What Should We Optimize, and What Must We Know?

Three Computational Experiments on Model Error, Future Goals, and the Cost of Learning

Björn Kenneth Holmström

4 November 2026

Revision note (v0.2): Incorporates the technical review of v0.1, especially correction of A1 H1″ as a difference-in-differences, scope of F2, explicit evidential status, experiment-sequence disclosures, and immutable artifact identifiers. No frozen experimental verdict has been altered.

Evidence convention. **[R]** denotes a result established within one of the specified computational models (with confirmatory or exploratory status stated separately). **[IP]** denotes an interpretive proposition derived from comparison or conceptual analysis. **[H]** denotes a hypothesis proposed for future investigation. Neither a mathematical model result nor its confidence interval is, by itself, evidence about real institutions.

Abstract

Optimizing a coupled system presupposes answers to different questions: what outcomes matter, how actions change the world, which future goals may arise, and what it costs to learn. A planner may optimize exactly against an inadequate representation. This paper synthesizes three synthetic computational experiments in *Systemic Meta-Optimization*. **A1** [confirmatory] investigates model errors in coupled flood management. A wetland-augmented objective does not establish a general hedge against unknown-direction parameter error. Its *difference-in-differences* advantage under an incomplete rather than full model equals 34.3% of the narrow model's regret; the narrow model jointly omits downstream retention effects, peak synchronisation, and rainfall correlation. Ambiguity aversion [confirmatory] recovers 83.1% of the direct planner's regret under a wetter evaluation shift **with exact planning parameters**; the benefit weakens substantially with parameter misspecification and reverses under a drier shift. **B** [confirmatory] keeps dynamics known but makes future goals uncertain. Reachability and diversity correlate with incremental option value (both approximately 0.73), yet at the frozen possibility weight no measure passes the joint test against premium-matched reserves. A uniform prior over an exactly correct 30-goal support recovers 69% of the value of the true goal distribution [exploratory]. **B5** [confirmatory] finds that support knowledge recovers an estimated 59.1% of full-information value (inconclusive against a 50% threshold), while period-displacing polling has negligible value. When surveys accompany ordinary actions [exploratory], they recover 20.9% of the full-information gap when free and 10.5% at a marginal cost of 0.075 reward units. A1's exploratory fixed-programme map finds no tested pair both mutually beneficial across nine strict accounts and synergistic. These findings neither establish a universal optimization theory nor validate institutional interventions in the field. They motivate a narrower programme: diagnose the *type of mismatch*, distinguish measures of possible futures from decision objectives, and model information-acquisition opportunity costs.

Keywords: systemic optimization; model misspecification; robust decision-making; option value; uncertain preferences; value of information; adaptive institutions; multiobjective interactions; computational preregistration.

1. The problem beneath the optimizer

An optimization problem appears to begin with an objective, constraints, and feasible actions. In a coupled system it begins earlier, with choices about which consequences count, which dependencies are represented, how uncertainty is encoded, and which future demands are worth

preparing for. Every omitted mechanism can distort predictions. Every omitted objective dimension can distort evaluation. Every unexamined information cost can make learning look either free or impossible.

A controller need not be computationally incompetent to fail. It may solve its formal problem exactly and still do poorly in the world. This is the starting point for the **Systemic Optimization** series, which complements *Governance as Engineering* (GaE) while asking a different question. GaE considers the architecture through which institutions observe, decide, coordinate and adapt. Systemic Optimization asks how one should *construct and evaluate the optimization problem itself*, especially when coupled outcomes, future preferences, or the world model are uncertain.

The programme's guiding question is more ambitious than this paper's findings: **what structural properties of interconnected systems indicate when mutual benefit is possible, and how should knowing those properties change optimization design?** The present experiments investigate prerequisites. Only A1's exploratory fixed-programme map directly examines mutual benefit, and it does so under a strict, model-specific ledger definition. No experiment here estimates the effect of real governance reforms.

This paper is a synthesis rather than three separate experiment reports. It also includes negative findings, tests that remain inconclusive, and ways in which the diagnostics themselves fail. Numbers from A1, B and B5 are **not pooled**; they have different simulated worlds, endpoints and denominators. The scientific contribution is the comparison of failure mechanisms and remedies, not a common numerical effect size.

1.1 Four objects that must not be conflated

Let a finite-horizon decision problem be described schematically by

$$\mathcal{D} = (\mathcal{S}, \mathcal{A}, P, u, \Gamma, \mathcal{I}, \mathcal{C}),$$

where $\mathcal{S}$ and $\mathcal{A}$ are states and feasible actions, P is the transition law, u is the evaluation criterion, Γ describes uncertain future goals where relevant, $\mathcal{I}$ is the information available to the planner, and $\mathcal{C}$ describes constraints and information-acquisition costs. An implemented planner instead operates on a represented problem $\widehat{\mathcal{D}}$. It selects a policy $\pi(\widehat{\mathcal{D}})$, subsequently evaluated under the true task.

We distinguish four potential inadequacies:

1. **Objective adequacy:** does the objective being optimized usefully express the evaluation criterion? An auxiliary term may improve the realized criterion without itself belonging to it.
2. **Model adequacy:** does the planner's transition law capture the causal structure and parameters that determine consequences?
3. **Goal-information adequacy:** does the planner know which future objectives are feasible or probable?
4. **Learning and adaptation adequacy:** can it acquire relevant knowledge and respond to it before decisions become irreversible, and at what cost?

These are analytical distinctions, not claims of statistical independence or a complete taxonomy. A real institutional failure may involve all four at once. The experiments separate them by design precisely because the remedies should not be assumed interchangeable.

1.2 The comparison standard

For any fixed true task, let $W(\pi)$ mean the expected value of policy π, oriented so that larger is better. For A1, which minimizes total evaluation cost U, set $W = -U$; for B and B5, $W = V$ already denotes reward. If π^* is optimal given the information explicitly available in the reference condition, the realized opportunity loss **for a policy with the same or less information than the reference** is

$$\mathcal{R}(\pi) = W(\pi^*) - W(\pi) \geq 0.$$

Crucially, the relevant oracle changes between experiments. A1's oracle knows the true transition model; B's informed planner knows the true **distribution** of future goals but not their realization; B5 separately reports the value of knowing that distribution, its support, or the realized goal. The common test is not whether a sophisticated planner beats a deliberately weak one. It is whether a proposed intervention adds value over a comparator that isolates its mechanism.

