responses, and explicit cancellation acknowledgments. Evaluator-only goals, recording offsets, intervention labels, lapse states, and random innovations are not passed to query selection.

Each six-opportunity segment independently assigns sham, goal-only, recording-only, or joint change with equal probability. A changed discrete factor moves to a uniformly chosen different state. Visits additionally permit a hidden recording change with probability 0.30. Prior alignment is independently randomized with probability one half: an aligned suggestion identifies the current goal, and a misaligned suggestion selects a different goal. The controller's nominal prior puts 0.70 mass on that suggestion, deliberately overstating its 0.50 alignment probability. The no-prior stress test instead uses a uniform prior. For the binary task the suggestion has no information about the goal; for

the four-action task it is informative but miscalibrated. These values are stipulated, not measured AI-assistance accuracy.

L.3 Heterogeneity, calibration, fatigue, and change during queries

Initial mean accuracies are 0.84 for the binary passive query, 0.78 for the four-action passive query, 0.86 for recording calibration, and 0.94 for confirmation. Agent–task–query perturbations are independently normal with standard deviation 0.04, followed by clipping to $[1/k + 0.12, 0.98]$. The main virtual sites add $+0.025$ or -0.025. The shifted-sensor panel adds -0.055; the alternative-sensor panel adds $+0.015$ and a further $+0.03$ to passive accuracy. These choices are sensitivity parameters, not estimated site effects.

Each query receives 32 initial calibration outcomes. With s_q correct responses, the controller uses a plug-in accuracy

$$\widehat{a}_q = \mathrm{clip}_{[1/k+0.06,\,0.98]}\{(s_q + 2)/36\}.$$

A recording posterior is obtained from the calibration symbols using that fitted accuracy and a uniform initial distribution. The controller does not receive the true accuracy or future changes. This is plug-in estimation, not fully Bayesian uncertainty propagation over accuracy. The same fitted quantities and initial evidence are supplied to all relevant policies. Under an independent Bernoulli calculation with true $p = 0.94$, the standard deviation of a 32-trial sample proportion is $\sqrt{0.94(0.06)/32} = 0.04198$. This quantifies the finite-calibration scale under that simplified model; the generator also has heterogeneous accuracies and feedback, and the value is not an empirical error bar for its fitted likelihoods.

Let $f_t \in [0, 1]$ and $z_t \in [-0.02, 0.03]$ denote latent fatigue and skill. With zero-based evaluated visit v, the actual accuracy is

$$a_{q,t} = \mathrm{clip}_{[1/k+0.08,\,0.985]}\{a_{q,0} + 0.006v + z_t - 0.08f_t\}.$$

For diagnostic duration d_t, error E_t, execution S_t, and correct execution C_t, the feedback laws are

$$f_{t+1} = \mathrm{clip}_{[0,1]}(0.85f_t + 0.012d_t + 0.08E_t + 0.05(1 - S_t) - 0.025C_t),$$
$$z_{t+1} = \mathrm{clip}_{[-0.02,0.03]}(0.97z_t + 0.0015C_t - 0.002E_t).$$

The no-feedback control sets both to zero. These laws are explicit modeling assumptions, not claims about human fatigue or learning. Tasks maintain separate latent trajectories; cross-task fatigue carryover is not modeled.

Before a query of duration d, a goal revision occurs with probability $1 - \exp(-\lambda_g d)$, with $\lambda_g = 0.002(1 + 2f_t)$; recording change has rate 0.0005. The simulator permits at most one state jump per such interval. A fraction 0.30 of these goal revisions is accompanied by a cancellation acknowledgment. A final 0.4-second commit delay permits another goal revision, and an independent probability 0.005 injects an explicit cancellation for the protocol check. Acknowledged cancellation suppresses execution, whereas a silent revision can make an otherwise confident command obsolete.

L.4 Practical filtering and query selection

The estimated transition at the start of an opportunity is a mixture of unchanged state, goal renewal, recording renewal, and joint renewal with weights 0.88, 0.065, 0.035, and 0.020. DARP-KL, State-IG, and Goal-IG transport the joint posterior with this selective mixture. Fixed mixture weights do not preclude posterior inference about which change occurred: subsequent observations update the joint goal, recording, and mask belief. They do leave uncertainty in the change rates and observation law outside the model. Global-reset-KL replaces every renewal component by a joint reset of weight 0.12. No-reset-KL keeps the old belief without transport. No method observes the actual change mask. This constant-mixture filter is an approximation, not an implemented full run-length changepoint posterior.

