

Thanks for Nothing

Gratitude, Model Revisability, and the Practice of Realizing Emptiness

A source-grounded study of gratitude, absence and revisable representations. Synthetic experiments clarify enabling operations; human gratitude and realization require separate evidence.



Conceptual cover illustration generated with AI; not a scientific result.

Daniel Ari Friedman

Active Inference Institute

ORCID: 0000-0001-6232-9096

daniel@activeinference.institute

Code and reproducibility materials: https://github.com/docxology/thanks_for_nothing

DOI: [10.5281/zenodo.23402517](https://doi.org/10.5281/zenodo.23402517)

Research manuscript, version 0.11.0 2026-10-22

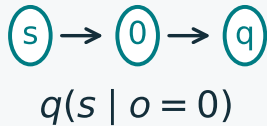
Paper: [CC BY 4.0](#); original code: MIT. Third-party rights remain separate.

What can ‘nothing’ mean?

Evidence, availability, counterfactual value and revisability answer different questions. Appreciation does not make them interchangeable.

Recorded absence

What did a functioning channel observe?

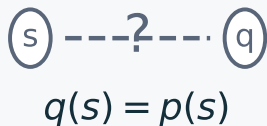


A recorded absence can be evidence when the channel discriminates between possible states.

A no-event record differs from missing data and from a claim that nothing exists.

Missing observation

Was a usable record available?

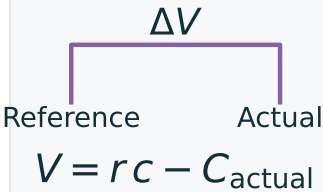


The missing-record sentinel supplies no observation likelihood in the declared implementation.

Ignoring missingness requires assumptions; other missingness processes can be informative.

Counterfactual appreciation

Against which alternative is an outcome valued?

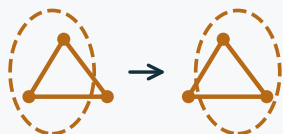


An uneventful outcome can acquire value relative to an avoided harm without supplying favorable evidence.

The valuation-only control preserves belief updates and predictive losses by construction.

Revisable representations

Can a useful distinction be examined and conditionally retained?



Conditional use

The finite counterpart tests access, retention and revision; it does not measure realization of emptiness.

Non-reification concerns how a representation is held, beyond recording an absence.

An enabling chain, not an attainment scale



Symbols: s is latent state, o observation, p prior and q posterior. In the value illustration, r is reference-event probability, c its cost and C_{actual} the actual cost. Textual traditions remain situated comparisons, not equivalent doctrines; see the atlas and edition notes.

Conceptual proposal → finite enabling operations → a separate human research obligation.

From mechanisms to bounded claims

Source reconstruction, corrected comparisons and new extensions remain separate. Recall cohort: 5120 independent agent-world units in 5 declared strata.

A: 3 x 2	Recall and independent prediction Matched exposure and replay dose separate positive selection from learning and independent prediction. Optimism shifts and predictive accuracy differ; effects depend on the world and comparator.
A: 2 x 2	Recorded absence and counterfactual value A no-event token can update belief; the missing-record rule supplies no likelihood. Valuing an avoided harm changes assigned value without creating favorable evidence.
A: 3 x 13	Native attention and learned-channel transfer Native GNN/pymdp runs distinguish agent and process; transfer bridges remain separate. The source supplies before-only gaze. A favorable attention shift need not improve prediction.
A: 2 x 3	Enacted inquiry and effective access Paid activity, effective access and evidence-based retention or revision are distinct operations. Blocked access can incur cost without revision; exact enumeration retains adverse selections.
A: 3 x 9	Adaptive scene sampling Complete query-conditioned GNN channels expose a context-dependent access advantage. Already-open and uniform controls do not establish a further inquiry-specific benefit.
A: 9 x 144	Joint scene and encoding inquiry Spatial and contingency queries test joint inference over the scene and its encoding. Equivalent laws remain unidentified; excluded support stays unavailable; the reference anchor is stipulated.

Execution and evidence remain inspectable



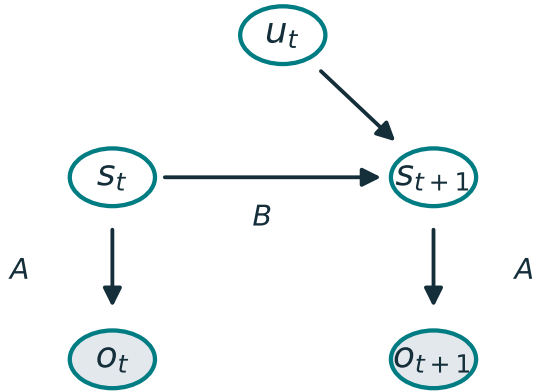
Glyphs unpack the admitted likelihood: s or z is hidden, o observed and q a known query. Dimensions come from executed GNN. The world glyph is one declared stratum; complete models and controls follow.

Synthetic model evidence does not establish human gratitude, clinical benefit, contemplative realization or philosophical equivalence.

From declarations to executed models

A time-unrolled POMDP view, parsed dependencies and the complete consumer inventory.

BAYESIAN TIME SLICE / conditional on the selected action

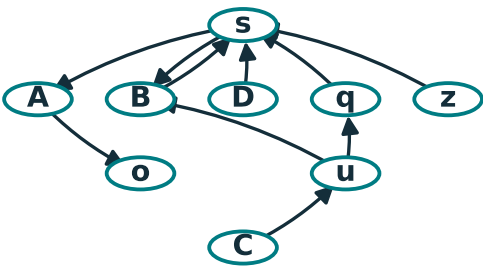


A observation | state
B next state | state, action
D initial-state prior
C log outcome preferences

Shaded: observations. Open: latent states and the conditioned action. D initializes the first state; C guides native policy selection, outside this conditional DAG.

$$P(s_{0:T}, o_{0:T} \mid u_{0:T-1}, A, B, D) = D(s_0) \prod_{t=0}^T A(o_t \mid s_t) \prod_{t=0}^{T-1} B(s_{t+1} \mid s_t, u_t)$$

DECLARATION GRAPH / source structure, not an instantaneous causal DAG



26 declarations; 2 topology groups. Complete query graph shown. B advances time; q and z define the joint basis. Full models are linked below.

The Bayesian diagram defines the tensor semantics. Finite consumers add their own declared task laws; a shared backend label does not establish native execution.

AUTHORING INPUTS / complete executed instances in the linked cards

Declaration / consumer	A	B	C	D	SHA
analysis Protocol carrier	1 x 1	1 x 1 x 1	1	1	93ac17
gratitude Finite learner	3 x 2	2 x 2 x 1	3	2	a2945f
nothingness Finite comparators	2 x 2	2 x 2 x 1	2	2	5cfa88
attention_positive Native GNN / pymdp	3 x 13	13 x 13 x 13	3	13	1d47aa
attention_negative Native GNN / pymdp	3 x 13	13 x 13 x 13	3	13	2b5411
emptiness_practice Finite inquiry	2 x 3	3 x 3 x 1	2	3	e27328
gaze Finite scene inquiry	3 x 3	3 x 3 x 1	3	3	2a370a
boundary_gaze Finite joint inquiry	9 x 9	9 x 9 x 1	9	9	4d0b48

A: observation x state; B: next x previous x action. SHA: source-digest prefix only. Full tensors, settings, annotations, native graphs, dependency revisions and evidence digests are in the linked complete card.

Scope: finite synthetic software models. These graphs do not establish human gratitude, contemplative realization, clinical efficacy or philosophical equivalence.

Contents

Abstract	2
1 Introduction	2
1.1 Nothing as evidence, missingness, value and revisability	3
1.2 Emptiness and appreciation: a situated dialogue across traditions	4
2 Methods	8
2.1 Reconstructing sources and admitting executable models	8
2.2 Separating recall, recorded absence and counterfactual value	11
2.3 Testing the chain from appreciation to effective inquiry	11
2.4 Transferring learned channels to active scene sampling	11
2.5 Joining scene evidence with questions about its encoding	11
2.6 Defining estimands and testing uncertainty procedures	12
3 Results	12
3.1 Positive recall changes optimism without a general predictive benefit	12
3.2 Effective access permits revision and appropriate retention	14
3.3 Access advantages do not establish an inquiry-specific gaze benefit	16
3.4 Learning an encoding differs from using it for prediction	18
4 Discussion	20
4.1 Gratitude for the possibility of examining one’s own constructions	20
4.2 What gaze, confidence and first-person reports can establish	20
4.3 From finite mechanisms to an identifiable human research program	21
4.4 Conclusion: appreciate, examine and remain able to revise	21
Supplementary information	23
S1. Conceptual distinctions and the scholarship behind them	23
S2. Declared models, executable interfaces and independent controls	27
S3. Experimental design, estimands and uncertainty procedures	41
S4. Complete results, including adverse and unavailable estimates	45
S5. Mechanistic interpretation and competing explanations	78
S6. What these models do and do not identify	80
S7. Human research: manipulations, measurements and recovery checks	80
S8. Source custody, operating diagnostics and complete interval tables	81
S9. Model settings injected from admitted declarations	95
S10. Textual editions, translation choices and comparison limits	96
References	97

Abstract

Gratitude can concern a benefit received, a harm that did not occur, or the conditions that make one’s own assumptions examinable. We develop the proposal that appreciating these conditions can orient practice toward realizing emptiness: examining how distinctions arise, relinquishing their assumed intrinsic necessity, and retaining their conditional usefulness. The argument distinguishes an observed absence, a missing observation, counterfactual valuation and non-reification, and places their differences in dialogue with explicitly situated philosophical and scriptural sources. We reconstruct Fisher, Sandved Smith and Hohwy’s gratitude and gaze notebooks, then separate faithful reconstruction from corrected experiments and new mechanisms. GNN-declared paired recall experiments use 5120 independent agent-world units, matched environmental exposure and replay dose, normalized likelihoods, and independent predictive assessment. Pinned Realizing Emptiness interfaces supply structural admission and finite mathematical comparators. An enacted inquiry model separates appreciation, costly activity, access to alternatives and evidence-responsive retention or revision; explicit learning-to-gaze and joint scene/encoding experiments examine transfer and identifiability. Exact count enumeration, mathematical nulls, paired controls and joint bootstrap inference make the resulting software claims independently checkable.

Selective positive recall reduces the declared optimism summary, while its predictive consequences depend on the world mixture and the specified transfer bridge. Counterfactual absence valuation preserves inference by construction. Accessible inquiry can revise a contradicted restriction and retain a useful one; activity without access instead incurs costs while preserving rigid forecasts. Exact analysis exposes rare adverse selections absent from seeded realizations. The original gaze code reproduces the pre-intervention occupancy but contains no implemented after intervention. Our active-scene extension finds a context-dependent access advantage without an additional inquiry-specific predictive advantage. Joint scene/encoding inquiry distinguishes learning an encoding from combining it with scene evidence: observationally equivalent labels remain unidentifiable, and excluded hypotheses cannot be recovered. Known-answer fixtures also retain skew-related undercoverage and unavailable rare-event estimates rather than treating replication size as a remedy for every inferential failure. The contribution is a reproducible account of enabling operations and their limits, together with a falsifiable human research program that separates inquiry value from patience, anticipated good news and instrumental information demand. Synthetic model results establish neither human gratitude nor contemplative realization or clinical benefit. Code, declared models and reproducibility materials are available at github.com/docxology/thanks_for_nothing.

1 Introduction

A functioning channel can record no event (observed absence), whereas an unavailable record leaves an outcome unobserved (missing observation). An actual outcome can also be valued against a specified alternative (counterfactual value). The Realizing Emptiness proposal introduces a further task: examining a useful distinction’s assumed intrinsic necessity and practicing its conditional use (non-reification). These operations motivate different questions about gratitude for avoided harm and for the conditions that enable inquiry. Their practical and mathematical consequences organize this study.

Computational models provide a way to specify those consequences without assuming that their state variables are direct measurements of lived experience. Fisher, Sandved Smith, and Hohwy propose that gratitude practice can be represented as deliberate precision allocation to positive observations, followed by learning that changes world beliefs and an optimism-task readout (Fisher et al. 2025). Their preprint also presents a modified Yarus attention task as a proof of concept. A subsequent journal abstract describes the same broad research program (Fisher et al. 2026). We inspect the archived preprint in full and the journal publication at abstract level; detailed reconstruction claims concern the preprint and its associated notebooks, not an assumed identical journal full text.

The proposal is useful because it connects an everyday practice to explicit operations on inference and learning. It also makes the limits of that connection visible. Increasing the probability assigned to a favorable state is not the same as improving predictive accuracy. An optimism score is not a direct measure of well-being. Rehearsing an event is not collecting an independent observation, and changing a likelihood parameter by hand is not showing that the parameter was learned and transferred to another task. Those distinctions determine which claims a reproducible extension can support.

We ask what follows from a declared finite model when positive-event selection, replay dose, environmental exposure, inference, and evaluation are separated. We then ask whether appreciation of an absence can be represented without converting an unobserved or hypothetical event into favorable evidence. Finally, we develop the more demanding proposal that gratitude can orient an agent toward examining its own model-making. The title “Thanks for Nothing” holds these questions together: appreciation of contingent benefits and avoided harms, and appreciation of the opportunity to realize that conceptual constructions need not be intrinsically or compulsorily established.

The work has three levels of analysis. Source reconstruction preserves the numerical behavior of supplied notebooks and documents differences between that behavior and a normalized generative-model interpretation. Corrected experiments hold environmental exposures fixed across conditions, normalize likelihoods, match replay dose, and evaluate frozen models on independent observations. New extensions distinguish observed absence from missingness, assign counterfactual value without changing beliefs, and enact a repeated inquiry policy over alternative representations. Keeping these levels separate prevents a correction from being reported as a faithful replication, or our proposed practice extension from being attributed to the original gratitude paper.

The resulting contribution is methodological and conceptual. A finite simulation can reveal how an intervention changes the model that was implemented, expose failures of an estimator under a known null, and make otherwise ambiguous uses of “absence”

explicit. Whether those operations characterize human gratitude remains an empirical question requiring construct validation, comparison activities, and outcomes beyond the simulation’s own favored readout.

The paper contributes an explicit appreciation–inquiry proposal, a modular finite counterpart imported through pinned GNN and Realizing Emptiness interfaces, and complementary evaluations through exhaustive evidence decisions, paired repeated-policy experiments and active scene sampling. It reconstructs both source notebooks, connects corrected learning to gaze through a declared likelihood embedding, and distinguishes structural access from information-seeking policy. The corrected recall study constrains favorable interpretations of selective rehearsal. Supplementary methods retain estimator failures and complete interval comparisons.

The psychological interpretation is a further research obligation. Construct validity concerns the testable relations supporting an interpretation, rather than the name attached to an operation (Cronbach and Meehl 1955). Calling a preference gratitude or a family change insight therefore cannot supply the missing evidence. Our contribution makes proposed operations discriminable; the validation account explains what would be needed to connect them to human constructs.

1.1 Nothing as evidence, missingness, value and revisability

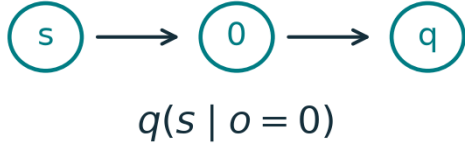
We use 4 stable terms, identified in Figure 1 and repeated in the task and result figures. They classify different questions rather than stages of an achievement scale. Here, observed absence concerns evidence; missing observation concerns availability; counterfactual value concerns evaluation relative to a reference; and non-reification concerns the status accorded to a representation.

Concept	Operation in this project	Evidential boundary
A: Observed absence	Update with the no-event likelihood.	Detection and exposure assumptions matter.
B: Missing observation	Retain the prior under our missingness rule.	State-dependent missingness needs its own model.
C: Counterfactual value	Assign value without adding an observation.	Reference-dependent value is not new evidence.
D: Non-reification	Examine its conditions; retain or revise its use.	Finite revisability does not measure realizing emptiness.

The distinctions also organize the scholarship. Goh and colleagues study contrastive silences within structured auditory contexts, rather than every possible absence (Goh et al. 2023). Missingness has a different inferential role: ignoring its mechanism requires explicit assumptions and an inferential target (Seaman et al. 2013). Our no-update sentinel is a declared modeling choice, not a consequence of calling a value missing.

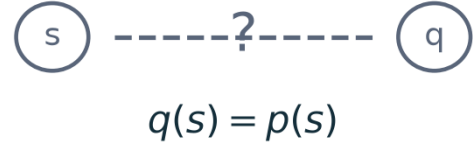
Counterfactual thinking can influence affect and action through different routes, including contrast with an alternative or assimilation toward an imagined event (Epstude and Roese 2008). Mental subtraction of an existing benefit is therefore a useful comparison for appreciation, while remaining distinct from detecting an absent harm or establishing gratitude (Koo et al. 2008).

A: Observed absence



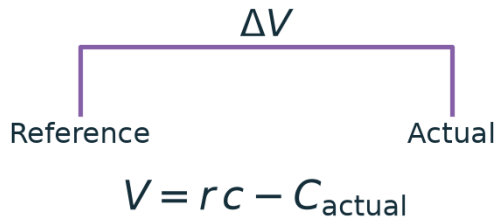
A working channel records no event.
Update with the no-event likelihood.

B: Missing observation



An observation record is unavailable.
Retain the prior under our missingness rule.

C: Counterfactual value



Actuality is compared with a specified alternative.
Assign value without adding an observation.

D: Non-reification



A useful distinction is not held as intrinsically fixed.
Examine its conditions; retain or revise its use.

Evidence, availability, value, and representation remain distinct.