2. Methods and evidential discipline

The three studies use finite synthetic worlds with exact dynamic programming or associated linear programmes, verification tests independent of the main solver, and experimental pools separated by role. Each had a frozen specification and analysis procedure before sealed-test verdicts were read. Every study evaluated **800 sealed test worlds**; B5 additionally ran a **200-world prior-misspecification arm**. These worlds are synthetic parameter draws, not a sample of observed municipalities or communities. Confidence intervals describe uncertainty across the declared world-generating distribution; numerical evaluation of a single world is exact within the solver's specified tolerances. The sealed pools used Latin-hypercube sampling, while the recorded percentile bootstrap resampled complete worlds as if independent; interval interpretation is consequently tied to that analysis approximation, not to observational sampling.

A1 uses a model of two upstream municipalities and one downstream municipality with probabilistic rainfall, wetland establishment, flood protection, emergency coordination and fiscal reserves. B replaces this with a compact activity/resource model in which a community may produce, invest, let capabilities atrophy or regenerate natural resources. B5 retains this second model in a smaller state space but adds a Bayesian goal hierarchy and surveying actions. The modifications are deliberately not an apples-to-apples evaluation of the same intervention.

Confirmatory and exploratory evidence are kept separate. A1's H1′, H1″ and H2′, B's HB1–HB3, and B5's condition-E H5-1–H5-3 have predefined decision rules. The later stress batteries, scenario maps, sensitivity analyses and B5 concurrent-learning amendment are exploratory; their effects can motivate new preregistrations, not revise old verdicts. This distinction is especially important in B5, because concurrent polling was added after verification worlds showed that the originally frozen, period-consuming survey was almost never selected.

Experiment	Main unknown	Main contrast	Primary caution
A1	Dynamics or causal structure of the world	Direct objective vs cross-system augmentation and ambiguity aversion	Different misspecifications favour different remedies
B	Future objective, while dynamics are known	Structural possibility measures vs explicit beliefs and premium-matched reserves	A measure of flexibility need not be a good objective
B5	Which goals are possible and how probable	Support knowledge and costly Bayesian learning vs a flat prior	Information's usefulness is distinct from the cost and timing of acquiring it

2.1 What a normalized recovery statistic means

The normalized percentages in this manuscript are generally **ratios of mean improvements to mean oracle gaps**, not improvements in total system welfare and not percentages of worlds. In B, for example,

$$R_M = \frac{\overline{V(\pi_M) - V(\pi_{\text{myopic}})}}{\overline{V(\pi_{\text{informed}}) - V(\pi_{\text{myopic}})}}.$$

In B5 the denominator instead uses the informed-minus-flat gap, and A1 uses the F0-minus-oracle *cost* gap for H1′ and H2′, whereas H1″ divides a **difference-in-differences** in absolute improvement by the **narrow model** F0-minus-oracle gap. A value greater than one is possible

for a comparison against a poor control. Negative values indicate harm relative to the relevant baseline; they do not imply negative absolute welfare. The experimental designs also have estimability rules for vanishing denominators. A raw ratio without a meaningful denominator must not be presented as a finding.

3. Experiment A1: the world may be misrepresented

3.1 The mechanism isolated

A1 evaluates an 11-period regional flood system with 31,104 states. The common evaluation criterion charges financial outlays, flood damage and monetized harm, with an additional worst-off component in a secondary version. **F0** minimizes that criterion under its planning model. **F1-eco** adds an auxiliary penalty for missing upstream wetlands; **F1-eq** adds one for the worst-off municipality's net damage. **F2** uses a robust hydrological transition ambiguity set. Premium-matched reserve policies and a myopic policy serve as controls. The same chosen policies are scored against the true world.

Two kinds of model imperfection are intentionally separated. **Parametric error** perturbs known mechanisms in parameter-appropriate coordinates, without prescribing which way the errors point. **Structural error** is a bundled simplification: the narrow planning model removes the downstream effect of upstream wetland retention, sets peak synchronisation to zero, and removes correlation between the two tributaries' rainfall. A separate evaluation shift makes heavy rainfall more or less likely than anticipated.

The distinction matters: an extra wetland reward is not universal insurance. In a dry world it may induce costly restoration. In a structurally simplified model, the same reward can steer the planner toward investments neglected by that simplification. **Because the narrow model changes three mechanisms at once, the experiment cannot identify downstream retention as the sole cause** of this substitution effect.

3.2 Confirmatory results [R within A1]

Claim and test	Estimated effect	Adjusted interval	Frozen verdict
H1′: wetland objective, unknown-direction parameter error ($s = 0.25$)	−0.021	[−0.121, 0.071]	Inconclusive
H1′: worst-off objective, same cell	0.001	[−0.015, 0.017]	Refuted: negligible
H1″: wetland objective substitutes for omitted structure	0.343	[0.293, 0.393]	Supported
H1″: worst-off objective	0.020	[0.007, 0.033]	Refuted: negligible
H2′a: ambiguity aversion under wetter shift	0.831	[0.805, 0.854]	Supported
H2′b: ambiguity aversion over premium-matched reserve control	1.727	[1.635, 1.828]	Supported

Intervals are 97.5% adjusted intervals under the frozen family rules. H1′ and H2′ use the specified F0-minus-oracle regret gaps; **H1″ is not a recovery share**: it is the additional absolute improvement under the narrow relative to the full planning model, divided by the narrow-model gap. A separate exploratory narrow-model recovery share is about 0.330. The P0 optimality null passed in all 800 worlds. Premium matching failed in 0.4% of worlds, below its pre-specified limit.

The first result is a limitation, not a triumph: adding a wetland term does not establish a reliable improvement under noisy parameters of unknown direction. Its interval spans benefits and harms that matter. By contrast, the *structural-substitution* contrast is positive: the wetland term improves performance substantially more in the jointly simplified model than in the full

model. This is a **substitution for a bundle of omitted mechanisms**, not an isolated causal effect of retained-water hydrology. The exploratory retention-tertile pattern is consistent with retention contributing, but it does not isolate its contribution. In other cells F2 also partly offsets structural misspecification (approximately 0.31 of regret in the narrow, exact-parameter, unshifted condition).