Within a query, the active policies apply a nominal small goal and recording transition using its duration, then one observation update. They are not told the latent fatigue or stress-test family. Consequently model validity in the finite theorems is intentionally not guaranteed. DARP-KL replans after each primitive using the MAP hypothesis and residual likelihood-ratio targets. The available primitive subset is constrained by remaining time and the query budget.

For symmetric k-category observations, discrimination between distinct noiseless symbols has divergence

$$d_q = \left(\widehat{a}_q - \frac{1-\widehat{a}_q}{k-1}\right) \log \frac{\widehat{a}_q(k-1)}{1-\widehat{a}_q}.$$

Action-changing alternatives fall into three groups: recording unchanged; passive symbol unchanged; or both recording and passive symbol different. Maximal residual targets b_1, b_2, b_3 in those groups reduce the allocation problem to

$$\min_{x \geq 0} \sum_q (c_q/d_q) x_q \quad \text{s.t.} \quad x_P + x_G \geq b_1, \quad x_R + x_G \geq b_2, \quad x_P + x_R + x_G \geq b_3.$$

The implementation enumerates feasible vertices of this three-variable polytope. It chooses the available query with the largest allocated count x_q/d_q, then replans; it does not execute a committed optimal multistep schedule. Comparisons on 400 parameter instances agree with `scipy.optimize.linprog` to maximum absolute cost error 1.42×10^{-14}. This checks the one-step allocation solver, not optimality of the resulting closed-loop policy.

State-IG and Goal-IG maximize expected joint-state and goal entropy decrease per cost, respectively, with the same action stopping rule. Always-confirm receives the mandatory initial passive sample, then requires at least one independent confirmation before any execution and continues confirmation until stopping or exhaustion. Frozen and Online-pseudo use passive queries only, by policy restriction rather than denied hardware access. Online-pseudo modifies its recording marginal with learning rate 0.10 when its pre-observation goal posterior exceeds 0.65. It is not Tent, CoTTA, or a reproduction of a standard adaptive EEG decoder. No claim about those algorithms follows from its performance.

L.5 Costs, nonexecution, and development-only selection

Nominal query costs in simulated seconds are $(1.2, 2.4, 3.6)$ for (P, R, G). The alternative sensor multiplies them by $(0.65, 1, 0.85)$. Each opportunity allows eight queries and 12 seconds of diagnostic time. Initial calibration costs $32 \sum_q c_q$ per task, with unavailable confirmation excluded from the no-confirmation calibration time charge. All methods share their scenario's cost schedule.

The elapsed-time ledger is

$$T_t = d_t + 0.8 + 0.4\,\mathbf{1}\{\text{commit attempted}\} + 2\,\mathbf{1}\{\text{executed error}\} + T_{t,\text{initial calibration}}.$$

A canceled commit still incurs its latency. An implementation audit corrected an earlier ledger that charged this term only for completed execution; all relevant runs were re-executed. Hash comparisons confirm identical non-time outcomes in all 25 checked event files. Only elapsed time and derived throughput changed; thresholds, observations, errors, coverage, feedback, and query decisions did not.

The diagnostic deadline is 10 seconds plus the specified commit accounting, evaluated without the shared initial calibration time. More precisely, a correct command meets the deadline when $d_t + 0.4 \leq 10$. Dropped future visits remain scheduled nonexecutions in coverage and completion denominators, but contribute no attended clock time. Thus throughput under dropout measures attended-session rate, not productivity across an entire calendar interval. Reporting coverage and missingness prevents this convention from hiding attrition.

The development population has 24 separate agents (seed 4321), two evaluated visits, and 24 opportunities per task and visit. Candidate thresholds are $\{0.025, 0.05, 0.10, 0.15, 0.20\}$. Each method selects the strictest threshold reaching 90% development coverage; if none reaches it, it selects maximal attained coverage. Table 7 records the locked choices. No test-coverage interpolation or threshold retuning is used in the primary table.