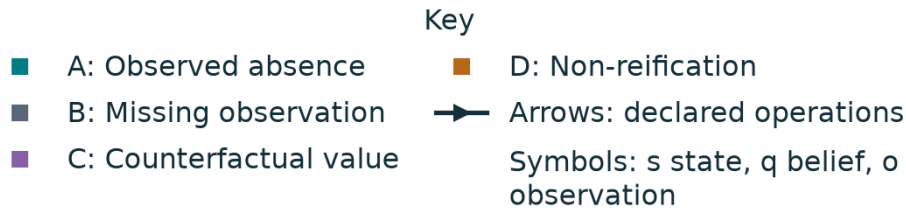


Figure 1: Different questions require different meanings of nothing. The 4 panels distinguish evidence, availability, reference-dependent value and the status of a representation. A: Observed absence records $o = 0$ through a functioning detector; B: Missing observation has no observation; C: Counterfactual value compares expected reference harm with actual cost. D: Non-reification redraws a self/world boundary while preserving useful dependencies. These are symbolic operations, with no sampled estimates or confidence intervals. Realizing Emptiness motivates examining assumed necessity; the diagram supplies neither a realization score nor evidence that appreciation causes inquiry.

1.2 Emptiness and appreciation: a situated dialogue across traditions

1.2.1 What is absent, what remains and what depends on conditions

Samyutta Nikāya (SN, the Connected or Grouped Discourses) and *Majjhima Nikāya* (MN, the Middle-length Discourses) are collections of early Buddhist scripture in the Pali Sutta Piṭaka. Their numbers identify discourses, not scientific measures. In *Kaccayanagotta Sutta: To Kaccayana Gotta (on Right View)*, SN 12.15, the Buddha is presented as teaching conditioned arising and cessation while rejecting opposed extremes. Ṭhānissaro Bhikkhu's translation says:

‘Everything exists’: That is one extreme.

The discourse also rejects the opposing assertion of universal non-existence (Thānissaro Bhikkhu 1997b, discourse paragraph beginning with the two extremes). In *Cula-suññata Sutta: The Lesser Discourse on Emptiness*, MN 121, the same translator

distinguishes what is absent from what remains:

Whatever remains, he discerns as present: ‘There is this.’

That passage occurs at the end of the wilderness-perception sequence; the later Release section examines refined concentration as conditioned and subject to cessation (Thanissaro Bhikkhu 1997a). Both quotations use the catalogued Access to Insight editions. These are translated textual arguments to interpret, not empirical evidence for a Bayesian model or a numerical definition of realization. The textual supplement identifies the editions and translation-sensitive qualifications.

A further discourse specifies the object of the negation. In *Suñña Sutta: Empty*, SN 35.85, sensory and mental domains are described as empty of self or what pertains to self (Thanissaro Bhikkhu 1997c). That qualification differs from claiming that no event occurred or that no observation was recorded. It constrains our interpretation of non-reification without making the finite support comparison a realization criterion.

The Mahāyāna *Heart Sutra* supplies a different textual setting. In Thich Nhat Hanh’s explicitly revised English rendering, the first address to Sariputra reads:

this Body itself is Emptiness and Emptiness itself is this Body.

The wording is specific to this contemporary translation, whose accompanying explanation identifies deliberate changes intended to avoid nihilistic interpretations (Thich Nhat Hanh 2014). Its paired formulation makes it difficult to treat emptiness as a separate object that abolishes embodied phenomena. It also cautions against treating a variable named empty as an adequate philosophical account. This is a comparison between differently situated texts, not a claim that the early discourses and all Mahāyāna traditions articulate an identical doctrine. The edition notes distinguish the translated passage from its translator’s explanatory argument.

1.2.2 Using distinctions without treating them as intrinsically necessary

The raft simile in *Alagaddupama Sutta: The Water-Snake Simile*, MN 22, distinguishes relying on a useful means of crossing from continuing to carry it afterward (Thanissaro Bhikkhu 2004, The Raft Simile). This motivates a question about appropriate reliance rather than indiscriminate rejection. Our retention controls test the narrower computational question of whether a working predictive family remains useful; the discourse’s liberative purpose is not reduced to predictive loss.

Lutz and colleagues distinguish object orientation, meta-awareness and dereification within a provisional phenomenological framework (Lutz et al. 2015). Noticing a construction and ceasing to take it as intrinsically necessary are therefore different hypotheses. Philosophical non-reification also requires identifying what is negated. Garfield and Thakchöe distinguish the rejection of intrinsic nature from the status of conventional phenomena, while comparing interpretations that differ on that status (Garfield and Thakchöe 2011). Our Realizing Emptiness connection concerns examining assumed necessity while retaining conditional usefulness. The finite model makes an enabling operation explicit; the complete philosophical comparisons and construct boundaries remain in the scholarship supplement.

Nāgārjuna’s *Mūlamadhyamakakārikā* (MMK) gives this caution a more specific philosophical basis. In our descriptive paraphrase of the inspected Sanskrit witness, the verse on dependent arising connects dependent arising, emptiness and dependent designation in its account of the middle way; the preceding discussion distinguishes two truths and makes reliance on convention necessary for teaching the ultimate (Nagarjuna 2020, MMK 24.18 and 24.8–10). This is our paraphrase, not a quotation from an English translation. Resolving the pronouns in the dependent-arising verse and determining what dependent designation entails require interpretation, for which the differing accounts discussed by Garfield and Thakchöe remain relevant (Garfield and Thakchöe 2011). The comparison therefore concerns how distinctions are held and used. It does not identify conventional truth with held-out predictive accuracy or make successful model selection a criterion of realization.

1.2.3 From gratitude to benefactors toward appreciation of enabling conditions

Our constructive proposal is gratitude for the conditions that permit examination, for an insight that changes how a distinction is held, and for the continuing possibility of revision. Appreciation can precede inquiry, follow it, or support its continuation. Those are different temporal relations. The *Kataññu* discourses in the *Aṅguttara Nikāya* (AN, the Numerical Discourses) connect gratitude with benefactors and ethical repayment, including parental care (Thanissaro Bhikkhu 2002). Appreciation of impersonal inquiry conditions is our constructive extension, rather than a translation of that benefactor-centered account. Sandved-Smith and colleagues’ account of examining an assumed self/world separation motivates the computational bridge (Sandved-Smith et al. 2026). The model implements access to alternatives and conditional use of a representation; it does not establish that appreciation produces contemplative realization.

The capacities that permit appreciation can themselves become its object. Epictetus considers perception and understanding within an argument for divine providence (Epictetus 1887, I.6, pp. 19–21). A Jewish liturgical comparison appears in the *Babylonian Talmud*: the blessing on opening one’s eyes includes “Blessed...Who gives sight to the blind.” The wording follows the annotated William Davidson English edition (Berakhot, n.d., Berakhot 60b:5). Both passages make an ordinary capacity salient without requiring a newly acquired possession. Their providential and liturgical meanings remain distinct from our declared value for examinability. In a human study, appreciation of a capacity, attribution to a benefactor and willingness to use that capacity would require separate assessment.

1.2.4 Useful hollowness, nonattached action and responsibility

The *Daodejing* offers a direct comparison for the usefulness of an unfilled condition. James Legge's translation states:

Clay is fashioned into vessels; but it is on their empty hollowness, that their use depends.

The surrounding wheel, vessel and room examples relate constructed form to what its unfilled space enables (Laozi 2008, ch. 11). Appreciating a useful hollow therefore need not mean preferring less structure everywhere. The arrangement matters to the use. This helps articulate a functional question about absence while leaving Daoist *wu*, *wuwei* and Buddhist *śūnyatā* distinct. Chan's discussion also cautions against reducing *wuwei* to mere inactivity (Chan 2025, secs. 6–7). Neither the vessel example nor our inference-preserving valuation condition establishes a shared ancient gratitude technique.

Action without a demand for a particular result supplies another comparison. Swami Mukundananda translates the *Bhagavadgītā* as follows:

You have a right to perform your prescribed duties, but you are not entitled to the fruits of your actions.

The same verse ends by rejecting attachment to inaction (Swami Mukundananda, n.d.-a, chap. 2, verse 47). Its duty-oriented and theological setting matters: nonattachment to fruits is not a definition of Bayesian uncertainty or Buddhist emptiness. For our proposal, it sharpens the distinction between valuing an opportunity to inquire and insisting that inquiry confirm a welcome answer. An inquiry policy may choose a costly examination whose evidence supports retention, revision or an unfavorable forecast. Evaluating these outcomes independently makes that distinction experimentally consequential without representing the model as an enactment of the *Gītā*.

The *Gītā* also resists reducing response to an internal feeling. In the verses on sacrificial cooperation, Mukundananda's translation places received gifts within a sacrificial relation involving humans and divine participants; enjoying gifts without return offerings is censured (Swami Mukundananda, n.d.-b, 3.10–13). Read alongside the earlier verse on duty and its fruits, these verses distinguish relinquishing entitlement to results from relinquishing responsibility to act. Their *yajña* is not generic ecological interdependence or a computational exchange rule. The comparison instead asks what follows recognition of a benefit: possession of a favorable attitude, a costly practice, an obligation to another, or some combination. Our finite inquiry policy addresses only the relation between valuing access, paying a cost and using available evidence; it implements neither sacrifice nor interpersonal repayment.

1.2.5 Contentment and humility without relinquishing reciprocal obligations

Appreciation need not take a new acquisition as its threshold. Ben Zoma's answer about wealth in *Avot*, in Joshua Kulp's translation, is: "Who is rich? He who rejoices in his lot" (Kulp, n.d.-b, 4:1). The saying invokes a psalm on labor's fruit and the enjoyment of labor's fruit; its setting includes learning, self-command and honoring others. It redescribes wealth through a relation to what is available rather than an increment in possessions. Rejoicing is not automatically gratitude to a benefactor, and the selected clause supplies no addressee. Nor does appreciation of an existing lot justify deprivation or injustice. This witness helps separate contentment, gratitude and the counterfactual valuation of an avoided harm before attempting to measure any of them.

Guru Nanak's *Japji Sahib* supplies a distinct theological comparison. Sant Singh Khalsa renders the conclusion of the second pauni: "O Nanak! One who understands His Command, does not speak in ego." Walking in divine Will in the preceding pauni supplies immediate practical context in the *Sri Guru Granth Sahib* (Guru Nanak, n.d., ang 1, paunis 1–2). Understanding *hukam* here qualifies self-assertion within an order one does not author. It is not an explicit instruction to feel gratitude, and neither *hukam* nor *haumai* can simply be replaced by natural law or a Bayesian self-model. Its relevance to our question is responsiveness without claiming authorship or mastery, rather than a vacant object to be celebrated. The text's Creator, grace and practical orientation remain part of that comparison.

A comparison should also admit an ethically relevant case without an explicit grateful response. In the *Analects*, Zigong asks for a rule of lifelong practice, and the answer specifies *shu*, reciprocity: "What you do not want done to yourself, do not do to others." The inspected Legge edition and Eno's teaching translation give the exchange different chapter numbers (Confucius 2002, XV.23; 2015, 15.24). Here a consequential absence is the unwanted imposition one refrains from making. Reciprocity constrains one's conduct; it does not wait for another person to repay a benefit. This ethical comparator supplies neither an emptiness doctrine nor a gratitude exercise. Keeping it in the comparison tests whether the phrase Thanks for Nothing can invite a useful question while leaving a passage's different answer intact.

1.2.6 Thanksgiving amid adversity: attending to the addressee

Theistic traditions bring the addressee and obligations of appreciation into focus. The *First Letter to the Thessalonians*, in the New Revised Standard Version Updated Edition, instructs:

give thanks in all circumstances, for this is the will of God in Christ Jesus for you.

The wording situates thanksgiving in circumstances and a relationship to God (New Revised Standard Version Updated Edition 2021, 5:18). Thanksgiving need not await a favorable outcome, but this sentence does not by itself say that every harmful event is a benefit. That distinction matters when interpreting appreciation of an avoided harm: favorable evaluation of a counterfactual difference cannot establish how someone should respond to an actual loss.

The cited paragraph of *Mishnah Berakhot*, in Joshua Kulp's translation, requires blessing for evil as for good (Kulp, n.d.-a, 9:5). The later Talmudic discussion preserves different formulas for good and bad tidings (Berakhot, n.d., Berakhot 60b:7). Blessing

across circumstances therefore does not erase the distinction between those circumstances. The comparison poses a question about the response and its divine addressee, rather than converting adverse evidence into favorable evidence. Blessing, gratitude, praise and acceptance can be related within a practice while remaining different candidate constructs for research.

In Mustafa Khattab’s *The Clear Quran*, a divine proclamation reads, “If you are grateful, I will certainly give you more.” The verse also warns against ingratitude; its surrounding account recalls deliverance and God’s independence from human gratitude (Khattab, n.d.-a, Sūrat Ibrāhīm 14:6–8). This religious relation cannot be replaced by a numerical optimism preference. Nor is the promise a quantified experimental law that appreciation produces wealth, predictive accuracy or clinical benefit. A comparison attentive to *shukr* asks about recognition, response and accountability, while preserving the theological setting of the translated passage.

The beneficiary and addressee of gratitude need not have symmetrical needs. In the same translator’s rendering, a verse in Sūrat Luqmān directs gratitude to Allah while locating its good in the grateful person and affirming divine self-sufficiency (Khattab, n.d.-b, 31:12). This complements, rather than first introduces, the qualification already present in Sūrat Ibrāhīm (Khattab, n.d.-a, 14:8). Gratitude is not represented as supplying something God lacks, yet obligation and accountability remain. That is a different relation from the *Gītā*’s return offerings, the *Analects*’ restraint and *Avot*’s rejoicing in one’s lot. Calling all of them positive valuation would conceal who responds, to whom, through what practice and for what reason. Those distinctions would need to enter a human study’s instructions and measures rather than being inferred from a common favorable score.

1.2.7 Apophatic language and the limits of computational comparison

Christian apophatic theology provides a further comparison concerning the limits of predication. In C. E. Rolt’s translation of *Mystical Theology*, Pseudo-Dionysius writes of the Cause: “It transcends all affirmation by being the perfect and unique Cause of all things”; the same final sentence also places it beyond negation (Pseudo-Dionysius the Areopagite, n.d., V). This concerns a transcendent divine Cause, not the same object of negation as Madhyamaka’s critique of intrinsic nature. Its relation between praise and the removal of predicates (Pseudo-Dionysius the Areopagite, n.d., II) nevertheless shows why appreciating something and claiming to describe it exhaustively are separable operations. Removing a predicate in this theological argument is not equivalent to deleting a hypothesis from a statistical model.

These comparisons refine the research questions through their differences. A useful absence concerns what an arrangement enables; nonattachment concerns one’s relation to action and its fruits; thanksgiving and blessing raise questions of addressee and obligation; non-reification concerns the status accorded to a distinction. None licenses replacing another tradition’s terms with a shared latent variable called wisdom. Our constructive bridge is narrower: appreciate the conditions of examination, act without requiring a preferred finding, and keep useful distinctions conditionally available. The comparative crosswalk, textual scope notes and prospective empirical design distinguish this proposal from historical identity, doctrinal consensus and demonstrated human effects.

1.2.8 Reading the atlas as a comparison of textual witnesses

Figure 2 places these witnesses beside one another under two questions: what is absent or relinquished, and what appreciation or response is articulated? Its cells are our interpretive summaries, not quotations or measurements. The selection follows the manuscript’s questions and the availability of verified named witnesses; it is not a census of world traditions, a geographical distribution or a proposed historical lineage. Separate Buddhist and Christian rows retain different textual problems within those broad labels. A passage that does not specify gratitude is marked accordingly in the companion, rather than being used to claim that its tradition lacks gratitude. Translator, locator, source link and interpretive limit travel with each row in the searchable source companion; the textual supplement and comparative crosswalk retain the surrounding argument.

The atlas exposes a limit of the computational bridge as well as a motivation for it. A vessel’s hollow, a relinquished entitlement, a withheld imposition and a divine Cause beyond predicates are different candidates for what nothing could mean. The ensuing response can be discernment, duty, rejoicing, blessing, restraint or praise. Our model asks only whether valuing the opportunity for examination can support costly inquiry while independent evidence governs retention or revision. The texts neither implement nor validate that mechanism. Their contribution is to discipline its interpretation and suggest distinctions a future empirical study would have to preserve.

Thanks for Nothing: distinct textual representations

Textual witness	Absence / relinquishment	Appreciation / response
Early Buddhist MN 121; SN 35.85	Empty of what? What remains is present.	Discern absence and remaining conditions.
Mahāyāna Heart; MMK 24.18	Emptiness is not a separate Nothing.	Retain phenomena and conventional discourse.
Daoist Daodejing 11	Unfilled hollow within a formed vessel.	Recognize what an arrangement enables.
Hindu / Gītā 2.47; 3.10-13	Relinquish entitlement to action's fruits.	Act; respond to gifts through offerings.
Stoic Epictetus 1.6	No new possession is required here.	Attend to perception and understanding.
Jewish Avot 4:1; Ber. 9:5	New acquisition need not define wealth.	Rejoice in one's lot; distinguish blessings.
Christian 1 Thess. 5:18	No favorable outcome is a prerequisite.	Give thanks in circumstances, toward God.
Islamic Qur'an 14:7; 31:12	God's need for repayment is denied.	Thank Allah; the grateful person benefits.
Sikh Japji, pauri 2	Egoic speech is relinquished.	Understand hukam; walk in divine Will.
Confucian Analects 15.24	Refrain from unwanted imposition.	Practice reciprocal restraint toward others.
Christian apophatic Mystical Th. II, V	Exhaustive predicates are relinquished.	Praise a Cause beyond affirmation / negation.

Selected witnesses; interpretive summaries, no shared measured construct.

Key

Rows: selected textual witnesses

Columns: questions, not measured quantities

Figure 2: Distinct textual responses to absence and relinquishment. Selected named textual witnesses are aligned by what is absent or relinquished and what appreciation or response follows. Read each row from the named witness and passage locator to the two interpretive summaries. Alternating shading guides row tracking and encodes no grouping or rank. Separate Buddhist and Christian rows retain different textual settings; the source companion gives translators, links, addressees, practices and limits. This is an editorially selected qualitative comparison, with no sampled estimates, shared scale, ranking or confidence intervals. Gratitude is not explicit in every passage. These witnesses neither exhaust their traditions nor establish doctrinal equivalence, historical influence or validation of the finite computational model.