The robust result is directional and cell-specific. The 83.1% reduction in F0 regret is at parameter-error scale $s = 0$, with a wetter evaluation shift $\psi = +0.1$. At that same shift the exploratory recovery falls to approximately 29.3%, 9.6%, and 6.7% as s rises to 0.1, 0.25, and 0.5. Under $s = 0$ and the drier shift $\psi = -0.1$, recovery is about -1.79, corresponding to **0.48 higher cost units** than F0; the large ratio partly reflects a small mean oracle gap (about 0.27). Critically, the wetter shift lies **outside** the F2 L1 ambiguity ball of radius 0.2: the post-result review computed distances of 0.23–0.32 in the normal regime and 0.21–0.24 in the wet regime across the sealed worlds. Thus the observed benefit is **not** an in-set coverage guarantee of robust-MDP theory (Iyengar, 2005; Nilim & El Ghaoui, 2005). The 1.727 statistic compares F2 with a premium-matched reserve policy that is itself worse than F0; it does not mean F2 recovered 173% of F0's regret.

3.3 Resilience without generalized adaptability [R, exploratory]

The separately run G6 stress battery adds a useful behavioral check. Its questions were outlined in advance, but the specific G6 cells, continuation definition, ledger conventions, tolerance, interpolation and PRIM settings were fixed **after G5 confirmatory verdicts were known** and before G6 results were calculated. These analyses are therefore exploratory. At $s = 0.25$, with the full model, the standard criterion and no rainfall shift, average early cost (periods 0–5) is **14.98** for F0 and **16.54** for F1-eco. The latter is *total modeled evaluation cost*, not simply cash expenditure. Under a later storm shock, excess cost with each policy's own continuation is **5.36** for F0 and **4.67** for F1-eco. F2 shows an intermediate pattern: early cost **15.27**, storm excess **5.11**. The associated full-horizon costs remain close: **5.98**, **6.03** and **6.02**, respectively. Earlier preparation therefore buys some shock protection, but its cost is visible elsewhere.

The proposed *value of acting after a shock* measure generated a methodological counterexample. Myopic and reserve-constrained policies, which had often failed to invest, appeared unusually adaptable because optimal corrective action from their neglected states was worth a great deal. Their high future action value reflected **unfinished protective work**, not necessarily superior response capacity. A1's pre-registered hoarding guard did not establish option preservation beyond matched reserve rules, and the diagnostic cannot distinguish genuine flexibility from accumulated need. A later study must measure response quality together with the damage already allowed to occur.

3.4 Mutual benefit is a separate question [R, exploratory]

A1 also compares three *fixed programmes*: wetland restoration (W), emergency-resource pooling (P), and hard protection (H). It evaluates changes in nine distinct financial, harm, wetland and habitat accounts rather than collapsing them into a single sum. **Mutual benefit** means that every applicable account is no worse and at least one is better. **Synergy** requires a positive interaction effect, which is not the same as Pareto improvement.

With fixed interventions and a stringent nine-account definition (every account non-worse to tolerance $10^{-9}\max(1, |Y(\emptyset)|)$), at the unshifted evaluation condition, restoration alone is mutually beneficial in **14.5%** of worlds. Pooling alone qualifies in **0.25%**, hard protection in **none**. For pairs, W×P is mutually beneficial in **4.25%**, W×H and P×H in **none**. W×P is synergistic in **2.25%**, but *not in the same worlds* where it is mutually beneficial. Across all 800 worlds and tested pairs, **no pairing is both mutually beneficial and synergistic**. W×H shows interfering effects in **42%** of worlds, reflecting substitution between two sources of downstream flood protection.

The tested programmes did **not** furnish a clear positive engineering complementarity: W×H was predominantly substitutive, while apparent P×H synergy in upstream relief can reflect the redistribution of a fixed emergency resource when downstream levees reduce downstream need. Consequently this is **not** evidence that strong engineering complementarities fail to create mutual benefit; it is a model-specific record of **no tested pair with both properties**. It is not an impossibility result. The ledger uses strict nine-account Pareto improvement, hard protection

mechanically lowers the river-habitat sentinel, and the fixed programmes exclude downstream **co-financing of restoration**, although that action exists in the full model. The next question is whether adaptive programmes, transfers or compensation alter those outcomes.

4. Experiment B: future objectives may change

4.1 Possibility as a measurement and as a decision rule

B removes model uncertainty to isolate another problem. Its community has three productive capabilities, a renewable resource, reserves, and an environmental regime: **1,080 physical states**. During an initial phase it pursues a known current goal. At a revelation time $\tau \in \{3, 5, 7\}$, a future goal is drawn from a distribution over **30 goals**. The goal is then known, and all planners adapt optimally for the remainder of a 13-period horizon. The informed reference knows the distribution before τ, not the realized goal; the myopic baseline assumes the current objective will persist.

The study compares four ways to value future possibility without knowing the true distribution:

- **M-reach**, the logarithm of the number of viable reachable states (resource $R \geq 1$) within the pilot-selected $H_M = 4$-period horizon;
- **M-div**, the number of distinct reachable capability profiles;
- **M-rec**, recoverability of resource stock following a standard shock;
- **M-AUP**, a penalty on changes in attainable values of random auxiliary reward functions, modeled on attainable utility preservation.

The principal test did **not** ask whether a measure outperformed myopia alone. It required each measure both to recover at least 5% of the informed–myopic gap **and** to exceed a reserve-based policy with the same foregone current-goal value by at least 5% of that gap. This second comparison matters: a system that merely holds back resources may look "open" without demonstrating anything beyond a familiar saving strategy.

4.2 Results: predictive measure, ineffective optimizer [R]

At the frozen measure weight $\bar{\lambda} = 0.25$, none of the four measures passed the preregistered joint test HB1. Capability diversity recovered **14.6%** of the informed–myopic gap relative to myopia, yet exceeded its premium-matched reserve control by only **0.6% of that gap**, well below the 5% margin. Reachability recovered **4.6%** and fell slightly below its reserve control. AUP recovered **2.7%**; recoverability actually lost approximately **49.6%** of the reference gap relative to myopia. All four failed the separate HB2 comparison with a specified incorrect-belief planner.

At the same time, the state-ranking validity test HB3 **supported both reachability and diversity**: each had a stratified mean Spearman correlation of approximately **0.730** with *incremental future-goal adaptation value*, after matching states on reserves, resources, regime and approximate current-goal value. The two correlations were also nearly indistinguishable; they should not be treated as evidence for two independent mechanisms.