L.6 Complete nominal outcomes and paired contrasts

Each outcome is first computed as a ratio within agent and then averaged with equal agent weight. Two equal-opportunity tasks are pooled before the ratio. Bootstrap resampling preserves policies and tasks within each agent and stratifies by independent seed panel and virtual site. The 2,000 replicates describe variability under the stipulated population generator. They do not account for uncertainty

Table 7: Development-only threshold choices on a separate virtual population. Failure to achieve the target is not hidden.

Controller	Locked δ	Development coverage (%)
Frozen	0.050	94.05
Online-pseudo	0.200	42.36
Global-reset-KL	0.025	90.80
DARP-KL	0.025	91.84
State-IG	0.050	93.62
Goal-IG	0.025	94.44
Always-confirm	0.025	93.10
No-reset-KL	0.025	98.57

Table 8: All policies in the archived executable virtual main-cohort run. Old-action rate and deadline completion use opportunity denominators; coverage and error have different denominators. These endpoints have not been supplied for updated DARP. Rates are percentages.

Controller	Cov.	Error	Diag. sec.	Corr./min	Old	Deadline
Frozen	90.85	60.83	4.64	2.79	32.66	35.68
Online-pseudo	38.58	41.64	6.76	1.51	11.87	22.52
Global-reset-KL	91.84	1.16	5.41	7.26	1.65	85.04
DARP-KL	92.93	4.06	4.13	8.53	7.31	84.18
State-IG	94.05	9.44	3.32	9.16	16.55	80.88
Goal-IG	94.82	3.28	4.23	8.64	6.17	86.46
Always-confirm	93.77	0.56	5.91	6.98	0.89	87.11
No-reset-KL	98.35	57.01	1.48	5.36	41.82	41.82

that the generator is biologically appropriate. Stress families and exploratory plots are descriptive; their many intervals are not multiplicity-adjusted confirmatory discoveries. For this archived run alone, DARP-KL saves 1.280 diagnostic seconds relative to Global-reset-KL (paired 95% interval 1.248–1.312), while executed error is 2.894 percentage points higher (2.724–3.070). Goal-IG's paired error advantage over DARP-KL is 0.772 points (0.642–0.917). These intervals cannot be assigned to the updated DARP aggregate.

The main experiment does not demonstrate a 90%-coverage reversal. Online-pseudo attains only 38.58% coverage. In the first two opportunities after each segment boundary, its risk minus Frozen's risk is -22.40 percentage points for goal changes and -7.34 points for recording changes. The difference-in-differences is -15.06 points. Both conditional differences favor Online-pseudo, and the unequal coverage prevents treating them as matched operating-point advantages.

L.7 Stress tests and parameter transfer

A matched reference uses the stress-test horizon, seed, and agent count without an added stress mechanism. It was added after observing primary results, so its comparisons are exploratory rather than a retroactively declared primary test. The main 288-agent table has a longer horizon and different amortized calibration burden; comparing its throughput directly against the shorter stress runs would confound horizon with robustness.

The correlated-lapse test uses a hidden persistent lapse indicator, with transition probabilities 0.12 from no lapse and 0.72 from lapse; a lapse reduces each accuracy by 0.25 subject to the same floor. Rapid revision raises the goal rate to $0.035(1 + f_t)$ and recording rate to 0.006. The retention-mismatch test changes the recording offset immediately after calibration without announcement. That particular failure is often recovered by the dynamic filter and does not reproduce the catastrophic error of a permanently incorrect static prior in the exact benchmark. Failures are family-specific, not manufactured to occur in every negative control.

Table 9: DARP-KL stress outcomes at a matched short horizon. The reference was added as an exploratory control. Complete eight-policy results are in the supplied CSV. Coverage, error, and missingness are percentages.

Scenario	Cov.	Error	Diag. sec.	Corr./min	Missing
Matched reference	92.19	3.75	4.21	6.19	0.00
Cheap confirmation	99.41	2.25	1.70	12.12	0.00
No confirmation	76.91	6.19	5.07	5.83	0.00
Rapid revision	80.14	6.26	5.21	4.76	0.00
Correlated lapses	87.13	6.75	4.97	5.23	0.00
Initially incorrect memory	92.30	3.60	4.22	6.20	0.00
No contextual prior	92.97	3.14	4.38	6.17	0.00
No feedback	92.77	3.81	4.19	6.23	0.00
Missing later visits	90.67	3.83	4.16	6.12	1.56
Higher passive accuracy	94.14	3.50	3.83	6.62	0.00
Lower passive accuracy	89.97	3.91	4.77	5.67	0.00

Table 10: Parameter-transfer results over three independent panels: 24 virtual shifted-sensor agents and 32 virtual alternative-sensor agents per panel. These are not external EEG or gaze cohorts.