2 Methods

2.1 Reconstructing sources and admitting executable models

We preserve the supplied gratitude and modified Yarus notebooks, reconstruct their available numerical behavior, and distinguish those reconstructions from corrected comparisons and new mechanisms. Every new categorical model and experimental design is admitted through packaged Generalized Notation Notation (GNN) declarations. The imported GNN version is 4.1.0

and the Realizing Emptiness version is 3.0.0, with exact commits and executing-module digests retained in the provenance record. A current-upstream check is documented separately; reproducible execution continues to use immutable pins.

The experimental overview separates each question, controlled manipulation, independent assessment and decisive limitation. The detailed task illustrations define the unknowns, observations, available actions and assessment targets. The diagrammed agents are finite models. Painting regions and locations are authored registration choices; the models do not infer scene semantics from pixels. **The full formal methods** retain every equation, likelihood, prior, policy, access intervention and independent oracle. **The task crosswalk** gives the same overview as a searchable table and preserves the concrete detector, timeline and query diagrams.

The imported interfaces have bounded responsibilities. GNN supplies the actual parser and admission of the declared categorical arrays; Realizing Emptiness supplies structural admission and finite numerical operations. We separately check those operations against mathematical controls whose targets do not move with the experimental parameters. This separates a dependency being called from its output being correct for the stated task. It also leaves the interpretation open to scrutiny: admitting a candidate family is a software operation, whereas identifying that family with a person’s self/world representation would require additional evidence. **The interface account** and **model-role notation** specify these boundaries before the experimental comparisons.

Checking a model family is also distinct from updating within it. Gelman and Shalizi argue for a cycle in which discrepancies can motivate changing the model, beyond posterior calculation inside an assumed family (**Gelman and Shalizi 2013**). Here independent controls test specified software and finite-model implications; they do not constitute an empirical check of a human gratitude theory. The implemented access gate exposes authored alternatives, while constructing a better candidate family remains outside the model.

Manipulate / match

Assess after choice

Control / limit

a Recall: does selected replay improve prediction?

Positive vs random replay; same exposure and matched replay dose.

1024 paired units/world; frozen Brier and optimism tests.

Value-only control: unchanged beliefs and predictive losses.

b Absence: no event, no record, or a different value?

Functioning detector vs missing record; actual vs reference cost.

Posterior change in bits; separate counterfactual cost value.

No record adds no evidence under the declared rule.

c Practice: does chosen activity provide usable access?

Value and cost affect activity; 24 calibration observations.

512 new outcomes/window; 512 paired units/context.

Sham or blocked access: activity can cost without revising.

d Transfer: what follows from a learned likelihood?

Freeze learned channels; compare likelihood and readout bridges.

5120 complete paired agents; attention and scene loss separately.

A favorable attention shift need not improve prediction.

e Scene inquiry: access or the sampling policy?

Restricted vs open; utility vs inquiry vs uniform; 48 decisions.

512 paired units/context; scene loss and movement cost.

Already-open comparators test an inquiry-specific increment.

f Boundary inquiry: can the inferred frame be used?

Spatial vs action-response queries; one 24-decision budget.

512 paired units/cell; encoding and reference losses.

Label-only, excluded support and matched-query erasure controls.

Key

■ Manipulation / matched input

■ Control or evidential limit

■ Independent assessment

Figure 3: What each experiment changes, scores and can rule out. Read each row from its question through the controlled comparison to its assessment and decisive limitation. Recall, absence/value, enacted inquiry, learned gaze transfer, active scene sampling and joint scene/encoding inquiry are separate experiments. Budgets and independent-unit descriptions come from the admitted designs; a new observation, a replayed event and a structural question have different roles. This is a design map, not an effect-size display; the main Results and complete supplementary tables report the estimates and uncertainty. The comparison column identifies software controls that constrain interpretation. No row measures human gratitude, eye movements or contemplative realization.

2.2 Separating recall, recorded absence and counterfactual value

The **graphical-model inventory** connects parsed GNN nodes and edges to admitted tensor dimensions and the actual numerical consumers. Shared graphical syntax does not imply shared execution or empirical interpretation.

The corrected recall study contains 5120 independent agent-world units: 1024 replicates in each of 5 fixed world strata. Each unit receives the same environmental exposures under 4 paired conditions. Selected positive recall is compared with no recall, dose-matched random recall and valuation without new belief learning. Training, replay, optimism-task readout and independent predictive assessment remain separate. Binary Brier loss and observation log loss are evaluated alongside the source-inspired optimism summary.

Each comparator addresses a different explanation. No recall supplies the ordinary-learning reference; random recall matches the amount of replay while changing its selection rule. The valuation-only condition changes the assigned value of an uneventful outcome without supplying a new observation or changing the learning path. Sharing environmental exposure within a unit prevents a different history of good and bad events from explaining the contrast. Assessment uses new observations after learning is frozen, so further exposure, a favorable readout and better prediction cannot be counted as the same outcome. The **paired-exposure protocol** describes the shared histories and random streams, while the **scoring definitions** distinguish latent-state Brier loss, observation log loss and the optimism task.

The absence experiment separates observed absence from missing observation: a functioning detector’s no-event outcome is distinct from an unavailable record. We assign counterfactual value through a separate reference calculation. Thus a record can change a posterior because the channel discriminates states, while a value can change because the chosen comparison changes. Neither operation licenses treating an unavailable event as observed favorable evidence. The **absence model** states the detector likelihood, missing-record convention and reference-value calculation separately.

2.3 Testing the chain from appreciation to effective inquiry

The practice experiment supplies a finite enabling counterpart to non-reification and uses 512 independent paired replicates per context and 12 windows. A policy can value examination, pay its cost and gain access to alternative predictive families; a sham can incur the activity cost without that access. Evidence determines conditional retention or revision, followed by 512 new assessment observations per window. Stationary and reversal contexts distinguish a restriction that remains useful from one that becomes inadequate. The open family is selected only when its evidence exceeds that of the restricted family; a tie retains the restriction. The selected forecast then faces assessment data that did not select it. The **enacted-policy model** and **Figure 13** make the action, access, evidence and assessment stages explicit. Prediction and expenditure remain separate outcomes.

This separation also informs the prospective measurement plan. Process-evaluation guidance distinguishes implementation, mechanisms of impact and context (Moore et al. 2015). Logging an activity alone therefore leaves its delivery and consequences unresolved. In this finite model, activity and informative access are distinct recorded quantities; their mapping to human measures remains a proposal. The **process-measurement argument** specifies that analogy and its limits.

Figure 10 retains the detailed mappings and distinct units in the technical supplement.

2.4 Transferring learned channels to active scene sampling

The source gaze code provides before-condition reconstruction only. New learned-channel transfer maps frozen recall channels to scene sampling through explicitly declared bridges. The active scene task uses 13 registered locations and 48 decisions per episode, with 512 paired replicates per context. Already-open utility and uniform sampling are active comparators; access to a larger candidate menu is distinguished from a preference for information.

Transfer is an authored hypothesis about how tasks are connected. The likelihood bridge transports the learned observation channel; the alternative readout bridge uses a different projection of the same learning result. Holding the learned channels fixed makes their downstream consequences comparable without presenting either bridge as empirically identified. Scene-wide predictive loss is evaluated separately from occupancy of positively labeled locations and from movement cost. The **transfer construction** specifies both bridges, and the **active-inquiry model** distinguishes opening the candidate menu from changing how queries are selected.

2.5 Joining scene evidence with questions about its encoding

The intersection supplies a further finite counterpart to non-reification by joining an unknown scene with an unknown predictive encoding. A policy chooses a spatial query or a nonspatial action-response question. The diagnostic regime contains encoding information in the joint action-response relation even though the separate marginals are uniform. The label-only regime removes that information. All 6 policies share the same potential query outcomes within each independent unit. Restricted support and a yoked sham test specific failure modes.

Figure 12 in the supplement specifies query budgets, conditional channels, and separate assessment targets.

The intersection contains 9216 independent paired units in 18 fixed cells, generating 55296 policy episodes with 24 dependent decisions each. The anchored prediction target stipulates a reference-frame intervention; it is not ordinary future prediction under an unchanged process. The sham follows the integrated policy’s query schedule but discards action identity and ignores the schedule likelihood. Its null is conditional on that updater, not proof that no information is available to any observer.

The common decision budget makes the competition between evidence routes visible. Spending a decision on an encoding question leaves fewer spatial observations, and learning the encoding alone need not identify the scene. We therefore retain separate scores for encoding judgments, the current scene and prediction under the stipulated reference intervention. Allocation and expenditure explain what a policy did without replacing those assessment targets. [The joint observation law](#) and [reference-target definition](#) state which information each route can supply and why the anchor is necessary.

2.6 Defining estimands and testing uncertainty procedures

Complete paired units are resampled within fixed strata; decisions, windows and policy conditions do not enlarge the independent sample. The recall family has 8 primary contrasts. Gaze and intersection analyses are separate exploratory families of 58 and 84 members. Joint rectangles use 16383 shared bootstrap maxima and retain a conditional 99% Monte Carlo quantile bracket. Proved constants, stochastic estimates and unavailable sample constants receive different statuses. Known-answer skew and rare-event controls retain undercoverage and unavailability.

The target is an equally weighted mean over declared world strata, not a population effect for people or naturally occurring environments. Changing the world weights changes that target, even when every within-world result is held fixed. Within each stratum, resampling preserves the paired outcomes and their shared exposures. This follows the simulation-study distinction between the generating mechanism, estimand, method and performance measure ([Morris et al. 2019](#)). [The weighting analysis](#) treats alternative mixtures as sensitivity scenarios, while [the joint-resampling methods](#) preserve dependence across contrasts rather than combining separate intervals after the fact.

Independent replication and resampling answer different precision questions. More simulated agents reduce Monte Carlo error for a declared generating process; more bootstrap draws refine the conditional numerical approximation to an interval. Neither supplies additional human observations. We therefore report [Figure 14](#) with an explicit standard-error calculation, rather than treating a visually stable trajectory as proof of precision ([Koehler et al. 2009](#)). [The simulation-design scope](#) separates this diagnostic from sensitivity across model settings.

An interval procedure can also fail to return an interval. Under our explicit rule that any zero-variance bootstrap resample makes studentization unavailable, additional resamples can reveal more such failures. We derive an [exact finite Bernoulli availability benchmark](#) and compare it with [Figure 40](#). Availability, coverage across all experiments and coverage among available intervals answer different questions; excluding failures from the denominator can conceal the practical limitation ([Morris et al. 2019, secs. 3.4, 5.1](#)).

Uncertainty has more than one source. Replicate variation concerns the simulated units; the Monte Carlo bracket concerns estimation of the bootstrap critical value conditional on those units. Neither addresses uncertainty over the authored model family or its human interpretation. A mathematical identity can establish an exact null, but an accidentally constant finite sample cannot justify inventing a stochastic interval. [The interval and decision rules](#), [intersection uncertainty](#) and [known-answer operating checks](#) retain these distinctions, including cases in which a nominal confidence procedure performs poorly.

[Supplementary statistical methods](#) specify pointwise intervals, fixed-stratum weighting, sign-flip assumptions, multiplicity and the representable-translation regression. [Complete interval tables](#) retain every reported comparison. A generated variable ledger binds each changing manuscript value to its producer; model settings and numerical findings are injected into both narrative and captions. Mathematical identities, scripture identifiers, historical source locators and bibliographic years are fixed references, not fitted model variables.

3 Results

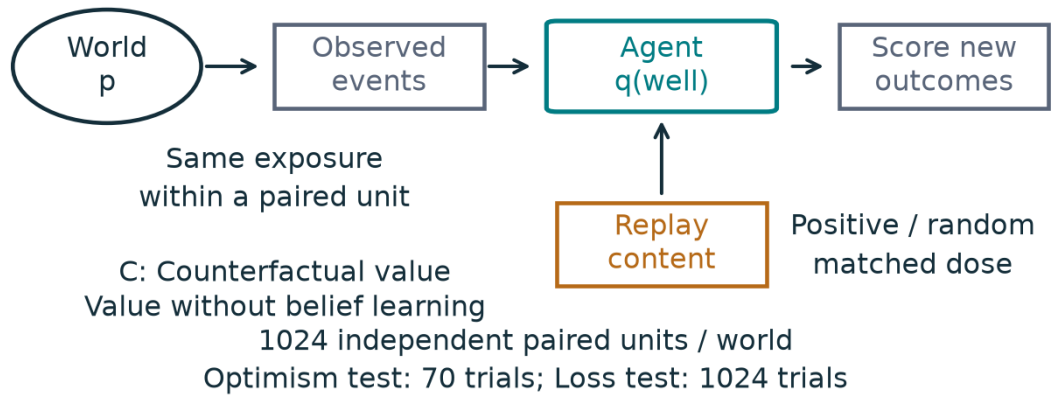
Each main result pairs the cognitive situation with selected readouts. The supplement retains the complete comparisons, bar charts, trajectories and operating checks. Schematic nodes and arrows describe the declared computational task, rather than observed psychological or contemplative processes.

3.1 Positive recall changes optimism without a general predictive benefit

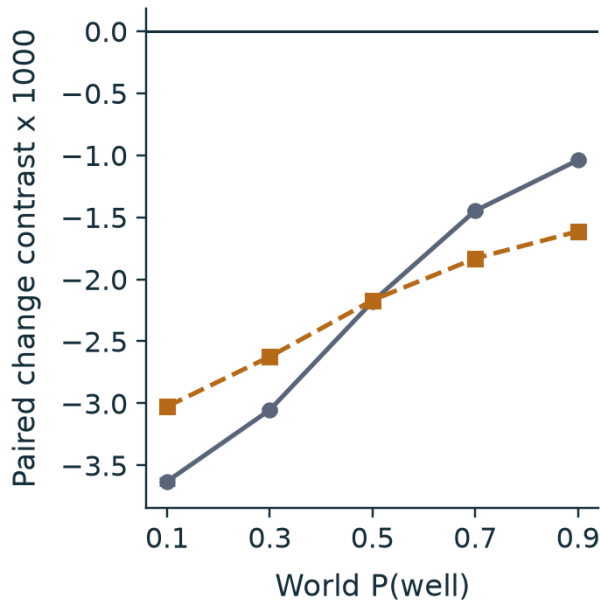
The source adapters expose implementation-sensitive labels, likelihood normalization and aliased snapshots. Corrected positive recall reduces the stipulated optimism summary. Its predictive consequences vary across the fixed world grid and the chosen comparison. Absence valuation leaves belief updates and predictive losses unchanged by construction. These findings concern the implemented operations, not a human well-being outcome.

a Paired experience, replay, and a new test

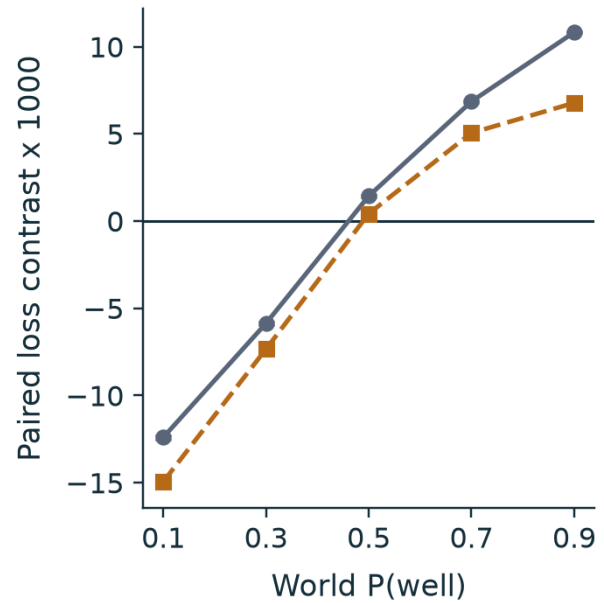
Test: does selected replay improve the model's independent predictions?



b Optimism readout



c Predictive Brier loss



Key

Positive minus no recall

Positive minus matched random recall

Arrows: declared learning / replay

Whiskers: pointwise analytic interval

Figure 4: Selective replay changes a readout without guaranteeing better prediction. Shared environmental exposure is held fixed within each paired unit; positive and random replay have matched doses, and frozen models face separate new tests. Panel a shows exposure, replay and independent assessment. Panels b (optimism) and c (Brier loss) show positive recall minus no recall (circles, solid) or matched random recall (squares, dashed) at each fixed world probability. Zero denotes no paired contrast. Both vertical axes multiply native differences by 1000; their scales differ. Whiskers are pointwise analytic 95% intervals over 1024 independent paired units per world; connections guide comparison across the fixed grid. Negative Brier differences favor positive recall; an optimism change is not a well-being effect. World-specific intervals are exploratory; pooled tests and every interval method remain supplementary.

Optimism-task change. Estimates and interval bounds are multiplied by 1000.

Contrast	Mean	Bootstrap t 95% interval	Holm p; MC decision
Positive - None	-2.270	[-2.278, -2.262]	0.00098; reject
Random - None	-0.017	[-0.031, -0.004]	0.23535; retain
Absence - None	0.000	Unavailable	1.00000; retain
Positive - Random	-2.253	[-2.267, -2.239]	0.00098; reject

Brier-loss change. Estimates and interval bounds are multiplied by 1000.