B planner	Share of informed–myopic value gap recovered	Evidential role
Uniform belief over 30 conceivable goals	**0.690**	Exploratory explicit-belief control
Wrong specified belief at the primary error level	**≈0.45**	Exploratory comparator
Capability diversity (M-div)	0.146	Fails joint HB1; supports HB3 as a measure
Minimax regret across goals	0.108	Exploratory robust comparator
Reachability (M-reach)	0.046	Fails joint HB1; supports HB3 as a measure
Attainable utility preservation (M-AUP)	0.027	Fails joint HB1
Recoverability (M-rec)	−0.496	Harmful in aggregate

The **uniform-belief control** is not merely another representation of generic openness: it speci-

fies a belief about *which goals are conceivable*, even while assigning them equal weight. Its 69% recovery therefore suggests that specifying a relevant goal space may be much more actionable than maximizing an unweighted structural count. The experiment does not show that uniform beliefs outperform every equally inaccurate concentrated belief; the comparator's actual distributional distance varies by world.

The distinction generalizes conceptually. Suppose a system has many possible future actions. The number of those actions is not their expected value. Their importance depends on which future goal materializes, how much each action contributes toward that goal, and what it costs to preserve the corresponding capacity. A measure may rank physical states by their potential well yet provide a poor marginal signal for choosing current actions.

[IP — Measurement–objective separation.] A statistic can be valid as a *diagnostic of opportunity* while failing as a *reward for producing opportunity*. Promoting any useful indicator into an objective is an additional causal claim about decision incentives, not a free inference from correlation.

Weight sensitivity and support caveats. Exploratory sensitivity at $\bar{\lambda} = 1.0$ raises M-reach's recovery to about 0.279, but it was **not reserve-premium-matched** at that weight; a stronger-weight version might or might not pass HB1. At $\bar{\lambda} = 0.05$, the measures lie within about ±0.03 of the gap. In B the 30-goal list is **exactly the true support** of the goal distribution, so a uniform prior has a particularly favorable specification advantage; B5 varies that support.

4.3 Boundaries and counterexamples

B contains worlds where specialization pays, unused capabilities decay, and future goals barely change. Its exploratory maps indicate that diversity's relative value declines when specialization returns or capability atrophy rise. Unrestricted positive-probability reachability correlates positively with incremental option value; a high-probability-path variant, under its specific transition threshold, correlates negatively. This is a warning against equating *likely paths* with *valuable options*, not proof that rare states should dominate planning.

The experiment's favorable assumption is that all planners can adapt **perfectly once the future goal is revealed**. That makes retained options especially usable. It also assumes that all conceivable goals, activities and actions exist in the model from the outset. It cannot test the capacity to *create* new possibilities, redefine goals through deliberation, or discover activities outside the original state space. Its negative result for these four formulations is not a theorem that possibility preservation never works.

The earlier literature already recognizes the importance of preservation and irreversibility. Real-options economics assigns value to waiting for information (Arrow & Fisher, 1974; Dixit & Pindyck, 1994). Relative reachability and attainable utility preservation provide formally different measures for discouraging undesirable side effects (Krakovna et al., 2019; Turner, Hadfield-Menell & Tadepalli, 2020). B is a deliberately narrow comparative experiment, not a claim to have originated the idea that retained options can matter.

Control comparison. The premium-matched reserve policy follows much the same tertile pattern: its recovery changes with specialization returns and atrophy in parallel to M-div (for example, about 0.188/0.124/0.114 versus M-div's 0.201/0.127/0.118 by specialization tertile). Thus this conditional pattern should not be attributed uniquely to the structural diversity objective.

5. Experiment B5: information is useful, but learning is an action

5.1 What does the planner not know?

B5 follows from B's unexpected strength of the uniform-belief control. Instead of supposing the same 30 goals can occur in every world, B5 draws a latent support set S of size $k \in \{2, 3, 4, 6, 8, 12\}$, uniformly from the 30-goal universe, then assigns that support Dirichlet weights with concentration $\alpha \in \{0.3, 1, 3\}$. Under this symmetric hierarchical prior, the prior predictive distribution over goals is *exactly uniform*. The planner knows the prior family but not the world's support or weights.

Three reference values isolate the source of information:

$$I = V_{\text{inf}} - V_{\text{flat}} = (V_S - V_{\text{flat}}) + (V_{\text{inf}} - V_S).$$

Here V_S is a policy optimized under a uniform predictive *within the true support*, and V_{inf} uses its actual probability weights. This is an accounting decomposition of value, not a general statement that one form of information is always more important.

The physical world is reduced to **324 states** to permit exact planning over up to **496 observation-history belief states**. The planner may obtain a poll that samples the true future-goal distribution, contaminated by known uniform noise. Its observations update an explicitly derived Bayesian posterior, in the tradition of Bayes-adaptive planning (Guez, Silver & Dayan, 2012). B5 then makes a critical distinction between *learning that displaces productive activity* and *learning that accompanies productive activity*.

5.2 What was actually preregistered [R, confirmatory]

The original **condition E** required a poll to occupy an entire decision period. Production or investment stops, normal capability atrophy still occurs, and polling may consume reserves. The Bayes-adaptive planner chooses whether to poll and is free to decline.

Frozen hypothesis	Estimate	Adjusted interval	Verdict
H5-1: support knowledge recovers at least half of I	**0.591**	[0.488, 0.682]	**Inconclusive**
H5-2: optimal exclusive polling recovers at least 5% of I	≈0.000	[−0.000008, 0.0000004]	**Refuted: negligible**
H5-3: directional associations with noise, support size and time	—	All include zero	**All inconclusive**

The half-information claim is **not confirmed** merely because its point estimate is 59.1%. The pre-registered adjusted lower bound is **0.488**, below the threshold **0.5**. The equally important negative result is that an optimal learner almost never chooses the original polling action: only **9 of 800** test worlds show any positive probability of polling, and the aggregate probability mass is tiny. With almost every gain zero, H5-3's formally estimable correlations are not substantively informative. Exploratory H5-4 finds essentially no change in capabilities retained at revelation.

The built-in optimality and calibration checks passed. These checks establish consistency with the mathematical model and the sampled prior; they do not certify the model as a realistic description of public opinion or future social needs.

5.3 A documented amendment that changed the interpretation, not the verdict [R, exploratory]

Verification-stage tests revealed that a period-consuming poll was prohibitively costly. Before the pilot and test pools were generated, a documented B5-0.1 amendment **preserved E's frozen hypotheses** and added two exploratory comparisons: **C**, where a survey accompanies an otherwise normal physical action and costs a fixed reward amount, and **F**, where one free observation arrives before the first action. The concurrent-survey costs were fixed at **0**, **0.075** and **0.375** reward units, corresponding to zero, one tenth, and one half of **the Phase 1 reward of one period producing activity 1 at capability 1** ($u_P = 0.75$ reward units), the same in every world. The investigators knew that a verification-world free-observation diagnostic had mean value **0.092** when setting this rule; that disclosure is part of the record. Condition-E **test-pool** verdicts were computed and locked before any C or F **development- or test-world values** were viewed. Before that lock, C polling-rate flags were seen on 16 verification worlds and 50 pilot worlds, and C/F values appeared in a handful of verification-world implementation smoke tests. No C or F development/test result was inspected before E's lock.