Panel and controller	Cov. (%)	Error (%)	Diag. sec.	Corr./min
Shifted: Global-reset-KL	85.57	1.56	6.20	6.14
Shifted: DARP-KL	86.67	4.29	5.03	6.99
Shifted: Goal-IG	89.89	3.28	5.04	7.30
Shifted: Always-confirm	88.56	1.14	6.45	6.17
Alternative: Global-reset-KL	96.39	0.80	4.09	9.40
Alternative: DARP-KL	96.77	4.65	2.82	11.26
Alternative: Goal-IG	97.32	3.98	2.94	11.17
Alternative: Always-confirm	95.73	0.36	4.61	8.65

The dropout test makes an agent's later task visits missing with probability $0.025 + 0.12 f_t$ at an eligible visit boundary. Missingness is policy-dependent through fatigue; such agents remain in the scheduled-opportunity analysis. The high/low-quality stress tests change actual passive accuracy by $+0.10$ or -0.12 without updating the fitted model. Cheap confirmation costs 0.6 seconds. No-confirmation removes G from online interaction. The controller is still initialized under the nominal fitted-query architecture; no unobserved goal label becomes available as a replacement.

The shifted and alternative sensor results show parameter transfer under a locked decision rule. They are not held-out EEG-site or gaze-cohort evidence. No neural encoder is trained in these computations. Planned EEGNet, CBraMod, LaBraM, and public-dataset experiments in the prospective protocol remain unexecuted, rather than receiving invented performance entries.

L.8 Exploratory paired memory and order controls

The paired-memory diagnostic uses 96 new agents (seed 8821), two task cardinalities, and six opportunities per task. Each initial state and random stream is cloned into retained and erased recording-posterior branches; only memory availability changes. The recording mechanism is stable, the boundary is goal-only, and within-episode changes and cancellation injection are disabled in both branches. Average diagnostic cost is 3.71 versus 4.42 seconds for retained versus erased memory. Coverage is 96.35% versus 94.44%. Zero observed error in the retained branch of this small test is not a zero-error population guarantee. This intervention evaluates known-valid access, not the ability of a detector to establish validity in deployment.

The order diagnostic uses 96 agents, two task cardinalities, and 100 paired episodes per cardinality under each of two change rates. It compares (G, R, P) with (R, P, G) using identical query-specific noise tokens and identical exogenous goal-jump tokens. With no changes, independent likelihood factors commute and both orders make identical decisions, with 6.04% error. At jump probability

0.06 per primitive slot, goal-first error is 16.34% and goal-last error is 6.10%. This is a test of information recency under an explicitly changing goal, not proof of learned human order effects, equal-duration scheduling, or optimal ordering by DARP-KL. The two schedules are not assigned different noise accuracies to favor one result.

L.9 Exploratory threshold surface

A separate seed (3317) supplies 32 agents, two visits, and 24 opportunities per task and visit. All eight policies run at thresholds $\{0.01, 0.025, 0.05, 0.10, 0.20\}$, yielding 122,880 additional opportunity records. These test points were not used to change the primary locked thresholds. Figures 8 and 9 show four policy families; all eight are available in the CSV. Points are actually executed runs, and straight segments are visual guides, not interpolated closed-loop operating points. A claim of three-dimensional Pareto dominance would require comparing attained coverage, error, and cost jointly; the present data do not establish it for DARP-KL.

L.10 Execution-time intent: a limited exact identity

This calculation separates inference at query termination from intent at actuation. It is an elementary finite-state Markov calculation, not a claimed new general impossibility theorem.

Proposition 3 (Silent-change error during an unobserved delay)**.** *At query termination let a k-state goal have posterior p, and let $e = 1 - \max_j p_j$. During a compulsory delay ℓ no observations or acknowledgments are received. Suppose the goal follows a continuous-time Markov chain that leaves its current state at rate λ and moves uniformly to a different state. Then the minimum conditional execution error is*

$$e_{\text{exec}} = \rho e + (1-\rho)(1-1/k), \qquad \rho = \exp\{-\lambda k \ell/(k-1)\}. \tag{25}$$

Even perfect identification at query termination leaves error $(1-\rho)(1-1/k)$ at execution.