Contrast	Mean	Bootstrap t 95% interval	Holm p; MC decision
Positive - None	0.171	[0.135, 0.207]	0.44312; retain
Random - None	2.177	[2.125, 2.229]	0.00098; reject
Absence - None	0.000	Unavailable	1.00000; retain
Positive - Random	-2.005	[-2.066, -1.946]	0.00098; reject

The contrast with random recall is central to interpretation. A change relative to no recall can reflect additional replay as well as its positive selection; the dose-matched comparison asks what positive selection adds under the same exposure. Neither comparison makes the optimism readout an independent definition of benefit. The table’s Brier contrast concerns probability prediction, whereas the optimism contrast concerns the task’s authored readout. Their different signs and world dependence must therefore be read as different properties of the model, not averaged into a single favorable verdict.

Figure 20 shows why the fixed-grid average is incomplete. Figure 21 describes heterogeneity within each world, and Figure 22 shows how a different weighting changes the aggregate target. Those spreads and sensitivity ranges are not confidence intervals for a human effect. The complete recall results retain the interval procedures, while the source reconstruction distinguishes implementation findings from the corrected experiment.

The valuation-only null answers a separate semantic question. Assigned counterfactual value can vary while the paired beliefs and predictive losses remain exactly unchanged. Conversely, the absence detector can change beliefs when its recorded no-event outcome is informative. Figure 23 and Figure 24 display these distinct operations. The null validates their separation in this implementation; it does not show that appreciation and belief are independent in people.

3.2 Effective access permits revision and appropriate retention

Costly activity helps only when it makes a relevant alternative available. In reversal contexts the inquiry policy can select a predictor excluded by the restricted family; sham and blocked access preserve rigid forecasts while charging activity costs. A useful working restriction can be retained or reused. The exhaustive count calculation also identifies rare adverse selections that a finite seeded sample can miss.

a Practice changes access to a predictor

Test: can costly activity expose a useful alternative before assessment?

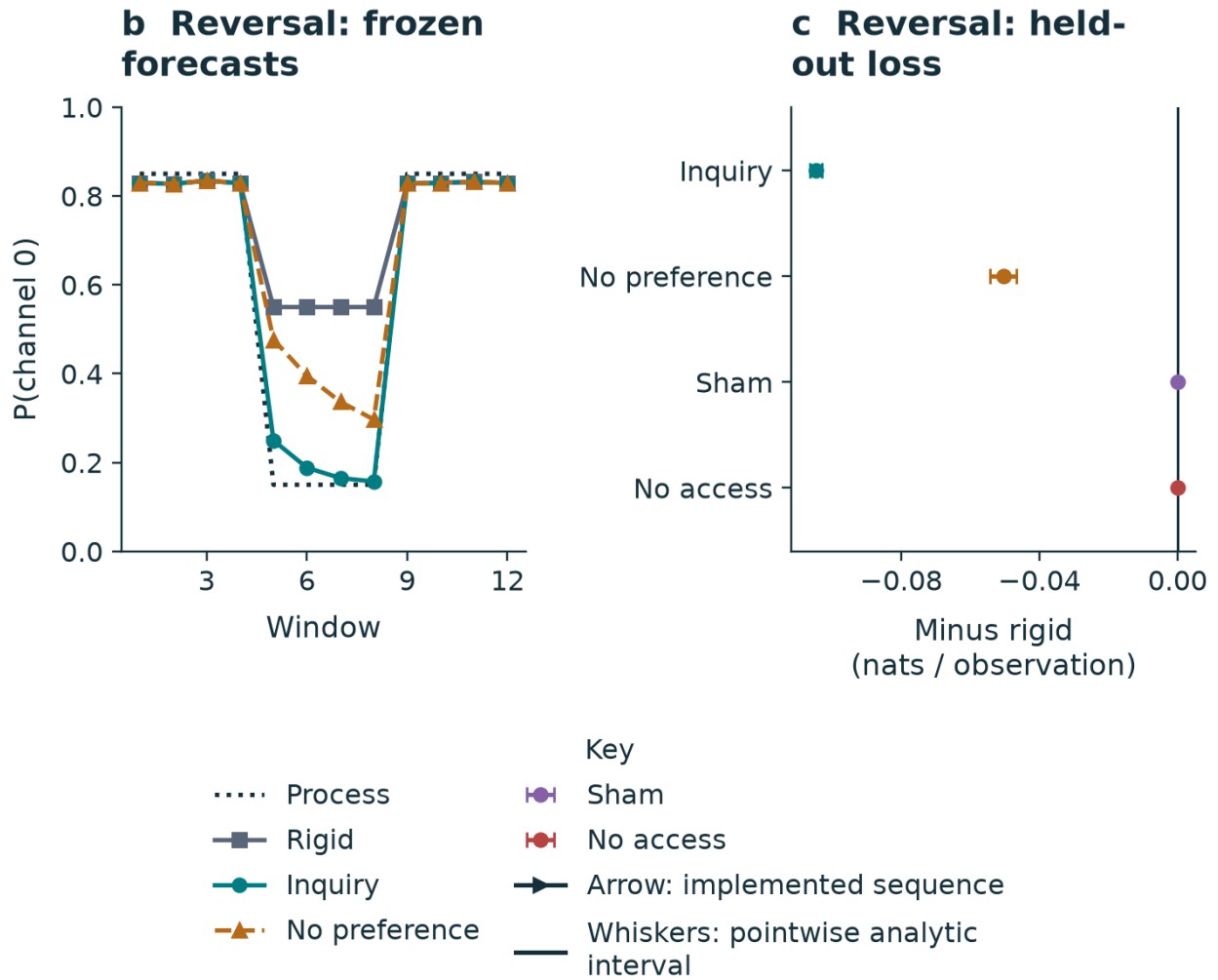
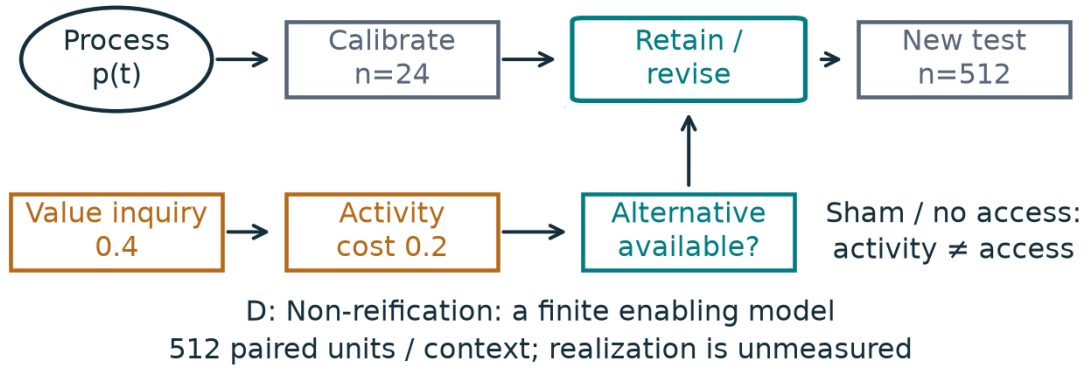


Figure 5: Access, rather than activity alone, permits evidence-responsive revision. The agent values an opportunity to examine a working family, chooses costly activity, may gain access, and selects a predictor before seeing new assessment outcomes. Panel a identifies these stages. Panel b compares reversal-process probabilities with frozen mean forecasts; panel c shows each policy's held-out log loss minus rigid prediction, with zero denoting equal loss. Sham and blocked access preserve rigid forecasts while charging activity costs. Loss whiskers are pointwise analytic 95% intervals over 512 independent paired replicate averages, each averaging 12 windows. Forecast curves are descriptive means. Negative log-loss differences favor the policy, in nats per outcome; costs are separate. The model can retain a useful restriction, and no stage measures realization.

The stationary context gives retention a substantive role. When the restricted family already supplies an adequate forecast, greater access need not produce a predictive gain, and the activity charge can remain. In the reversal context, access permits evidence-responsive use of an alternative and later return to a useful restriction. [Figure 16](#) shows those different consequences; [Figure 17](#) shows why choosing an activity cannot stand in for gaining effective access. The no-preference comparator further separates the declared value of examinability from inquiry driven by previous instrumental gain.

The exhaustive analysis qualifies any interpretation based only on the seeded trajectories. It evaluates the possible calibration counts under the declared process, including rare counts at which the open family is selected adversely. A seeded stationary null can therefore coexist with a nonzero exact probability of an unhelpful selection. **The exact-decision result** and **Figure 15** distinguish conditional selection from how often the relevant evidence occurs. The figure repeats each count distribution on linear and logarithmic axes so rare events remain visible alongside the modal evidence. Its selection probability describes how often the rule chooses the open family under a specified process, conditional on effective inquiry; it is not a posterior probability that this family is true. These results establish a property of this finite candidate menu, not that a larger menu or another process would behave similarly.

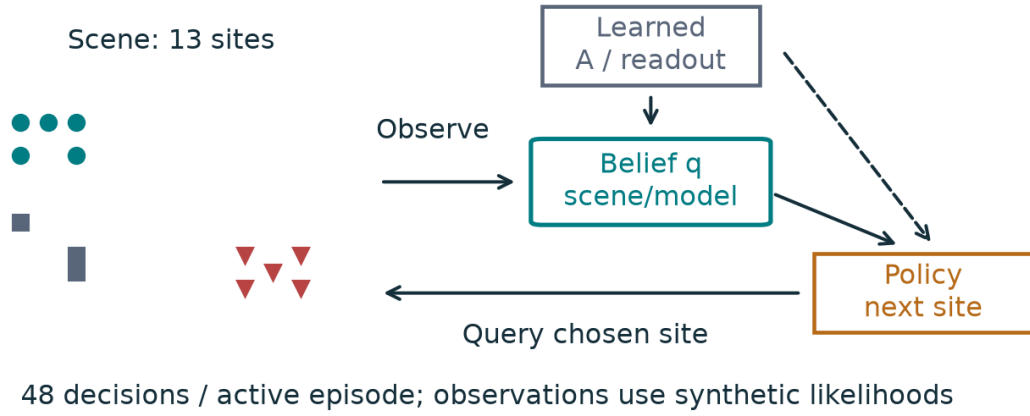
Together these findings establish a conditional enabling mechanism. They do not identify the declared opportunity value as human gratitude, nor make every relaxation beneficial. **Complete practice results** retain the costs and policy contrasts; **the complete practice interval table** retains all interval procedures, including unavailable studentized intervals. Success requires effective access and appropriate evidence use, with independent prediction and cost kept visible.

3.3 Access advantages do not establish an inquiry-specific gaze benefit

The preserved source code reproduces published before occupancy and contains no after-practice intervention. The new transfer experiments link actual learned channels to a separate scene task, where a change in positive-location sampling can coexist with worse prediction. Already-open policies improve reversal prediction relative to an excluded candidate, but the direct comparisons do not establish an additional predictive advantage specific to inquiry.

a Look, observe, update, and choose a site

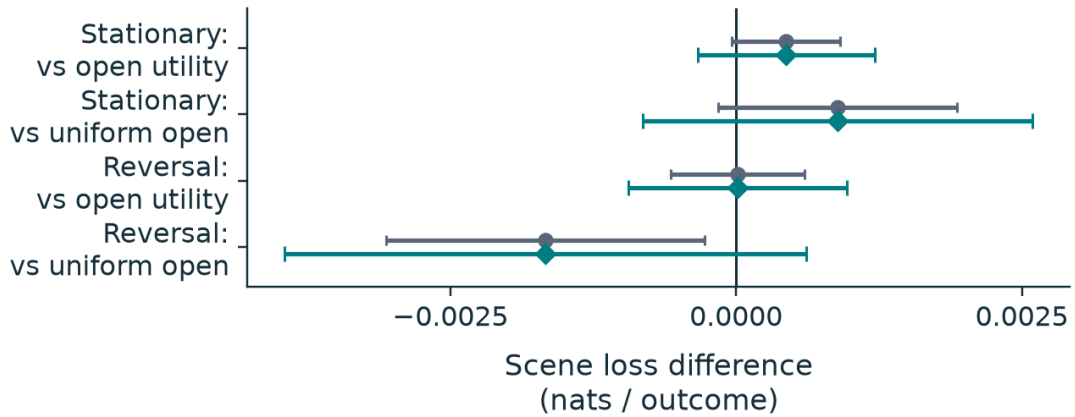
Tests: learned transfer; inquiry against already-open comparators



b Learned likelihood bridge



c Inquiry vs open sampling



Key

- ◆ b: Positive recall
- ◆ b: Matched random recall
- ◆ b: Absence valuation
- ◆ c: Pointwise analytic interval
- ◆ c: Joint bootstrap bounds
- Circle / triangle / square: positive / negative / neutral sites
- - - Dashed: alternative transfer bridge

Figure 6: Looking differently and predicting better are separate tests. An agent samples declared scene sites and updates a finite belief; learned-likelihood transfer and active inquiry are evaluated as distinct experiments. Panel a shows the sites and sampling loop. Panel b uses the likelihood bridge to compare positive-site policy changes with no recall, in percentage points. Panel c shows open-inquiry log loss minus already-open utility or uniform sampling, in nats per outcome; negative values favor inquiry. The panels have different endpoints and scales. Transfer whiskers are pointwise analytic 95% intervals over complete paired agents. Active-scene circles show analytic bounds; diamonds use the full exploratory 58-contrast joint rectangle over complete replicate averages. The dashed bridge is an alternative readout projection. Intervals spanning zero do not establish an inquiry-specific advantage, and registered model sites are not measured eye tracking.

The source and transfer findings answer different questions. [Figure 25](#) concerns the executable before condition; no reconstructed

after effect is available from that code. **Figure 27** instead reports a new experiment whose learned channels and bridge are explicit. Its positive-location change and predictive-loss change use different units and can point toward different judgments of success. The alternative bridge provides a sensitivity analysis of that connection rather than another independent replication of the source claim.

The active-scene comparison then asks whether an access advantage also identifies an information-seeking advantage. Opening a previously excluded candidate can improve forecasts in a reversal context even when the sampling rule is uniform or guided by utility alone. **Figure 29** shows the consequences separately. The direct inquiry-versus-already-open contrasts are needed before attributing an additional predictive gain to the information preference. Their bounds do not establish such a gain in the admitted design; this is a bounded lack of evidence, not a general equivalence theorem for inquiry policies.

The full gaze record retains both bridges and all paired comparisons. **Figure 30** keeps the 58-member exploratory family together, and **Figure 31** shows why nominal confidence is not a universal guarantee. Skew-related undercoverage and unavailable rare-event regions remain reported results. **The complete joint table** preserves each estimate’s status rather than dropping unfavorable or uninformative cases from the display.

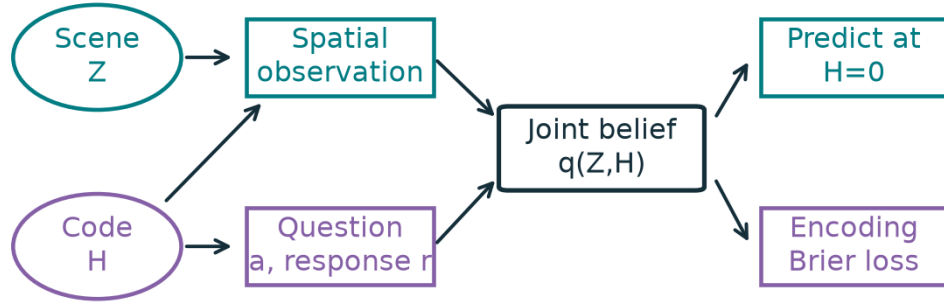
A practical-equivalence claim would require justified, prospectively specified bounds and an analysis that can exclude effects outside them (**Lakens 2017**). No such test is claimed here. **The equivalence boundary** separates these inconclusive stochastic comparisons from the model’s proved identities.

3.4 Learning an encoding differs from using it for prediction

The joint experiment separates learning an encoding from using it with scene evidence. Label-equivalent encodings remain observationally indistinguishable. Restricted zero support prevents recovery of an excluded encoding even when a distinguishing channel exists. The matched-query sham and boundary-only policy retain their exact conditional identities; adverse nonzero offsets remain visible.

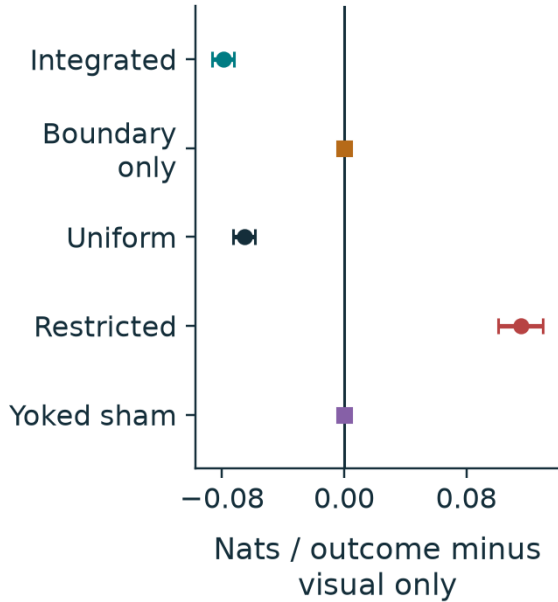
a Ask about the scene and its encoding

Test: do scene samples and encoding questions support reference transfer?