Figure 1. Exploratory 95% paired-bootstrap intervals from `figures/b5_learning_cost.png` (figure generator: `wp0_build/figure.py`, see source bundle). E's interval [−0.000007, 0.0000002]

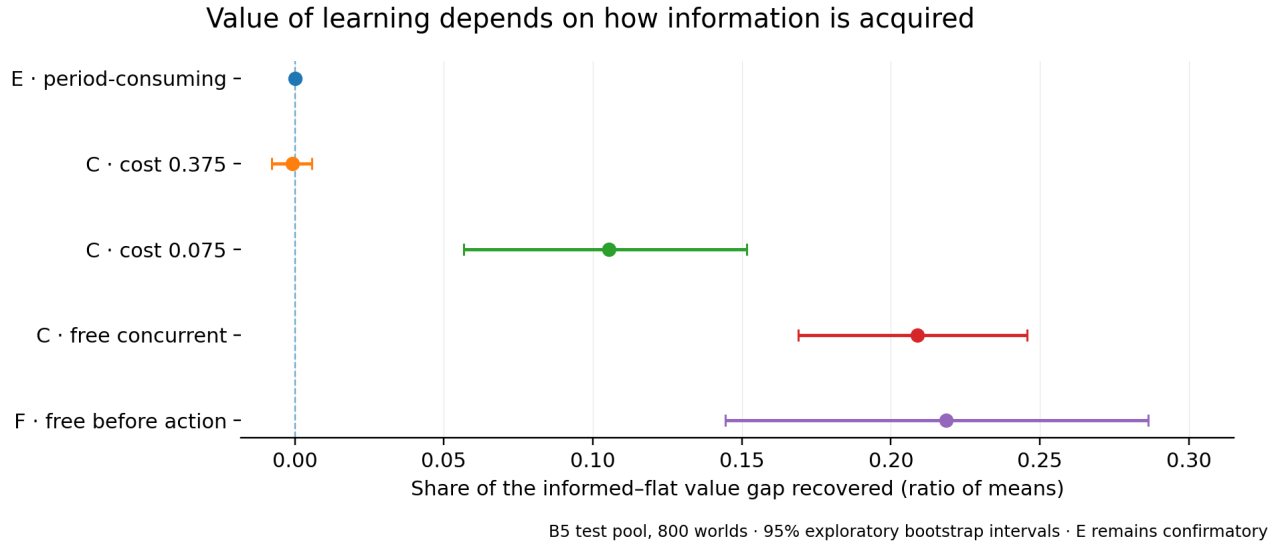


Figure 1: B5 information acquisition results, exploratory

is too narrow to be visible at this scale. Confirmatory H5 tests use adjusted 97.5% intervals, **not** these exploratory intervals.

Information condition	Share of I recovered	Worlds where learning occurs	Expected polls per world
E: survey displaces a period	≈0.000	1.1%	≈0.000005
C: alongside action, cost 0.375	−0.001	3.1%	0.009
C: alongside action, cost 0.075	**0.105**	60.9%	0.327
C: alongside action, free	**0.209**	100%	1.600
F: one free observation before action	**0.219**	100%	1.000
Support known exactly	**0.591**	Not an acquisition policy	—

These are **exploratory** ratios of mean gains against the flat planner to mean I across 800 test worlds, not new confirmatory verdicts. The mean ex-ante value of one free observation is **0.087** reward units (95% interval [0.078, 0.096]) and is positive in all 800 worlds. Yet an observation that displaces a whole period is rarely purchased. Even a marginal cost of 0.075 roughly halves the realized value of free concurrent learning; at 0.375 the advantage nearly disappears.

Timing also matters. One observation before acting (F) recovers about as much as up to two surveys taken alongside production (C at zero cost). This is not a general dominance claim between early information and two later signals; it describes the chosen goal family and action horizon.

On **a separate, unpaired sample of 200 additional worlds with clustered rather than uniformly selected supports**, knowing the support recovers about **0.914** of the full-information gap, while free concurrent learning recovers **0.263** and one free early observation **0.280**. The mean informed-minus-flat gap is larger there (0.640 versus 0.438), and the free-concurrent gain is about 0.168 rather than 0.091 reward units. This single misspecification regime illustrates that being calibrated to one's prior does not universally upper-bound *improvement over a flat baseline*; it is not a general claim about wrong priors. The misspecification arm is exploratory and does not authorize retuning the posterior.

5.4 The crucial separation [IP]

B5 does not show that research is unprofitable. It shows that **a particular information source, purchased through a particular interruption, is not worth its cost under these model**

parameters. The Bayesian learner is not irrational: information increases expected decision quality, but obtaining it may destroy an even more valuable opportunity. A policy that can observe while acting faces a different constraint and may rationally learn.

The expected marginal value of an experiment is decision-relative. A signal may predict the future goal accurately and still have no effect if every possible posterior implies the same action; conversely, a coarse signal may be consequential if it changes an irreversible early choice. This is consistent with Bayesian decision theory and sequential preference elicitation (Blackwell, 1953; Boutilier, 2002). The cost of learning is not always monetary: delay, foregone production, coordination burden, and the decay of unexercised capability all matter.

[IP — Learning architecture matters.] It may be more productive to ask *whether information can be obtained without suspending action* than to ask whether information has positive abstract value. B5 gives one model-based reason to examine that design variable. It does **not** prove that dedicated research, deliberation or democratic consultation should be made subordinate to ongoing production.

6. What the three experiments jointly say—and do not say

6.1 No universal remedy follows from a legitimate concern

One can advocate ecological consideration, robustness, diversity and learning for excellent reasons. The experiments test a different proposition: does translating that concern into a particular objective term, penalty, or action improve a specified evaluation criterion under a specified uncertainty?

A1 answers **sometimes and conditionally**. An ecological augmentation can partly substitute for a bundled omission of three hydrological features, but it can be costly when the modeled risk is overestimated. Ambiguity aversion helps under the adverse evaluation shift explored in A1, even though that shift lies outside its nominal ambiguity ball, but costs under the opposite shift. A generic safety-looking reserve floor can lose badly by delaying necessary protection.

B answers **not necessarily**, even when a measure is statistically informative. Counting accessible states or diversity captures some properties of incremental option value; optimizing those quantities can still be inferior to an informed belief about future objectives or a reserve policy with comparable current-goal sacrifice.