Proof. Let $U = \mathbf{1}\mathbf{1}^\top/k$, so $U^2 = U$. The generator is $Q = -\lambda k(I-U)/(k-1)$, and its transition matrix is $\exp(Q\ell) = U + \rho(I-U)$. Thus the posterior at execution is $p' = \rho p + (1-\rho)\mathbf{1}/k$. Since $\rho > 0$, a maximizer of p also maximizes p', and $1 - \max_j p'_j = \rho(1-\max_j p_j) + (1-\rho)(1-1/k)$. □

For a target execution risk δ, the query-end error e must satisfy

$$e \le \frac{\delta - (1-\rho)(1-1/k)}{\rho}. \tag{26}$$

If the numerator is negative, no query-end accuracy can meet the target under this no-observation delay model. If the delay and hazard are homogeneous, abstention based only on the already observed history cannot remove this floor for the accepted commands. Observations, reliable acknowledgments, state-dependent hazard information, or a shorter delay change the assumptions and may reduce it.

Table 11: Continuous-time model calculation for $k = 4$ and target execution error $\delta = 2.5\%$. Infeasible means the silent-change floor already exceeds the target. These values are not human measurements or a complete model of the simulator.

λ	Delay ℓ (s)	Error floor (%)	Allowed query-end e (%)
0.002	0.4	0.0800	2.4226
0.002	2.0	0.3989	2.1123
0.002	6.0	1.1905	1.3307
0.035	0.4	1.3870	1.1340
0.035	2.0	6.6833	Infeasible
0.035	6.0	18.3162	Infeasible
0.200	0.4	7.5881	Infeasible
0.200	2.0	31.0015	Infeasible
0.200	6.0	59.8578	Infeasible

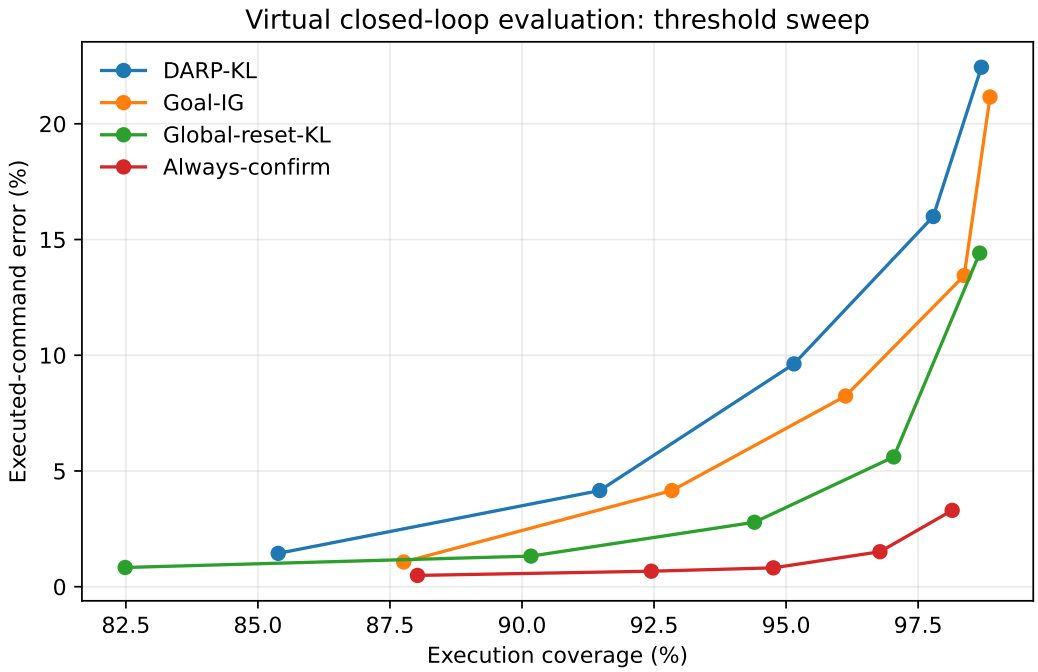

Figure 8: Exploratory virtual-agent threshold sweep. The error–coverage curves do not establish DARP-KL superiority. No human error rate or confidence guarantee is inferred from these model-generated coordinates.