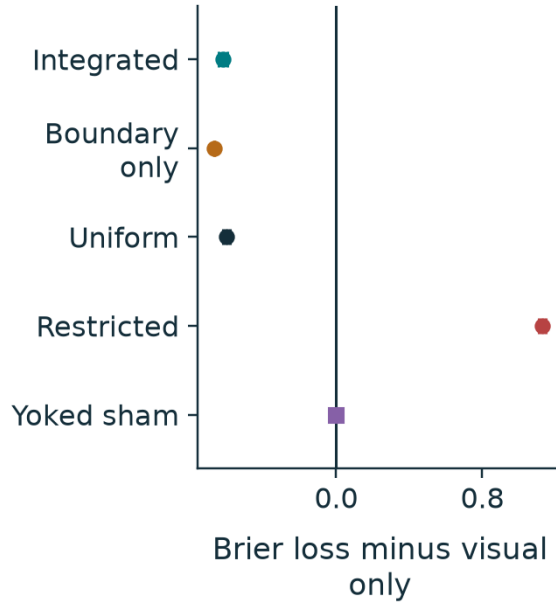


One 24-decision budget; 512 paired units / fixed scene stratum

b H2: reference prediction



c H2: encoding estimate



Key

- Circle: stochastic estimate
- Square: mechanism-exact value
- × Cross: interval unavailable
- Whiskers: joint bootstrap bounds
- ➔ Arrows: scene / encoding queries

Figure 7: Learning an encoding is different from using it for reference prediction. Spatial observations and paired randomized-action responses update a joint belief over scene and encoding under one 24-decision budget. Panel a separates evidence routes and targets. Panels b and c show policy-minus-visual-only differences for the preselected H2 case: anchored excess log loss and encoding Brier loss, respectively. Zero denotes equal loss; the endpoints have separate scales. Bounds come from the complete 84-member exploratory family, with 512 independent paired replicates per fixed scene stratum. Circles are stochastic estimates, squares proved constants and crosses unavailable intervals. Negative differences favor the named policy. Encoding improvement alone does not establish reference transfer; adverse restricted-policy offsets and exact sham values remain visible.

The diagnostic channel explains why the response record matters. An action and a response can each be marginally uninformative while their relation identifies the encoding. Erasing the action identity removes that route for the specified updater. In the label-only regime, by contrast, the encodings have the same observation law, so more sampling cannot distinguish them. **Figure 33** makes this difference visible; **Figure 34** shows how spatial observations and encoding questions share the budget. Neither a query count nor low posterior uncertainty alone establishes that the relevant distinction has been learned.

Learning an encoding is also different from using it for the anchored prediction. A boundary-only policy can improve its encoding judgment while lacking scene evidence needed for the reference target. Conversely, an excluded encoding remains unavailable

to a restricted prior even when an informative channel exists. [Figure 32](#) retains encoding and anchored losses side by side, and [Figure 35](#) keeps behavioral descriptions separate from those scores. Direct integrated-versus-uniform anchored comparisons do not establish an additional predictive advantage in this design.

The 84-member family contains 30 proved constants and retains all remaining stochastic or unavailable estimates. [Complete joint bounds](#) preserve the exact offsets and uncertainty statuses. [Figure 36](#) characterizes precision within the declared design; it cannot turn observationally equivalent encodings into identifiable ones. The separate [native boundary checks](#) establish specified QRF and Gaussian properties. [Figure 37](#) distinguishes failure of marginal-threshold blanket recovery from recovery using the known full covariance; neither is evidence of a measured human metacognitive boundary.

4 Discussion

4.1 Gratitude for the possibility of examining one’s own constructions

The stronger proposal is that appreciation can support the practice of examining what seems compulsory in one’s way of construing experience. Gratitude for an opportunity to inquire need not demand a favorable conclusion. A practice can reveal that a distinction is contingent while preserving its everyday usefulness; it can also reveal that a particular alternative does not help. This orientation differs from collecting agreeable evidence. It values the possibility of revising the terms in which evidence is understood, while continuing to answer to consequences.

The finite extension makes this proposal more demanding than a verbal endorsement. Preference must affect action; action must actually provide access; accessible comparison must consume evidence; and a selected working representation must face independent observations. The sham and no-access nulls show where the chain can break. Their role is not to devalue aspiration, but to prevent its presence from being counted as an enacted mechanism. The no-preference condition also shows why instrumental inquiry and appreciation for its possibility need separate identification. Here the latter is declared, not inferred from simulated behavior or validated as human gratitude.

Useful self/world models can persist after their assumed necessity is questioned ([Sandved-Smith et al. 2026](#)). Our model reflects that point by allowing a restricted working family to be retained or reused when its evidence is better. Returning to useful distinctions need not reinstate the rule that alternatives are inadmissible: future inquiry remains available. This is the substantive difference between a representation that is employed and a restriction that cannot be examined. The finite family switch supplies a constrained analogy; it does not exhaust the source’s physical or contemplative claims.

Practice also has a temporal structure. A moment of insight would not remove changing conditions or the need to examine future constructions. The repeated reversal design therefore asks about continued revisability rather than a one-way progression through named attainments. Its costs matter as well. More inquiry can be predictively useful while consuming resources; if an adequate model already works, the charge can remain without a predictive gain. Our separate cost column makes that possibility visible instead of declaring every relaxation beneficial.

This positive proposal requires more than being thankful that a harm did not occur. The object of appreciation can be an enabling relationship, a discrepancy, an opportunity to question an assumed boundary, or the continued possibility of revision. Appreciation following an insight is another possibility. Neither causal direction is measured by this model. Lutz and colleagues’ framework permits meta-awareness without dereification ([Lutz et al. 2015](#)); our accessible candidate menu therefore supplies a testable enabling condition rather than a sufficient realization criterion.

4.2 What gaze, confidence and first-person reports can establish

Natural gaze depends on tasks, learning and changing information requirements ([Tatler et al. 2011](#)). A fixed painting partition therefore needs an explicit task description, not an implied natural-vision interpretation. Võ and colleagues additionally distinguish recorded fixation locations from retrospective knowledge of those locations ([Võ et al. 2016](#)). Their task-specific findings motivate measuring monitoring independently rather than inferring it from a scanpath.

The joint model implements selection of evidence about a supplied encoding. It does not learn the representation language or a metacognitive architecture. Nelson and Narens’ monitoring/control distinction and Fleming and Lau’s separation of confidence bias, sensitivity and efficiency remain central to empirical translation ([Nelson and Narens 1990](#); [Fleming and Lau 2014](#)). A person could look differently without understanding the boundary differently, or report insight without improved independent monitoring.

These distinctions require different observations in a human task. Mean confidence concerns a response tendency; sensitivity concerns how well confidence distinguishes correct from incorrect judgments; efficiency concerns metacognitive performance relative to the available first-order performance ([Fleming and Lau 2014](#)). A change in one cannot be substituted for evidence of another. Likewise, looking longer at an informative region does not by itself establish awareness of why that region matters. The model’s encoding Brier score evaluates an explicitly defined first-order probability judgment, not a measurement of metacognitive efficiency or dereification.

An empirical bridge would therefore retain the actual action-response pairs, independent scene predictions and confidence reports, rather than interpreting a changed scanpath retrospectively as insight. It would also test whether any apparent advantage depends on the stipulated reference anchor and whether an observer can use information in an adaptive query schedule that the software sham ignores. [The extended gaze interpretation](#) explains the conceptual limits, and [the coupled human-task proposal](#) specifies

the distinct measurements and comparator requirements. The useful contribution is a task in which evidence routes can be separated and challenged; the task does not settle the meaning of a lived realization.

First-person descriptions would supply another evidential stream. In a visual-task study, trained participants’ reports organized trial groups for behavioral and EEG analysis (Lutz et al. 2002). Such reports could constrain our empirical interpretation alongside predictions and confidence, without making either a report or a score a sufficient criterion of realization.

4.3 From finite mechanisms to an identifiable human research program

The models assume finite authored candidate menus, known likelihoods, explicit access gates and declared costs. The scene/encoding target also requires a stipulated anchor. Useful prediction, informative questions, low entropy and flexible switching are different achievements; none is a sufficient measure of contemplative realization. Larger simulated samples characterize Monte Carlo precision within these assumptions and cannot resolve observation-equivalent models.

The gratitude literature makes comparator choice consequential beyond the simulation. The inspected preregistered meta-analysis reports small aggregate gains, but its comparison with positively valenced activities does not establish a significant advantage or equivalence (Choi et al. 2025). Its country-based cultural proxy and English-language search also constrain generalization. Our matched replay and exposure controls address a narrower computational question: what a selection rule does under a declared process. They do not isolate human gratitude from positive reflection, engagement, social attribution or expectancy. The comparator synthesis retains those study-specific qualifications alongside favorable findings and disputed optimism measurements.

A human study would separately manipulate appreciation, access to alternatives, instrumental information and delay or effort. It would record scene predictions, encoding judgments, confidence, inquiry choices and eye movements, with an independent transfer target. Practice instructions and the interpretation of process measures must be specified rather than treating all contemplative activities as one intervention (Van Dam et al. 2018). Patience, anticipated good news, expectancy and compliance remain rival mechanisms. The preserved empirical-design supplement gives the interaction estimand, identification problem, manipulation checks, alternative explanations and model-recovery requirements.

Practice specificity matters for the proposed causal chain. Dahl and colleagues distinguish attentional, constructive and deconstructive practice families and their proposed mechanisms (Dahl et al. 2015). An instruction to appreciate an enabling condition therefore cannot be assumed to recruit the same processes as attentional training or inquiry into an assumed self/world boundary. The empirical proposal would need to specify what participants actually do, how they recognize an alternative, and what counts as appropriate retention as well as revision. The inquiry-rival analysis separates valuing examinability from seeking useful information, anticipating good news or tolerating delay.

Fitting the software policy to behavior would introduce another identification problem. Different combinations of declared value, cost and action precision can produce the same initial choices; exact replay of those choices does not identify their psychological cause. Parameter recovery and model recovery address different questions, and both must be evaluated before assigning substantive meanings to fitted parameters (Wilson and Collins 2019). Equation (5) shows the confounding in this model; the prospective inquiry study specifies independent manipulations and recovery checks that would be needed to move beyond it. A menu member winning a comparison still need not explain the behavior adequately.

Relative model ranking also needs an adequacy check: simulations should reproduce the behavioral pattern they purport to explain (Palminteri et al. 2017). Our exact oracles and adverse controls create opportunities to reject specified mechanisms, but a successful finite demonstration cannot substitute for that human test. The validation account separates reconstruction, numerical verification, model criticism and construct interpretation.

4.3.1 Generalization beyond the declared experiment

Generalization would require a declared target beyond the tested sample. Yarkoni’s argument concerns variation over stimuli, tasks and settings as well as participants (Yarkoni 2022). For this proposal, instructions, social attribution and the contexts in which retention or revision is useful belong in that target. More paired simulations improve conditional precision without sampling those human differences.

The detailed interpretive argument and all source-access qualifications remain in the discussion supplement and scope limits. This is a targeted scholarly synthesis, not a systematic review. The software supports finite-model claims; human gratitude, lived realization, clinical benefit and philosophical equivalence require distinct evidence.

4.4 Conclusion: appreciate, examine and remain able to revise

“Thanks for Nothing” names more than relief that no harm occurred. It proposes appreciation for the conditions that make one’s own constructions examinable, and an orientation toward practicing realization of emptiness while retaining useful representations. The executable counterpart distinguishes the value of inquiry, its cost, chosen activity, effective access, evidence-based revision and independent consequences. Its semantic nulls prevent aspiration from being mistaken for enactment, while its retention controls prevent indiscriminate relaxation from being counted as progress.

The gratitude-recall experiments remain a complementary constraint: positive selection did not yield a general optimism or predictive-performance benefit, and statistical conclusions depended on world composition and distributional assumptions. Together, the constructions support a clear research program: appreciate without inventing favorable evidence, inquire without assuming the preferred answer, retain distinctions where useful, and remain able to revise them. The numerical results establish

finite conditional mechanisms. Whether this orientation supports human gratitude or realization requires the separate conceptual, phenomenological and empirical work the manuscript specifies.

Supplementary information

The technical record preserves the original formal methods, complete numerical results, source diagnostics and extended interpretation. Stable equation and section identifiers link these details to the IMRAD argument. Changing settings and estimates are hydrated from their producers.

S1. Conceptual distinctions and the scholarship behind them

Appreciating the conditions of inquiry

Our central proposal concerns the object and direction of appreciation. One can appreciate the conditions that permit examination: a teacher, a conversation, a stable interval, a difficult discrepancy, or the availability of an alternative description. Those conditions need not already feel pleasant or confirm a favorable self-image. Gratitude, on this proposal, can support the willingness to look more carefully at how experience is being organized. Preferring opportunities for such examination is distinct from preferring that a flattering hypothesis turn out true. Practice makes that preference consequential through repeated allocation of time and attention, with costs and the possibility of finding that a familiar distinction remains useful.

Sandved-Smith and colleagues' *There is no self-evidence* supplies the conceptual bridge (Sandved-Smith et al. 2026). Their account treats the assumed necessity of a self/world separation as a structural commitment that can become available to inquiry. Their discussion of opacification concerns bringing model-making into view; dependent-origination inquiry concerns the conditions on which a designation depends (Sandved-Smith et al. 2026, secs. 4.1–4.3, pp. 12–15). The proposed shift is not the destruction of every representation. Useful models can continue to operate while the claim that a particular division must organize every possible situation is relinquished. We take this as a reason to model access to alternatives and ongoing conditional use, rather than to reward indiscriminate uncertainty.

This gives the proposed relationship a direction: appreciation of enabling conditions can motivate practice; practice can expose a previously unexamined restriction; evidence can then support retention, revision or renewed inquiry. A gratifying thought alone does not perform these operations. Likewise, simply asserting that distinctions are empty does not show that a restriction has become examinable. Gratitude may also arise after an inquiry, as appreciation of newly visible possibilities, but that feedback is a future empirical hypothesis. The implemented feedback below concerns observed model-evidence gain only, not a simulated emotional response.

The practical target remains realization of emptiness, not appreciation of a metaphysical thing called nothing. The finite counterpart operationalizes one proposed enabling mechanism—revisability of a structural commitment—while leaving the lived realization and the philosophical adequacy of the computational interpretation open. This distinction permits a substantive positive proposal without treating a software score as its completion.

Dahl, Lutz and Davidson propose distinct attentional, constructive and deconstructive practice families (Dahl et al. 2015). Their framework separates meta-awareness, reappraisal and self-inquiry while retaining ethical, interpersonal and cultural context. In our synthesis, appreciation can prepare or motivate examination without itself constituting deconstructive insight. This distinction prevents a single practice score from standing in for every proposed link.

Social gratitude and broader appreciation

Relational evidence supplies a different target from a favorable numerical value. Algoe, Haidt and Gable's naturalistic gift-giving study associated perceived benefactor responsiveness, gratitude and later relationship ratings (Algoe et al. 2008). Only its abstract was inspected here. That social association does not identify causal mediation or a benefactor-free appreciation mechanism; the existing relational theory and intervention comparisons therefore retain separate roles.

Gratitude is not a unitary computational primitive. An influential moral account emphasizes recognizing another person's beneficial action, motivating prosocial behavior and reinforcing generosity (McCullough et al. 2001). A relational theory emphasizes finding and maintaining responsive partners (Algoe 2012). Broader accounts include habitual appreciation of positive aspects of life (Wood et al. 2010). A model of favorable and unfavorable states can represent a stipulated appreciation operation while omitting benefactor recognition, reciprocity, obligation and interpersonal interpretation. The absence of a benefactor is therefore a substantive construct boundary, not a naming detail.

This boundary also affects measurement. He and colleagues compared writing about a significant other, writing about health and no writing; highly correlated gratitude and appreciation word ratings did not establish that the constructs were interchangeable (He et al. 2022). Their selected indebtedness comparison must be read alongside a nonsignificant omnibus test and uncontrolled writing time and length. Yu and colleagues distinguish recipient benefit from benefactor cost in social gratitude (Yu et al. 2018). These findings motivate separate measures of appreciation, gratitude, relief and obligation. They do not license calling the present counterfactual value a validated gratitude score.

Bartlett and DeSteno's interpersonal experiments found greater costly helping following gratitude induction than neutral or amusement conditions; source attribution constrained its incidental extension to strangers (Bartlett and DeSteno 2006). These results motivate separating action and social interpretation from pleasant affect. They do not test contemplation or establish that benefactor-free appreciation motivates inquiry.

Gratitude may also affect the costs a person will tolerate. In DeSteno and colleagues' randomized autobiographical-recall study, a planned gratitude-versus-pooled-comparators contrast favored greater patience in incentivized intertemporal choice (DeSteno et al. 2014). The inspected accepted manuscript concerns 75 young adults, an immediate task and that planned

contrast; it does not establish enduring adherence or every separate pairwise effect. For the present proposal, patience is both a plausible bridge to sustained practice and a rival explanation: greater willingness to wait does not identify a greater value of examinability.

What intervention comparisons identify

Comparison activity determines the explanation a study can address. Cunha and colleagues' randomized writing trial had substantial attrition, and its neutral-event activity often included positive material (Cunha et al. 2019). A post-test positive-affect advantage therefore does not imply superiority on every outcome or at follow-up. A guided internet intervention against a wait list addresses a different comparison, which does not isolate gratitude from guidance and expectancy (Heckendorf et al. 2019). Favorable diary associations add another design rather than settling the mechanism (Hartanto et al. 2023).

Meta-analyses and reviews emphasize comparator-dependent conclusions, generally small effects and the importance of positive or psychologically active controls (Davis et al. 2016; Cregg and Cheavens 2021; Kerry et al. 2023). The inspected Kerry review is itself bounded by restrictive inclusion and a single-researcher search. A newer preregistered synthesis reports small aggregate gains and between-country variability, providing a counterweight to a uniformly skeptical account without validating a specific computational mechanism (Choi et al. 2025). Its full text reports no significant advantage over positively valenced activities; that imprecise comparison does not establish equivalence. Country is a proxy for culture, and the English-language search limits generalization. Matching replay dose and exogenous exposure in our model isolates a selection rule under a declared process. It does not isolate human gratitude from expectancy, task engagement, social meaning or demand characteristics.

Optimism is a further, distinct target. Daily-life findings identify overlapping but distinguishable associations of gratitude and optimism (Newman et al. 2021). Belief-updating tasks motivate the source readout, but their specificity and measurement remain contested (Garrett et al. 2014; Hobbs et al. 2022). Shah and colleagues argue that apparent updating asymmetries can arise from statistical task properties; Garrett and colleagues dispute aspects of that analysis (Shah et al. 2016; Garrett and Sharot 2017). This disagreement motivates estimator controls and independent predictive scores rather than treating a favorable task score as construct validation.