B5 answers **not at any cost**. The information source is useful, yet the originally specified learning action almost never passes the opportunity-cost test. Allowing concurrent learning changes the answer, without changing the laws of Bayesian inference or the set of possible goals.

These are results of different models. One should not place the A1 figure of 83%, B's 69%, and B5's 21% on a common league table: each has a different reference regret, opportunity set, and experimental task.

6.2 Three separations the evidence requires

Separation A: causal model and objective. An augmented reward may compensate for a missing causal pathway, but it remains distinct from explicitly representing that pathway. Treating the two as interchangeable would conceal the directional dependence of A1's ecological objective. A model improvement and an incentive change can generate similar behavior for different reasons; their response under a new shift may differ.

Separation B: option measurement and option valuation. In B, HB3's ranking validity does not imply HB1's policy value. Reachable options must be interpreted through possible future goals, their probabilities, their conditional payoffs, and the costs of preserving them. This is a reason to measure structural capability and decision value separately.

Separation C: information and acquisition architecture. In B5, an observable signal and a practical learning channel are not equivalent. A theoretically informative survey does not guarantee beneficial learning if it consumes time, resources or institutional attention before a decision can be revised.

6.3 A provisional diagnostic framework [IP]

The following is a **proposal for future operationalization**, not a validated four-factor theory:

Question	Diagnostic object	Failure example	Candidate response to test
What matters?	Evaluation accounts and explicit trade-offs	Mutually beneficial totals obscure losses in a habitat account	Multi-account reporting; legitimate preference elicitation
How does the world work?	Causal pathways and uncertain transitions	Omitted upstream retention in A1	Improve representation; compare targeted augmentation and robust planning
What might matter later?	Goal support and distribution	B's possibility counts are weak objectives	Represent plausible future goals; value retention conditionally
Can we learn in time?	Signal quality, delay, costs, policy updates	B5's exclusive poll interrupts productive activity	Compare integrated feedback, protected inquiry, and adaptive timing

The table is not a checklist that guarantees success. It is a way to prevent selecting remedies on the basis of slogans. It asks the analyst to identify the exact information loss, model error or incentive mismatch before choosing a formal intervention.

6.4 Negative results that should survive publication

A credible synthesis must retain four inconvenient observations.

First, **the worst-off auxiliary term F1-eq barely moves A1 outcomes** in most relevant cells. Including an equity-sounding term does not establish equity improvements; here, its induced policy changes were small under the selected normalization and environment.

Second, **B's possibility measures did not pass the primary joint test**. Strong correlations of state ranks with future option value coexist with poor action incentives and failure against reserve-matched controls.

Third, **B5's first polling channel was effectively unused**. The amended channels are informative exploratory contrasts but cannot retroactively rescue H5-2.

Fourth, **A1's preferred adaptability metric sometimes rewards unfinished work**. Any subsequent paper that uses value of future corrective action as "capacity" must account for accumulated liabilities. A large capability-to-recover number may indicate a fragile, underprepared state.

7. The programme's unanswered question: mutual benefit

Mutual benefit is not a scalar objective unless one has already specified whose benefits matter and how conflicting outcomes will be compared. A policy can reduce total harm while imposing disproportionate costs on a particular community. A technically synergistic pair of interventions may be politically unacceptable; a non-synergistic or substitutive pair may still be jointly beneficial under a suitable compensation scheme.

A1's strict ledger is a valuable first discipline precisely because it does not hide these tensions. It reports the financial and harm consequences separately across upstream and downstream municipalities, with ecological sentinels. Its negative intersection of synergy and Pareto improvement should prompt **mechanism design**, not despair. A fixed combination of wetlands, levees and pooling may fail to satisfy every account, while contingent co-financing, transfers, timing or negotiated rules could change the feasible set. None of those institutional arrangements was tested in the fixed-programme map.

[H1 — Conditional mutual benefit.] Under heterogeneous physical externalities, the set of Pareto-improving joint programmes will depend not only on engineering synergies but on compensation, timing, cost assignment and the set of rights or accounts that must be protected.

A follow-up study should specify alternative rules *before* evaluation and check the distribution of gains, not merely the aggregate.

[H2 — Cost-adjusted adaptable capacity.] A measure of adaptable capacity that jointly accounts for response cost, prior damage avoided, and unrealized needs will distinguish resilient preparation from deferred repair more faithfully than A1's raw value-of-action statistic. It should be tested against deliberate underinvestment counterexamples.

[H3 — Goal-weighted optionality.] Possibility-preserving planners that weight retained capabilities by an explicit, uncertainty-aware model of *plausible goals* will perform better than unweighted reachability in some regimes, but can still lose when their goal model is wrong or the cost of preservation is high. This requires new worlds and new hypotheses, not post hoc tuning of B.

[H4 — Integrated learning.] With comparable signal quality, institutions able to gather and use feedback during ordinary operations will have a greater region of positive net learning value than institutions that must halt core activities to acquire information. This does not imply that protected, dedicated learning is never worth its cost; that alternative belongs in the comparison.

Each hypothesis is falsifiable in a constructed model but remains untested in this paper. The step from synthetic experiments to an empirical institutional claim would require operational measures, field data, explicit welfare judgments, and attention to legitimacy and asymmetric power. The review chain here is AI-assisted throughout; no independent human solver-and-evaluator audit has yet been reported.

8. Research agenda and relation to Governance as Engineering

A defensible next sequence should preserve the same methodological restrictions used here: sealed worlds, independent verification, negative controls, explicit nulls, and clear separations between model findings and interpretations. Three directions follow naturally.

A2: structure-aware optimization under model incompleteness. Compare better causal representations, targeted auxiliary objectives, robust ambiguity sets and combinations of these. Vary the relevance of omitted mechanisms independently of the direction of parameter error. The principal question should be *when* a surrogate objective helps, rather than whether complexity-affirming objectives are generically superior.

A3: mutual benefit under adaptive programmes and bargaining constraints. Replace fixed interventions with adaptive regional and municipal strategies. Track the full nine-account consequences, check transfers and compensation, and report which accounts block feasible mutual benefit. Distinguish positive engineering interaction from gains that everyone can accept. In particular, test whether the fixed-programme substitution of wetlands and levees persists when policies can adjust timing and intensity.

B6 or a future epistemic-options study: allow agents to investigate possible goals, but also to *create* new activities or learn new capabilities that expand the space of feasible futures. Compare the value of expanding that space with merely maintaining existing capabilities. This study would test a phenomenon B and B5 exclude by construction.