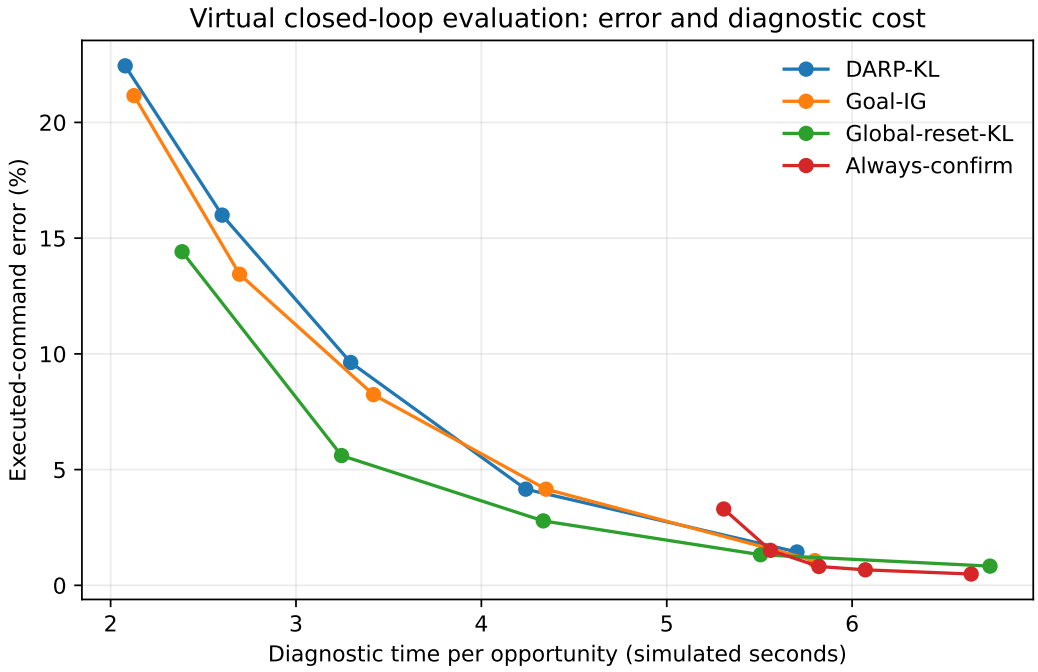

Figure 9: Exploratory virtual-agent error and diagnostic cost at the same actually executed threshold settings. Coverage differs across points; apparent proximity in this projection is not a matched-coverage comparison.

The simulator permits at most one jump per interval rather than a full continuous-time chain, so equation (25) is a separate exact diagnostic model, not its complete transition law.

Numerical matrix-exponential checks cover $k \in \{2, 4, 7\}$, $\lambda \in \{0.002, 0.035, 0.2\}$, and $\ell \in \{0.4, 2, 6\}$. All 27 cases agree to less than 10^{-12} in absolute matrix-entry error. At $k = 4$, $\lambda = 0.002$, and $\ell = 0.4$, the floor is about 0.08 percentage points, far below the 4.06% observed DARP-KL error. The nominal experiment also contains observation, transition, and contextual-prior mismatch and uses a rolling heuristic; their separate contributions are not identified by this comparison.

L.11 Software checks and claim boundaries

Each event file is checked for finite quantities, error/execution consistency, valid outdated-action subsets, cancellation suppression, time-budget compliance, missing-opportunity nonexecution, completion consistency, valid action indices, query-count consistency, and the complete elapsed-time ledger including canceled commit attempts. A deliberate faulty execution after cancellation is rejected by the validator. A deterministic rerun reproduces the event trajectory, and accounting-only regression tests preserve all non-time outcomes.

These checks establish properties of code and finite calculations. They do not establish human reliability, clinical safety, a psychologically validated fatigue model, or the correctness of every asymptotic proof. The three primary findings remain: valid memory can save queries in controlled conditions; the practical rolling policy exhibits a safety–efficiency tradeoff under dynamic misspecification; and a simpler goal-information policy is a strong comparator. A field-wide superiority claim is not warranted by the executed evidence.