Absence, mental subtraction and relief

An observed absence can be perceptually structured: the inspected silence experiments concern contrastive auditory silences within structured sound (Goh et al. 2023). Their discussion leaves absolute silence as a further question and qualifies analogies to other modalities. These results do not test a failed recorder. An imagined alternative is different again. Koo and colleagues asked participants to imagine that an existing positive event had not occurred (Koo et al. 2008). Their affect and relationship outcomes should not be summarized as uniformly gratitude-specific: the first two studies did not show significant differences on the direct event-gratitude item. Mental subtraction of an actual benefit is not the same manipulation as detecting an uneventful interval, establishing that a harm was prevented, or assigning an externally chosen hazard reference.

Epstude and Roese's functional account further separates content-specific pathways, which can supply information for action, from content-neutral pathways through affect, motivation and processing style (Epstude and Roese 2008). The emotional consequence depends on how the alternative is used: comparison with a worse possibility can produce favorable contrast, whereas vividly imagining it can produce assimilation. This is a reason to measure the process and its outcome separately. Our counterfactual value condition implements a reference calculation with a designed belief null; the review does not establish that human counterfactual thinking preserves beliefs or that its affective consequence is necessarily gratitude.

Counterfactual gambling feedback (Coricelli et al. 2005) and conditioned pain relief (Seymour et al. 2005) concern still other contrasts. The common reference dependence makes them relevant comparisons, but does not identify relief, gratitude and appreciation as one latent mechanism. Our valuation-only condition is deliberately narrower: it asks whether the software can assign a counterfactual value while preserving its evidence update. Human reflection could change recall, attention or hazard beliefs as well; the designed null is an experimental contrast to test, not a prediction imported from those studies.

Emptiness and the continued use of representations

The main-text introduction identifies the Buddhist collections and named discourses and supplies short, edition-specific quotations. Those textual readings constrain interpretations that equate emptiness with total non-existence or a blank numerical state. The distinction between what is absent and what remains, and the discourse's treatment of refined concentration as conditioned, matter to the interpretation (Thanissaro Bhikkhu 1997b, 1997a). This is philosophical source interpretation, not a claim that the text describes Bayesian computation.

Garfield's translation and commentary and his work on cross-cultural interpretation supply further philosophical context (Garfield 1995, 2001). Our access to those books is limited to verified publisher identity and synopsis, not verse-by-verse analysis. The inspected emptiness preprint develops a specific physics and active-inference account; it does not settle every philosophical interpretation (Sandved-Smith et al. 2026). Its emphasis on continuing revision after a structural change (Sandved-Smith et al. 2026, secs. 5.1–5.3) motivates repeated opportunities to examine a working model. Its broader policy-scope discussion (Sandved-Smith et al. 2026, secs. 6.1–6.2) motivates a normative question about whose consequences matter. We keep that normative question distinct from prediction: a wider action-influence scope is not, by itself, a measure of compassion.

We use non-reification as a modeling discipline: a chosen representation can function without being treated as the only admissible description. Consistent label permutation checks a prerequisite representational invariant, while the new practice experiment

tests the consequential availability of a genuinely different predictive candidate. Neither renaming states nor increasing entropy is counted as realization. In particular, a distinction that predicts well should be retainable. Valuing examinability is therefore compatible with precise, useful everyday inference.

Garfield and Thakch e’s chapter gives the object of negation a precise role (Garfield and Thakch e 2011, 73–87). Negating intrinsic nature and negating conventional phenomena have different implications; emptiness itself must not simply replace intrinsic nature as another intrinsic essence. Their comparison of Tsongkhapa and Gorampa also shows why preserving conventional efficacy does not settle every Madhyamaka interpretation. The chapter presents that disagreement without adjudicating it. We use the limited analogy of withholding intrinsic status from a useful construction, rather than claiming that every tradition equates emptiness with representational flexibility. The source catalog distinguishes the cited print edition from its earlier online publication and preserves both publication dates.

This sharpens the positive practice proposal. Gratitude can concern the contingent conditions that enable investigation and the continuing possibility of using distinctions without making them compulsory. Practicing inquiry includes attending to how a distinction is established, acting to examine its consequences, and remaining able to retain or revise its application. Our software executes a finite portion of this proposal. Its assessment concerns prediction, access and cost; the philosophical and experiential claims require arguments and observations beyond that execution.

From conceptual distinctions to model tests

Distinction	Implemented comparison	Evidence still needed
Replay content versus activity dose	Positive versus matched random recall	Human controls matching expectancy and engagement
Detected absence versus missing record	Absence likelihood versus missingness sentinel	Detection reliability and recording-failure manipulations
Counterfactual value versus new evidence	Separate valuation with a designed belief null	Joint valuation, belief, relief and gratitude measures
Model policy versus gaze	Declared learned-channel transfer and paired scene-risk controls	Empirical transfer and independent gaze prediction
Scene content versus its encoding	Coupled scene/contingency queries, erasure and label-equivalence controls	Measured gaze, action-pair access and experimentally instantiated reference transfer
Confidence versus metacognitive sensitivity	Posterior uncertainty and proper prediction scores kept separate	First-order judgments, corresponding confidence and sensitivity/efficiency estimands
Motivation versus access	Preference and costly activity versus an effective access gate	Separate social attribution, engagement and metacognitive process measures
Access versus enacted revision	Available repertoire versus evidence-selected working family	Noticing and using alternatives, with curiosity and compliance controls
Useful representation versus compulsory restriction	Exact prior comparison and held-out prediction after reversals	Behavioral model comparison and alternative explanations of flexibility
Representation versus realization	Label invariance as a prerequisite, not the endpoint	Philosophical argument and phenomenological and empirical work

This mapping is our synthesis of the inspected literature, not a combined claim established by any cited study. The source catalog records whether each item was inspected in targeted full-text sections, at abstract level or through bibliographic metadata. The review is targeted, not a systematic search.

Validation, model criticism and generalization

Different checks warrant different conclusions. Reconstructing a source tests fidelity to an artifact; independent mathematical controls test a specified implementation; behavioral validation tests a model’s explanatory claims. Cronbach and Meehl’s construct-validity argument adds a further obligation: an interpretation must participate in testable relations, and adverse evidence may require reconsidering the measure, theory or design (Cronbach and Meehl 1955). Our terms appreciation and examinability identify proposed operations, not validated human constructs.

Relative fit cannot settle adequacy. Palminteri and colleagues recommend simulation-based tests of the behavioral effect a model is meant to explain, alongside discrimination and recovery (Palminteri et al. 2017). Gelman and Shalizi distinguish posterior calculation from checking model implications and using discrepancies to revise a family (Gelman and Shalizi 2013). Our candidate-menu controls implement narrower checks. They neither generate a new representation language nor establish that the selected family describes a person. Likewise, mathematical revisability supplies no independent philosophical proof of non-reification.

The target of a generalization must also match what the design addresses. Yarkoni includes tasks, stimuli and settings among relevant sources of variation (Yarkoni 2022). Our fixed worlds and authored scene partition therefore define a conditional computational target. Their weighted average is not an estimate over human cultural or contemplative settings. The textual

comparison serves a different purpose again: its named editions discipline interpretation, without turning selected passages into sampled evidence about populations. **The empirical program** must establish its own target and measurement plan.

Textual witnesses, useful absence and gratitude

The ancient comparisons sharpen the object of appreciation and the purpose of inquiry. They do not provide interchangeable definitions of nothing. In the *Kataññu* discourses, gratitude concerns received care and its ethical repayment (Thanissaro Bhikkhu 2002). Extending appreciation to a detector, an opportunity to examine assumptions or an impersonal enabling condition changes that interpersonal structure. A future human study must therefore establish which construct is elicited, rather than validating it by a scriptural label.

The *Daodejing*, in James Legge's inspected digital edition titled *The Tao Teh King*, describes the use of a wheel, vessel and room through the relation between their constructed form and unfilled space (Laozi 2008, ch. 11). This offers a comparison for appreciating enabling conditions. The useful hollow of a vessel is neither an unavailable record nor, by itself, evidence of freedom from reification. Chan distinguishes divergent interpretations of *wu* and relates *wuwei* to ethical and political questions; he explicitly resists reducing it to total inactivity (Chan 2025, secs. 6–7). We therefore retain the transliterated terms rather than treating nonbeing, nonaction and Buddhist emptiness as synonyms.

Epictetus makes the capacities for perception and understanding objects of gratitude within an argument for divine providence (Epictetus 1887, I.6, pp. 19–21). That comparison gives appreciation for cognitive capacities a concrete historical counterpart, while its providential commitment differs from our account of contingent inquiry conditions. The declared opportunity-value parameter neither measures Stoic gratitude nor imports that metaphysics.

Textual comparison	Question sharpened here	Computational use	Inference not licensed
<i>Kataññu</i> discourses (Thanissaro Bhikkhu 2002)	Who or what is appreciated, and why?	Separate benefaction from declared opportunity value	Impersonal appreciation is canonically identical to interpersonal gratitude A family switch is liberation
Raft simile (Thanissaro Bhikkhu 2004)	When is reliance useful, and when is relinquishment appropriate?	Retain useful working representations while keeping alternatives examinable	
<i>Suñña</i> discourse (Thanissaro Bhikkhu 1997c)	Empty of what?	Specify the object of a non-reification analogy	No event, no data and non-self have the same meaning
<i>Daodejing</i> (Laozi 2008; Chan 2025)	What can an unfilled or unasserted condition enable?	Keep functional absence and its value distinct	<i>Wu</i> , <i>wuwei</i> and Buddhist emptiness are equivalent
Epictetus (Epictetus 1887)	Can the capacities for inquiry themselves be appreciated?	Make the proposed object of appreciation explicit	A numerical inquiry preference establishes gratitude or providence
<i>Heart Sutra</i> (Thich Nhat Hanh 2014)	How does a specified rendering relate emptiness and embodied phenomena?	Question a reified interpretation of an empty-state label	All Buddhist textual traditions supply the same realization criterion
<i>Bhagavadgītā</i> (Swami Mukundananda, n.d.-a)	What differs between acting and claiming entitlement to its fruits?	Separate inquiry value from a demand for favorable evidence	Nonattachment means inaction or an uninformative posterior
Thanksgiving and blessing (New Revised Standard Version Updated Edition 2021; Kulp, n.d.-a; Berakhot, n.d.; Khattab, n.d.-a)	To whom is appreciation directed, and what response is required?	Distinguish addressee, circumstance and evaluative response	Harm becomes favorable evidence, or a religious promise becomes an efficacy estimate
<i>Mystical Theology</i> (Pseudo-Dionysius the Areopagite, n.d.)	How can praise coexist with limits on predication?	Distinguish appreciation from exhaustive conceptual description	Apophatic theology and Madhyamaka negate the same object
<i>MMK</i> (Nagarjuna 2020)	How do dependent designation and conventional discourse bear on emptiness?	Keep a useful distinction's status open to examination	Conventional truth is held-out accuracy
<i>Gītā</i> offerings (Swami Mukundananda, n.d.-b)	What response is owed for received gifts?	Distinguish valuing a benefit from undertaking a practice	The model implements yajña or divine reciprocity
<i>Avot</i> (Kulp, n.d.-b)	Must appreciation await an additional possession?	Separate existing-lot contentment from a new reward	Contentment automatically measures gratitude or justifies deprivation
<i>Japji</i> (Guru Nanak, n.d.)	What changes when one understands an order one does not author?	Ask about responsiveness without assuming mastery	Hukam and haumai are a likelihood law and a self-model
<i>Analects</i> (Confucius 2002, 2015)	Can a response consist in withholding an unwanted imposition?	Retain a relevant ethical comparator without inventing gratitude	Shu is repayment, emptiness or gratitude practice

Textual comparison	Question sharpened here	Computational use	Inference not licensed
Qur'an on gratitude's beneficiary (Khattab, n.d.-b)	Who benefits, and does the addressee need repayment?	Distinguish beneficiary, addressee and obligation	Divine self-sufficiency is absence or an empirical utility function

These are deliberately asymmetric comparisons. Early Buddhist translated discourses, later philosophical accounts of emptiness, Daoist interpretive traditions and Stoic providential reasoning have different questions and aims. The proposed model borrows no authority for an empirical effect from their age. Its contribution is to expose a limited sequence of choices and consequences that a subsequent empirical design can challenge.

S2. Declared models, executable interfaces and independent controls

Parsed graphical structure and execution roles

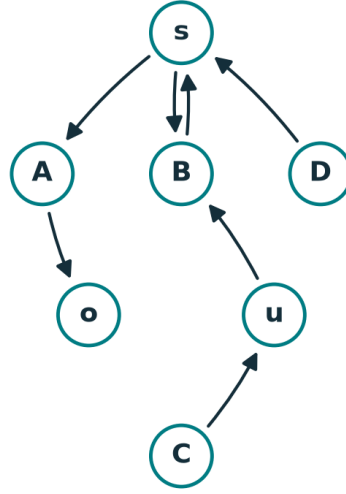
[Figure 8](#) is generated from the actual pinned GNN parser’s variables and connections, rather than a separately authored edge list. Every cataloged model is also rendered by GNN’s native network renderer. Admission checks native graph statistics against parsed topology and compares retained pixels with a fresh isolated native render, in addition to binding image digests and the numerical record’s source identity. The inventory in [Figure 9](#) focuses on complete runtime families, with dimensions and instance counts from the admitted arrays and execution catalog. The opening model card separately lists authoring inputs and distinguishes native attention execution, finite numerical consumers and the neutral analysis-protocol carrier. A declared backend identifies a supported representation; it does not by itself prove that a particular numerical backend executed an experiment.

These graphs express declaration dependencies. The state–transition cycle is not an instantaneous causal DAG, and the shared syntax does not identify a psychological mechanism. Passive preferences remain inactive where the finite consumer does not select actions. The analysis carrier’s neutral tensors carry settings rather than a new model of a person. The software checks reject stale model identities, altered topology and missing native renders before accepting the publication products.

The opening Bayesian model card unrolls the categorical tensor definition across time, conditional on selected actions; it does not reinterpret the declaration cycle as an instantaneous causal network. Its inventory distinguishes the same numerical consumers shown here. The accompanying complete model cards preserve all admitted arrays and scientific settings, original declaration annotations, native graph assets, dependency revisions and exact consumer digests. Finite task-specific likelihoods and policies remain defined by their own declared settings and methods rather than inferred from the common tensor syntax.

The executable model set also includes complete standalone GNN instances for exogenous worlds and observation schedules, news readouts, the fixed attention process, transfer embedding and alternative prior supports. Scene and joint-inquiry declarations expand all query-conditioned channels into a controlled categorical model whose state combines the known query with the latent candidate basis. Its transition sets the next query while retaining the candidate; numerical consumers recover conditional likelihoods from the admitted GNN array. Diagnostic and label-only laws are separate complete models. Each instance is preserved and read back through real GNN admission before use, with originating declaration versions, consumer identities and evidence digests retained. Fitted agent channels remain learned data under the declared family. Historical notebook behavior and independent mathematical oracles retain their separate evidential roles.

Parsed topology / 20 declarations



Parsed topology / 6 declarations

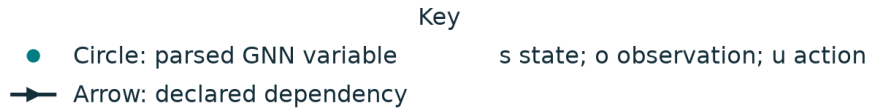
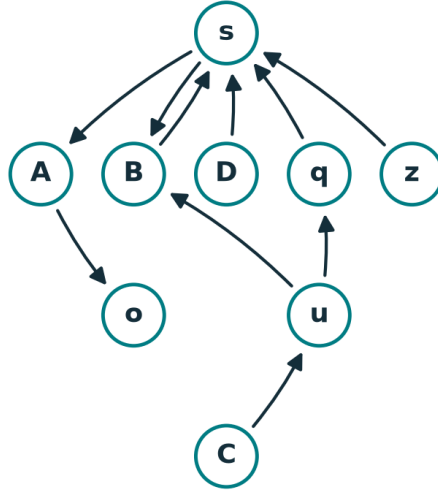


Figure 8: Graph structure comes from the executed GNN declarations. The graph exposes the admitted StateSpaceBlock variables and Connections edges. Only declarations with identical parsed nodes and edges share a panel. Complete executable instances expose known-query and latent-basis variables as well as their joint state. Arrows retain the declared direction; native per-model GNN renders and topology statistics accompany the publication view. This is a declaration dependency graph, not an estimated causal DAG or an uncertainty interval. The state–transition cycle is not an instantaneous causal claim. A common graph syntax does not imply common psychological meaning, action selection or numerical execution. The tensor and consumer inventory distinguishes those roles.

Runtime family / instances	A	B	C	D
attention process 1 instance(s)	3 x 13	13 x 13 x 13	3	13
boundary diagnostic 1 instance(s)	9 x 144	144 x 144 x 16	9	144
boundary diagnostic (R) 1 instance(s)	9 x 144	144 x 144 x 16	9	144
boundary label only 1 instance(s)	9 x 144	144 x 144 x 16	9	144
boundary label only (R) 1 instance(s)	9 x 144	144 x 144 x 16	9	144
optimism poor 1 instance(s)	2 x 2	2 x 2 x 1	2	2
optimism well 1 instance(s)	2 x 2	2 x 2 x 1	2	2
practice process 2 instance(s)	2 x 12	12 x 12 x 1	2	12
practice (R) 1 instance(s)	2 x 3	3 x 3 x 1	2	3
recall world 5 instance(s)	3 x 2	2 x 2 x 1	3	2
scene open 1 instance(s)	3 x 9	9 x 9 x 3	3	9
scene (R) 1 instance(s)	3 x 9	9 x 9 x 3	3	9
transfer embedding 1 instance(s)	2 x 13	13 x 13 x 13	2	13

Key

Dimensions: admitted tensor shapes

R: restricted prior support

Rows: complete runtime families

Figure 9: Complete runtime GNN families expose the full declared state space. Each row binds a complete executable family to admitted dimensions and numerical evidence. Instances share a row only when family and dimensions agree; their parameter values can differ. The front-matter model card separately lists the authoring inputs, including native attention models and the analysis carrier. A is observation by state; B is next state by previous state by action; C contains observation preferences; D is the state prior. R denotes restricted prior support. Dimensions and instance counts are injected from admitted tensors and the execution catalog. Shapes describe finite model structure, not sample size or empirical uncertainty. These local finite consumers use GNN-admitted arrays; native GNN/pymdp execution is separately recorded for attention. The generic GNN backend label does not establish that a native policy executor ran every experiment. Passive preferences and neutral protocol tensors do not acquire a scientific role from appearing in the graph.