These directions have precedents in adaptive policy pathways, where decisions are deliberately staged as information arrives (Haasnoot et al., 2013). They are compatible with GaE but do not validate its governance mechanisms by themselves. The link is architectural: an institution's observation channels, objectives, decision rights and feedback costs determine what optimization problem it can even formulate. GaE concerns the organizations and boundaries through which those choices are made; Systemic Optimization investigates their decision-theoretic consequences.

A working-paper series can make this complementarity productive by treating governance extrapolations as **[IP]**, not **[R]**, until appropriate empirical or institutionally realistic models test them. The series should welcome genuinely negative results, since the point is to improve problem construction, not to demonstrate that one favorite vocabulary always wins.

9. Conclusion

The question "what should we optimize?" cannot be answered by selecting a more expansive objective function alone. A planner may have the wrong causal model, the wrong expectations about future goals, or the wrong way of paying for information. Three experiments test those mismatches separately.

A1 shows that objective augmentation and ambiguity aversion have conditional, sharply different strengths and failures. B shows that, at the frozen possibility-measure weight, structural option metrics can be diagnostically informative without beating the specified premium-matched reserve controls. B5 shows that information may be valuable while **its modeled method of acquisition** costs more than the information is worth: in this reduced environment, a period of foregone productive activity often overwhelms the value of two noisy signals. Changing the acquisition channel alters the practical decision without changing the signal's statistical meaning.

The joint conclusion is deliberately modest. **[IP] The architecture of an optimization problem should be matched to the particular uncertainty and institutional constraint it faces.** No single hedge, diversity score, or learning policy is justified across all four dimensions by these results. The design task is to ask which world mechanisms are represented, which future values are plausible, what commitments are reversible, and when information can still change an action.

That is the proposed foundation of *Systemic Optimization*: not a universal method for finding the best outcome, but a disciplined way to investigate **why an apparently correct optimization problem may not be the right problem to solve**.

Appendix A. Audit trail and reproducibility

A.1 Experimental source files

This paper was prepared from the uploaded frozen specifications, aggregate outputs and archived world records. The **technical review of v0.1** reproduced A1 G5 and B B4 verdicts using each study's frozen analysis script and checked B5 against its locked report; these are arithmetic/provenance reproductions, not independent implementations of the solvers. At the drafting stage, the recorded aggregates were inspected and key mean effects independently re-aggregated; this is **not** a complete independent rerun of the solver or Monte Carlo verification suites. The supplied study authors' audits report that the original test suites and provenance gates passed.

Study	Primary supplied evidence	Sealed-test reference
A1	A1 v0.2.2 G4-frozen revision, with *post-G5* report additions §§19–20; `g5_verdicts.json`; `g5_worlds.zip`; `g6_exploratory.json`; diagnostic code archive	800 worlds; frozen G4 code fingerprint `22568e14ddcb5d2cc0763d784709305542d5a2abf`
B	B v0.2 B3-frozen analysis and B2 horizon record; `b4_worlds.zip`; B4 verdicts reproduced from records with frozen analysis code during the technical review; B4 exploratory output	800 worlds; $H_M = 4$; frozen analysis code fingerprint `b57c71f4979715fd5cc053d0104a4ab6cae971a67`
B5	B5 v0.2 revision 23 plus B5-0.1 amendment revision 31 and post-test §14; `b54_E_verdicts.json`; `b54_exploratory.json`; `b54_lock.json`; `b54_worlds.zip`; `b53_worlds.zip`	800 test worlds plus 200 clustered-support worlds; frozen B5-3 code fingerprint `a2a8f827e36e3aa602dca2f1a9141a387d18d76dc`

B5's original condition-E verdicts file has SHA-256 `0252bffa7adf2afcb62e6152b58cf1083e591b7a067f6fcfa0a` matching `b54_lock.json` and the reference carried in the exploratory output. The B5 test-world

archive includes **800** original test records and **200** misspecification records; its separate development archive includes **100** development records. The aggregate B5 means reproduce from those test records, including the informed-minus-flat mean gap of **0.4376338** and the concurrent-survey gains of approximately **0.091376** (free) and **0.046154** (cost 0.075) reward units. A1's G6 report records 800/800 passed replay and ledger checks.

A.2 A reader's checklist for checking the claims

1. Check the recorded fingerprints and pool IDs before computing any outcomes, and retain the signed-off configurations and amendments.
2. Reproduce the confirmatory statistics first, using the pre-frozen estimability and multiple-comparison rules. Do not use exploratory contrasts to select confirmatory verdicts.
3. Reproduce every normalized effect as a ratio of **paired means** with its specified oracle gap and interval; do not average per-world ratios unless the protocol requires it.
4. Verify A1's wet and dry counterexamples, B's reserve-matched controls and HB3 ranking validity, and B5's exclusive versus concurrent survey distinctions.
5. Separate deterministic model identities and numerical tests from world-prior conclusions, then separate both from claims about real governance.

A.3 Methodological disclosure

The study programme was initiated and directed by **Björn Kenneth Holmström**. The experiments and specifications were developed with substantial AI assistance, including implementation and technical analysis in Claude. This synthesis draft was prepared by ChatGPT through review of the uploaded research artifacts, with selected numerical cross-checks, and is intended for subsequent author revision and an independent technical source audit. **Draft v0.2 is not a peer-reviewed publication, and AI cross-checks are not an independent human code audit.** A later public version should identify the definitive repository location, software environment, exact frozen artifact set and any material changes made during author review.

A.4 Frozen artifact identifiers and environments

The download-copy filenames above are locators, not persistent identifiers. The following recorded values should be checked against the canonical freeze records **before** a public release. The A1 G6 script fingerprint and A1/B workstation package versions remain to be included in a finalized release manifest; they should not be inferred from a later computer environment.

A1 (frozen G4). Document hash prefix `f07b3584…`; G0 configuration SHA-256 `fe9d52be094fcb40baeff5cedd61fb5f25676967cc5dfbd4714f53463a8effe9`; test-pool SHA-256 `a788bf1b85f9c9f9c978ff21f730f8b80ff337aba05c23ec0112b513aed70ba8`; G4 analysis fingerprint `22568e14ddcb5d2cc0763d784709305542d5a2abf4493c5ce0bf0c51cee257c4`.