Input		Model operation		Readout / assessment
A: Observed absence				
Recorded no-event	→	P(low hazard) 0.600 → 0.871	→	Information update 0.258 bits
B: Missing observation				
Record unavailable	→	P(low hazard) 0.600 → 0.600	→	Declared rule 0.000 bits
C: Counterfactual value				
Specified alternative	→	Fixture value 0.3 cost units	→	Paired belief null max $ \Delta \text{belief} = 0$
D: Non-reification				
Examinable distinction	→	Costly inquiry cost = 0.2	→	Retain or revise; assess new predictions

A-C: finite audit readouts. D: an enabling mechanism; realization is unmeasured.

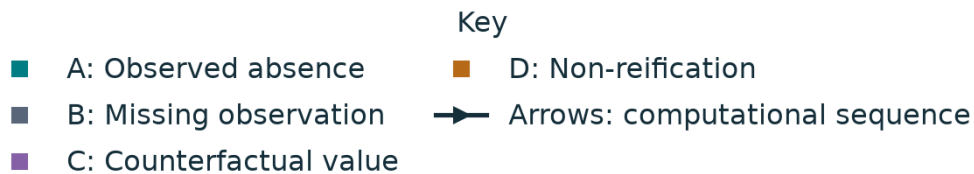


Figure 10: From each meaning of nothing to its implemented operation. The rows connect the shared vocabulary to distinct inputs, model operations and assessment targets. A: Observed absence and B: Missing observation show posterior probabilities and KL information in bits. C: Counterfactual value places the fixture's reference value in cost units beside the corrected study's maximum paired belief difference. D: Non-reification links costly access to independent prediction. These are exact fixture readouts or declared settings, with no sampling intervals; values in different units are not a common effect size. The first rows test numerical semantics. The non-reification row is a finite enabling counterpart to Realizing Emptiness, rather than a realization score.

Experimental questions and independent assessment

The crosswalk below is generated from the same typed task registry as the main overview. It separates what changes, what is assessed and which control constrains the interpretation.

Experimental question	Manipulation / matched input	Assessment	Control / limit
Recall: does selected replay improve prediction?	Positive vs random replay; same exposure and matched replay dose.	1024 paired units/world; frozen Brier and optimism tests.	Value-only control: unchanged beliefs and predictive losses.
Absence: no event, no record, or a different value?	Functioning detector vs missing record; actual vs reference cost.	Posterior change in bits; separate counterfactual cost value.	No record adds no evidence under the declared rule.
Practice: does chosen activity provide usable access?	Value and cost affect activity; 24 calibration observations.	512 new outcomes/window; 512 paired units/context.	Sham or blocked access: activity can cost without revising.
Transfer: what follows from a learned likelihood?	Freeze learned channels; compare likelihood and readout bridges.	5120 complete paired agents; attention and scene loss separately.	A favorable attention shift need not improve prediction.
Scene inquiry: access or the sampling policy?	Restricted vs open; utility vs inquiry vs uniform; 48 decisions.	512 paired units/context; scene loss and movement cost.	Already-open comparators test an inquiry-specific increment.
Boundary inquiry: can the inferred frame be used?	Spatial vs action-response queries; one 24-decision budget.	512 paired units/cell; encoding and reference losses.	Label-only, excluded support and matched-query erasure controls.

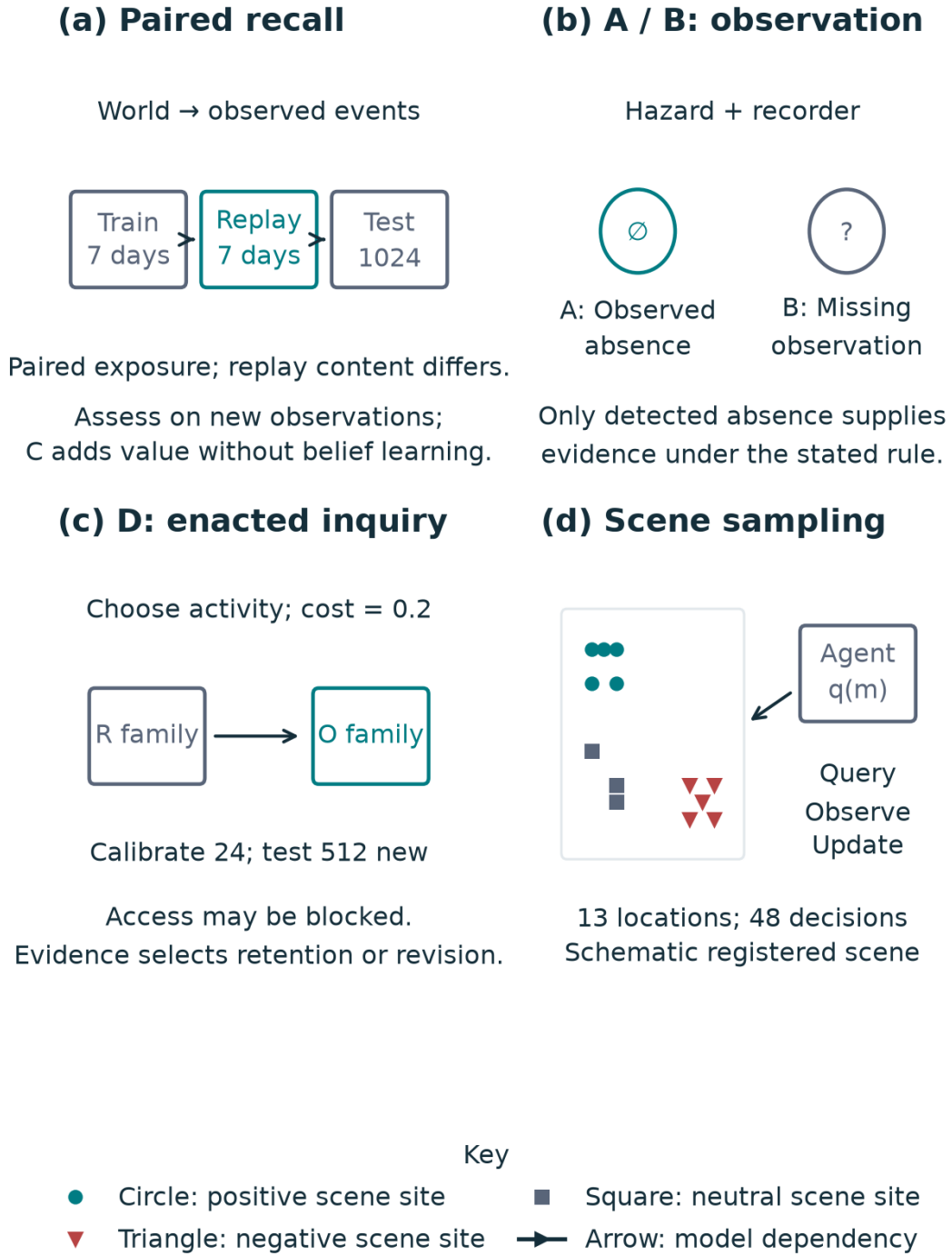
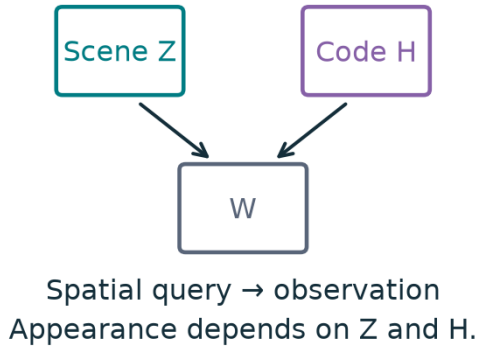
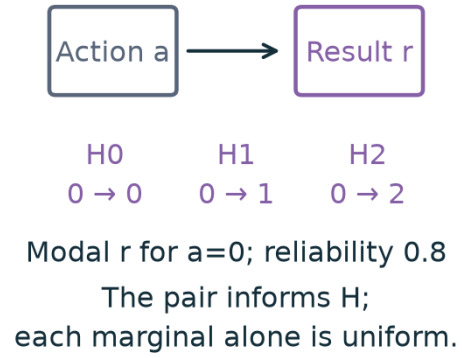


Figure 11: Concrete settings behind the experimental comparisons. The panels introduce paired recall, a detector with an absent or unavailable record, costly access to predictive alternatives, and discrete scene sampling. Training and replay durations, assessment sizes, costs, locations and decision budgets are generated from the executed records. In the scene, circles, triangles and squares identify positive, negative and neutral registered sites; R and O denote restricted and open candidate families. The timelines and icons describe the protocol and carry no inferential intervals. The scene is a schematic registration; its pixels do not generate observations. Counterfactual value does not update beliefs, and activity does not guarantee access.

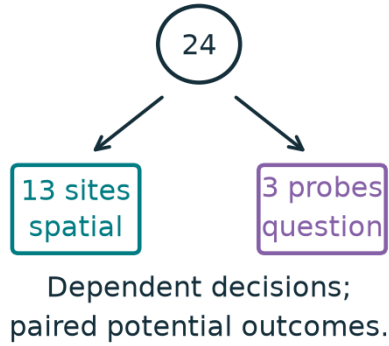
A Scene and encoding



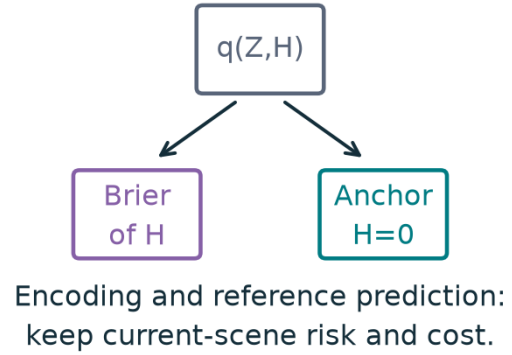
B Question the encoding



C One decision budget



D Separate assessments



Key

- Scene and spatial observations ➔ Arrows: declared dependencies
- Encoding and structural questions

Figure 12: Two evidence routes answer different questions about the same scene. Scene content Z and encoding H jointly determine appearance W ; an encoding question returns a randomized action and its conditional response. The middle probe uses reliability 0.8. Spatial samples and 3 question types share 24 decisions. Separate scores assess the encoding and prediction under the declared $H=0$ reference intervention. All boxes describe an admitted protocol, not an estimated effect; arrows denote data dependencies. Neither evidence route alone identifies the full reference target in the diagnostic setting. The pinned Realizing Emptiness mathematical fixtures are separate checks, not extra observations or participants.

GNN declarations and executable upstream interfaces

Packaged Generalized Notation Notation declarations supply the categorical tensors and experimental parameters. The project imports the pinned GNN implementation and the Realizing Emptiness model-admission interface rather than maintaining a parallel parser. Admission checks finite entries, nonnegative probabilities, normalized likelihood and transition columns, and the expected dimensions. Each result retains the hash of its model declaration. The admitted design serialized with the results records the seed, phase lengths, observations per day, replay limit and precision, learning and forgetting rates, concentration, assessment sizes, and resampling budgets; these values are the numerical specification for replay. Realizing Emptiness also supplies finite KL and information-ordering comparators and a binomial interval routine. These imports are software dependencies with bounded mathematical roles; their use does not transfer empirical or philosophical validation to this project.

The learning model has 2 hidden states, labeled going well or going poorly, and a 3-category observation alphabet, labeled positive, neutral, and negative. Labels identify an operational coding scheme. They do not independently establish an event’s value to a person. Environmental states are sampled with a fixed stratum-specific probability of going well, followed by observations from the declared environmental likelihood. This produces a stationary artificial world within a stratum. The agent’s transition model is separately declared and need not equal the environment’s state frequency. Environmental luckiness is therefore a design parameter, not a diagnosis or a personality measure. The practice extension uses a different, valence-neutral binary channel; its predictions are never labeled happiness or safety.

The imported Realizing Emptiness 3.0.0 source is commit pinned (Friedman 2026). We use its `BoundaryScreen`, `QRFDeployment` and `SeparationPrior.admissible` interfaces to validate a common screen and admit a restricted repertoire; `FreeEnergyTerms` computes actual KL complexity and expected log likelihood; `BMRComparison` supplies the difference algebra. The package also executes `practice_protocol_map()` and `build_ongoing_revision_audit()` and retains their receipt and executing-module hashes. The upstream practice mapping supplies a vocabulary, not an applied intervention engine: its GNN deltas remain metadata. The upstream revision diagnostic selects and scores the same windows. Our enacted, costly policy and independent assessment are new mechanisms in this project, not improvements silently attributed upstream.

Notation and distinct modeling roles

The recall, absence and practice models share categorical machinery but have different scientific roles. The following notation keeps state inference, family comparison and action choice distinct.

Symbol	Meaning and units
o, s	Observed category and hidden state in the recall/absence models
a_{os}, A_{os}	Likelihood pseudocount and normalized observation probability
p_t, q_t, β	Prior, posterior and generalized-Bayes observation precision
ρ_f, η	Per-update forgetting multiplier and learning increment
m and $j \in \{R, O\}$	Predictive candidate and restricted/open working family
n_{to}, Z_{tj}, F_{tj}	Calibration count, sequence evidence and free energy in nats
$\rho_e, c, \lambda_g, \beta_a$	Examinability value, activity cost, gain-to-utility weight and practice choice precision
$g_{t-1}, \hat{p}, L$	Previously observed gain, frozen forecast and assessment log loss
h, r, H, R	Fixed world stratum, paired replicate and their counts
$E, a, x_a, r(a)$	Learned-state embedding, scene location, unit-square coordinate and its region
$\eta_H, \lambda_I, \kappa$	Scene ambiguity weight, information weight and movement charge
γ_T, γ_A	Transfer-policy and active-scene choice precisions, respectively

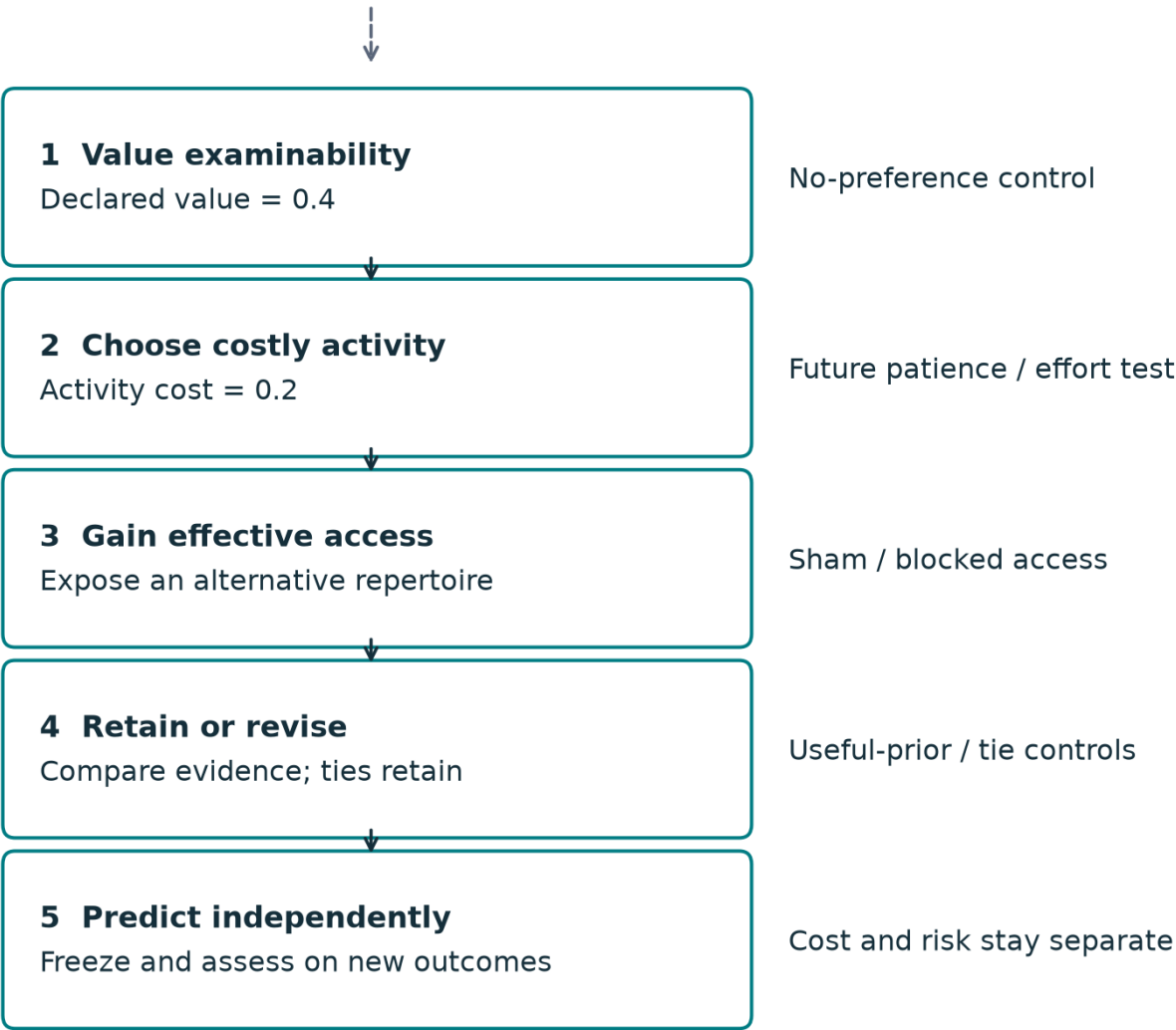
Observation precision β , choice precision β_a and upstream structural-prior precision have different roles. No common scalar called precision is treated as a measure of realization. The forgetting multiplier ρ_f is unrelated to examinability value ρ_e .