B (frozen B3). Document revision 49, hash prefix `8fd0cf3a…`; B0 configuration SHA-256 `69ebe82650a4384dee53e1d64ddbf58295d5c6eac2ff22bd0f301d20c32f9fda`; horizon $H_M = 4$; sealed-pool SHA-256 `d2a8fdf01854e91d7f9c9803f2eecd06c361b7e2e586ca66d2ccb4b530457600`; analysis fingerprint `b57c71f4979715fd5cc053d0104a4ab6cae971a6756bee58ab90964a0b177d6f`.

B5 (frozen B5-3). Document revision 23 with export SHA-256 `6e339fa90e00a12d928827cad1cc8f8c4e88b108fea4fc30b59167302cada92b` and amendment revision 31; configuration SHA-256 `248e1e9dcff8f9609d96c002ba6c63a1e87bc116d99170dadfc780a9e03488b7`; amendment SHA-256 `347db34c86910f68f3c75983babd34f766a81ecbcb4e6b53674697b13fe99c99`. Sealed test-pool SHA-256 `4b393fdf0c25eb6f7551c1e49c5b176f7e26022f22b90ad6911f6b94afefa3cb`; misspecification-pool SHA-256 `8e88994f5fbf221c0cf27f09c4c746240395888be64d5c9b67eebeeaae768414`; B5-3 fingerprint `a2a8f827e36e3aa602dca2f1a9141a387d18d76dc44b235d48a185dd342f034d`; condition-E verdict lock `0252bffa7adf2afcb62e6152b58cf1083e591b7a067f6fcfa0afb913032984d1`.

The B5 pool record specifies **Python 3.10.12, NumPy 2.2.6, SciPy 1.15.3** on x86_64. This environment matters for regenerating the sealed pools: RNG algorithms and numerical order can change across versions. Matching a CSV or JSON statistic does not by itself certify matching RNG state or code identity. The A1 and B workstation environments should be recorded from their original execution machines, not reconstructed from the B5 versions.

A.5 Sequence of analysis and audit

A1's G0–G5 confirmatory procedures were frozen before the sealed G5 verdicts. **G6's runner, analysis code, and several diagnostic operational choices were written after those verdicts were available**; G6 is exploratory. B's B4 confirmatory analysis followed its B3 freeze. B5's B5-0.1 amendment was made after seeing zero-polling behavior in verification worlds and an exploratory free-observation diagnostic, but before the development and sealed-test comparisons. For B5-4, E's confirmatory test values were locked before any C/F **development or sealed-test** values were viewed; earlier verification/pilot flags and a few verification smoke-test values were available.

The experiments were designed and directed by the author with **substantial AI assistance**. Claude implemented and analyzed the studies, ChatGPT reviewed successive gates, and Claude later technically reviewed this manuscript, including numerical cross-checks. This is a documented multi-agent review process, **not independent human peer review**. An external human audit of at least one frozen solver and evaluator, together with public release of executable analysis code and original pool manifests, remains an important next validation step.

A.6 Reproducibility update (5 November 2026)

Following completion of draft v0.2, three **author-run computational audits** strengthened the provenance of the experiments. **Audit 1:** all 47 frozen source-file hashes and four sealed-pool hashes matched their records, the B5 condition-E verdict lock matched, and all 158 included tests passed. **Audit 2:** all 2,400 confirmatory world records matched their archived ZIPs; 22 selected point estimates were independently checked; and the frozen analyses reproduced all three complete confirmatory verdict JSON files **byte for byte**, including confidence intervals. **Audit 3:** the frozen simulators regenerated four selected sealed worlds (A1 test 000; B test 000; B5 test 000 and misspecification 000), with **identical scientific fields** after excluding execution-time metadata.

These checks demonstrate source integrity, archived-outcome analysis reproduction, and **sample-level** simulation reproduction. They do **not** constitute a full-pool simulator rerun, independent solver/evaluator validation, human peer review, or evidence that the models represent real-world institutions. No experimental verdict has been changed. The methods and reports are documented in the research repository (Audit 1, Audit 2, Audit 3); access depends on its publication status.

References

Arrow, K. J., & Fisher, A. C. (1974). Environmental preservation, uncertainty, and irreversibility. *The Quarterly Journal of Economics, 88*(2), 312–319. https://doi.org/10.2307/1883074

Blackwell, D. (1953). Equivalent comparisons of experiments. *The Annals of Mathematical Statistics, 24*(2), 265–272. https://doi.org/10.1214/aoms/1177729032

Boutilier, C. (2002). A POMDP formulation of preference elicitation problems. *Proceedings of AAAI-02*, 239–246. https://cdn.aaai.org/AAAI/2002/AAAI02-037.pdf

Dixit, A. K., & Pindyck, R. S. (1994). *Investment under Uncertainty*. Princeton University Press. https://press.princeton.edu/books/hardcover/9780691034102/investment-under-uncertainty

Guez, A., Silver, D., & Dayan, P. (2012). Efficient Bayes-adaptive reinforcement learning using sample-based search. *Advances in Neural Information Processing Systems 25*. https://proceedings.neurips.cc/paper_files/paper/2012/file/35051070e572e47d2c26c241ab88307f-Paper.pdf

Haasnoot, M., Kwakkel, J. H., Walker, W. E., & ter Maat, J. (2013). Dynamic adaptive policy pathways: A method for crafting robust decisions for a deeply uncertain world. *Global Environmental Change, 23*(2), 485–498. https://doi.org/10.1016/j.gloenvcha.2012.12.006

Iyengar, G. N. (2005). Robust dynamic programming. *Mathematics of Operations Research, 30*(2), 257–280. https://doi.org/10.1287/moor.1040.0129

Krakovna, V., Orseau, L., Martic, M., & Legg, S. (2019). Penalizing side effects using stepwise relative reachability. *AI Safety Workshop at IJCAI*. https://ceur-ws.org/Vol-2419/paper_1.pdf

Nilim, A., & El Ghaoui, L. (2005). Robust control of Markov decision processes with uncertain transition matrices. *Operations Research, 53*(5), 780–798. https://doi.org/10.1287/opre.1050.0216

Turner, A. M., Hadfield-Menell, D., & Tadepalli, P. (2020). Conservative agency via attainable utility preservation. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 385–391. https://doi.org/10.1145/3375627.3375851

Original research artifacts (Holmström, 2026; unpublished specifications and frozen computational results): *Experiment A1 — Specification v0.2.2*; *Experiment B — Future Possibility Optimization, Design v0.2*; *Experiment B5 — Learning Goals versus Preserving Options, Design v0.2* and B5-0.1 amendment. These are the primary sources for all experimental numbers, whereas the external references establish conceptual and methodological precedents.