An enacted policy for examinable representations

The central practice experiment asks whether valuing the availability of inquiry can sustain a practice that makes a structural restriction revisable. It implements a finite counterpart of an enabling mechanism for realizing emptiness; it does not encode realization as a hidden state or an optimization target. The 3 declared deployments share a 3-channel screen. Sector assignments distinguish deployments using only `self` and `env` from the deployment using `sensation`, `action` and `context`. The imported separation-prior interface admits 2 deployments when enforced and the complete menu when relaxed. These labels are operational assignments, not measurements of a person’s self/world experience. Distinct predictive likelihoods make the alternative consequential; mere relabeling remains a negative control.

Concept D: Non-reification | Realizing Emptiness

Appreciation-to-preference is unmeasured; the model declares it.



Continuing access permits future inquiry; no stage measures realization.

Key
 ➡ Solid: implemented dependency ➡ Dashed: unmeasured proposal

Figure 13: A proposed human motivation and an executed software mechanism. The diagram asks which links are needed for appreciating examinability to support useful inquiry. Solid arrows connect declared value, costly activity, effective access, evidence-based retention or revision, and prediction on new outcomes. Named software controls and prospective behavioral rivals identify ways these links can fail. The dashed appreciation-to-preference link is an unmeasured proposal; the diagram contains no effect estimates or confidence intervals. Future access remains possible after a working-family choice. Activity, an open menu and a favorable score are each insufficient as a measure of realization.

Declaration	Value / role
Candidate channel-zero probabilities	0.85, 0.55, 0.15
Restricted / open prior	(0.5, 0.5, 0) / (0.333333, 0.333333, 0.333333)
Transition / observation preference	Identity / neutral
Stationary context	Channel-zero probability 0.85 for 12 windows
Reversal context	0.85 for 4 windows; 0.15 for 4 windows; 0.85 for 4 windows
Replicates / observations	512 per context; 24 calibration and 512 held-out per window
Examinability value / cost	0.4 / 0.2 declared utility units
Action precision / previous-gain weight	6 / 1 utility unit per nat
Access probability when available	1; no-access and sham controls block effective access
Resampling	4095 bootstrap draws over replicate averages

These values are declared in `emptiness_practice.gnn.md`; the generated audit retains the exact admitted design and source hash. The upstream structural-prior precision of 2 is passed as validation metadata, not converted into a hand-added free-energy penalty. This extension does not claim a continuous precision-ablation result.

Let m index the candidates, A_{om} their declared channel likelihoods, and n_{to} the calibration counts in window t . For each candidate prior $p_j(m)$, with $j \in \{R, O\}$ denoting restricted and open, define

$$\ell_{tm} = \sum_o n_{to} \ln A_{om}, \quad q_{tj}(m) = \frac{p_j(m) \exp \ell_{tm}}{\sum_{m'} p_j(m') \exp \ell_{tm'}}. \quad (1)$$

Both posteriors are recomputed on the same calibration observations. The imported distribution-derived free-energy decomposition, evaluated consistently in nats with `base=e`, is

$$F_{tj} = D_{\text{KL}}[q_{tj} \| p_j] - \sum_m q_{tj}(m) \ell_{tm} = -\ln \sum_m p_j(m) \exp \ell_{tm}. \quad (2)$$

Here Z_{tj} is the prior-weighted likelihood of the ordered calibration sequence. Its logarithm is evaluated with a support-aware shift and an independent scalar evidence sum. Native KL-minus-accuracy components remain in the receipt, but the strict decision uses the scalar log evidence ratio; subtracting rounded components can spuriously break an exact tie. **Equation (2)** is exact for these finite posteriors, not a fitted psychological law. The comparison retains $\Delta F_t = F_{tO} - F_{tR}$, its accuracy and complexity components, and both forecasts. An accessible inquiry selects the open working family only if $\ln(Z_{tO}/Z_{tR}) > 0$; otherwise it retains or returns to the restricted working family. A tie retains the restriction. We import `BMRComparison` for its difference algebra, while recomputing both posteriors: relaxing a zero-support restriction is not an application of the analytic posterior-reuse BMR identity. No gain is manufactured by simply deleting a scalar cost.

Valuing an opportunity is represented separately from selecting a model. Before the current window's comparison is made available to the policy, chosen inquiry $a_t \in \{0, 1\}$ has probability

$$P(a_t = 1 \mid g_{t-1}) = \text{logistic}[\beta_a (\rho_e - c + \lambda_g g_{t-1})]. \quad (3)$$

Here ρ_e values examinability, c charges activity, and λ_g maps previously observed evidence gain into utility. Initially $g_{-1} = 0$. After a genuinely accessible inquiry, $g_t = \max(0, \ln(Z_{tO}/Z_{tR}))$; otherwise the previous observed gain is retained. **Equation (3)** receives neither current-window counterfactual gain nor held-out outcomes. This explicit cost addresses a boundary in the emptiness source, whose model-reduction discussion assumes that metacognitive access has already become available and sets aside its acquisition costs ([Sandved-Smith et al. 2026, sec. 4.3](#), footnote 7, p. 15).

5 conditions distinguish the mechanism. **Rigid** never chooses inquiry. **Inquiry** values examinability and can access alternatives. **No preference** sets $\rho_e = 0$ while retaining costly stochastic and instrumentally reinforced inquiry; it is not a zero-activity control. **Sham** can value and choose activity but exposes no alternative repertoire. **No access** likewise retains preference and activity while blocking metacognitive access. A chosen activity always incurs its declared cost, even when access fails. These semantic interventions specify what would have to differ in an empirical comparison; their activity names do not establish that human practices are matched.

All conditions receive the same calibration and held-out observation arrays and common action/access random numbers within a replicate. Separate random streams generate calibration and assessment conditional on the exogenous context schedule. A working-family choice persists between inquiries; within each window every agent refits its posterior from the declared prior for that family on the current calibration batch. The reset is an explicit episodic design choice, not cumulative lifetime learning. The frozen forecast is then

$$\hat{p}_{tj}(o) = \sum_m A_{om} q_{tj}(m), \quad L_{tj} = -\frac{1}{N_H} \sum_{h=1}^{N_H} \ln \hat{p}_{tj}(o_{th}^H). \quad (4)$$

Equation (4) is evaluated on observations that never select the family or update its posterior. Counterfactual comparisons are retained in the research receipt even for rigid or blocked agents, but are not supplied to those agents’ policies. The inferential unit is the independent paired replicate: losses are averaged over its windows before taking policy-minus-rigid contrasts. Analytic, percentile and studentized interval records are retained; the main table presents analytic pointwise 95% intervals and the supplement retains all retained families. This exploratory extension is separate from the original 8 primary contrasts and does not create a second significance-testing family.

Known-answer controls require useful restrictions to survive favorable evidence, contradictory evidence to permit revision, equal likelihoods to produce no strict preference, invalid channels to fail, consistent candidate permutations to preserve forecasts, and sham/no-access forecasts to remain identical to rigid forecasts. The policy receipt is checked against only previously observed gain. No gate requires superior predictive loss: opportunity value, epistemic benefit and incurred cost remain distinct quantities. The candidate menu deliberately omits a low-probability channel from the restricted family, so reversal is a constructed resolving case, not evidence that relaxation is universally beneficial.

Access and control also have a computational phenomenology distinct from task success. Sandved-Smith and colleagues model perceptual, attentional and meta-awareness levels, separating the ability to access a cognitive state from controlling it (Sandved-Smith et al. 2021). Our binary access gate is an authored simplification, not an implementation of that hierarchical architecture. Neither exposing an alternative in software nor observing a changed response establishes a participant’s reflective awareness.

The action policy is a stipulated logistic choice rule with reinforcement from previously observed evidence gain. It is not expected-free-energy planning over future observations. Its utility units, action precision and gain weight must be interpreted together. Initially, choice depends on the product $\beta_a(\rho_e - c)$: equal increases in opportunity value and cost leave that probability unchanged. These fixed design parameters are consumed by the simulator; they have not been independently identified from behavior.

The identification limit extends beyond the initial choice. Conditional on a fixed gain history, the entire logistic policy depends on only two combinations:

$$\begin{aligned} \eta_0 &= \beta_a(\rho_e - c), & \eta_1 &= \beta_a\lambda_g, \\ (\beta_a, \rho_e, c, \lambda_g) &\mapsto (\beta_a/k, k\rho_e + d, kc + d, k\lambda_g), & k &> 0. \end{aligned} \quad (5)$$

For admissible nonnegative values and costs, this transformation preserves every conditional choice probability. Choices alone therefore cannot identify all four original parameters. Independently varied or measured costs, an explicit utility scale and parameter-recovery checks are necessary before interpreting a fitted preference as gratitude. Observed expenditure can supply information beyond choices; the invariance does not claim that differently charged activities have identical costs.

Exact evidence decisions and conditional operating characteristics

The finite family comparison admits a derivation independent of its software decomposition. Let $\mathcal{R}$ be the retained candidate set, let $w = \sum_{m \in \mathcal{R}} p_O(m)$ with $0 < w < 1$, and let p_R and p_X be the open prior conditioned on $\mathcal{R}$ and its complement. For positive likelihoods, marginal evidence satisfies

$$Z_O = wZ_R + (1 - w)Z_X, \quad \ln \frac{Z_O}{Z_R} = \ln \left[w + (1 - w) \frac{Z_X}{Z_R} \right]. \quad (6)$$

Thus the open family wins exactly when $Z_X > Z_R$. Enlarging support is not automatically rewarded: the evidence of the added candidates must exceed the retained family’s average evidence. The same-data evidence comparison remains conditional on the declared candidate menu and priors; it does not establish the menu’s absolute adequacy (Wilson and Collins 2019). Nor is it the posterior-reuse identity for Bayesian model reduction: a candidate with zero restricted prior mass requires a new posterior fit when admitted.

In the admitted candidate menu, $w = \frac{2}{3}$. Writing $L_m(k) = A_{0m}^k A_{1m}^{n-k}$, the rule becomes $2L_3(k) > L_1(k) + L_2(k)$. The coefficients follow from the admitted open prior and its conditional restriction. The binomial coefficient is absent from each ordered-sequence likelihood and cancels in a comparison of count likelihoods. It must be included when assigning probability to a count under the process. If $\mathcal{K}_O$ contains the winning counts, then

$$P_p(\text{select open} \mid \text{effective inquiry}) = \sum_{k \in \mathcal{K}_O} \binom{n}{k} p^k (1 - p)^{n-k}. \quad (7)$$

The implemented oracle enumerates all $n + 1$ counts and checks native decisions and forecasts against rational products and sums. Decimal strings of the admitted GNN tensor entries define its rational coefficients; open and conditional restricted priors are normalized exactly. The rational count masses sum to one. Logarithms and predictive losses remain numerical evaluations and retain an explicit error check against the native path. This exhaustive finite check replaces sampling error for this calculation, not uncertainty about whether the model represents people.

Let $f_R(k)$ be the restricted channel-zero forecast and $f_S(k)$ the forecast after the evidence-selected family is fitted. With $\ell(p, f) = -p \ln f - (1 - p) \ln(1 - f)$, the conditional expected predictive contrast is

$$\Delta\mathcal{L}(p) = \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} [\ell(p, f_S(k)) - \ell(p, f_R(k))]. \quad (8)$$

A negative value favors evidence selection relative to restriction. Both predictors use the same calibration count; the loss expectation concerns a new observation from the declared process. Selection probabilities and this expected loss are conditional on an effective inquiry. They do not integrate the repeated policy’s action probability, availability, history or cost. The repeated experiment evaluates those additional mechanisms separately.

An observed zero in repeated simulations does not establish impossibility. Zero-numerator bounds require an explicitly defined event and independent trials with a common event probability (Hanley and Lippman-Hand 1983). For an adaptive trajectory, an event such as any revision over the entire trajectory has the complete independent replicate as its trial; correlated windows cannot inflate that denominator. The exact conditional count oracle answers a separate question under its admitted finite channel. It can retain a rare positive selection probability even when a repeated experiment happens to observe no corresponding event.

Recall inference, normalization and learning

Let $a_{os} > 0$ denote the agent’s likelihood pseudocounts. The expected observation probabilities are normalized separately for each hidden state:

$$A_{os} = \frac{a_{os}}{\sum_{o'} a_{o's}}. \quad (9)$$

For prior $p_t(s)$, observed category o_t , and precision $\beta > 0$, the implemented generalized-Bayes update is

$$q_t(s) = \frac{p_t(s) A_{o_t s}^\beta}{\sum_{s'} p_t(s') A_{o_t s'}^\beta}. \quad (10)$$

The operation powers the likelihood values for the observed category. It does not first renormalize every likelihood column after exponentiation. Ordinary online inference uses unit precision; replay uses the declared positive precision. An unsupported observation or invalid probability array raises an error. A missing observation, represented separately from any categorical outcome, returns a copy of the prior. This last behavior is an explicit missingness convention, not a theorem that missingness is always uninformative.

Learning applies forgetting and a belief-weighted increment:

$$a_{os} \leftarrow \rho_f a_{os} + \eta \mathbf{1}\{o = o_t\} q_t(s), \quad 0 < \rho_f \leq 1, \quad \eta > 0. \quad (11)$$

The first observation of each day uses the declared initial prior; subsequent online observations use the transition prediction from the preceding posterior. Initial counts are the declared likelihood multiplied by a prior-concentration parameter. Thus probability shape and accumulated concentration remain distinct. The POMDP and process-theory literature provides the general lineage for this construction (Friston et al. 2017; Sajid et al. 2021; Smith et al. 2022), but the particular update and replay choices above define this experiment.

At day’s end, positive recall samples existing positive observations without replacement, up to the declared recall limit. When fewer positive observations occurred, the replay dose is smaller; the algorithm does not invent favorable experiences. Random recall samples the same number of time points from all observations of that day. Selected indices are processed in temporal order. Both replay conditions use the stored prior from the original time point together with the current learned likelihood. Replay adds another pseudocount update: it is deliberate reweighting of existing experience, not replacement of the original update and not acquisition of independent evidence. Because forgetting is applied per update, sensitivity to forgetting must also be interpreted in relation to replay dose.

Absence observations and counterfactual valuation

The separate absence model includes an observation category explicitly designated as absence. When its likelihood differs between hidden states, observing that category changes the posterior. In contrast, the missing-observation sentinel preserves the prior under the implemented convention. An uninformative likelihood provides a third case: an actual observation occurs, but it does not distinguish states. KL divergence from the prior records the magnitude of updating; the imported routine expresses this comparison in bits. These distinctions prevent “nothing was measured” from silently becoming “harm was absent.” Missing-data methods provide context for the need to state observation assumptions (RUBIN 1976).

Counterfactual value is defined separately as

$$V_{cf} = h_{ref} c_{harm} - c_{actual}. \quad (12)$$

The reference hazard and harm cost are declared inputs, not estimates recovered from the absence itself. In the main study, absence appreciation uses the intervention-phase frequency of the neutral category multiplied by reference hazard and harm cost.

Treating that category as an uneventful interval is a new coding decision, not a claim that every neutral experience is absence of harm. A nonzero value therefore expresses evaluation relative to a reference scenario; it does not prove that an actual harm was prevented.

Controls set reference hazard or harm cost to zero, supply uninformative observations, reject invalid probabilities, and permute hidden-state labels with their corresponding probabilities. Label permutation should preserve predictions. This is a finite representational consistency requirement, offered as one modest operational aspect of non-reification. It is not a comprehensive model of conceptual flexibility. An imported Blackwell Bayes-risk comparison supplies an additional, separate information-ordering control; it must not be mistaken for a theorem that the main gratitude experiment measures philosophical insight.

From learned likelihoods to scene sampling

The gaze extension begins with the exact numerical class in the preserved modified Yarus notebook. Its 13 locations belong to 5 positive, 5 negative and 3 neutral states, with observations ordered positive, neutral, negative. The source comparison alternates its positive and negative likelihood templates for 5 pairs of 1000 states under a single seed of 42. Its state occupancy includes the initial neutral location. We reproduce that draw order rather than seed each profile separately. The notebook contains no gratitude import, parameter learning or before/after intervention. Its cached bars also disagree with its code. Consequently, agreement with the published before bars is a reconstruction result; a new intervention cannot be presented as reproduction of the paper’s after bars.

Task-directed scene viewing supplies relevant background (Yarus 1967; Castellano et al. 2009). The latter experiment manipulated viewing instructions, not gratitude or contemplative practice. Likewise, active-inference models can formalize context-sensitive sampling (Mirza et al. 2019) without establishing that our valence partition is a perceptual fact. Our regional categories and coordinates are stipulated. We inspect Yarus’s publisher chapter metadata/abstract, Castellano and colleagues’ full article, and Mirza and colleagues’ targeted full-text model sections; no claim of reading the restricted Yarus chapter is made.

We retain each corrected agent’s normalized likelihood before and after intervention, frozen before assessment. A declared column-stochastic embedding $E \in \mathbb{R}^{2 \times 13}$ assigns each scene location a distribution over the learned latent states. Each column is nonnegative and normalized; its weights come from the preserved GNN `transfer_state_map`. The primary bridge is

$$A_{i,c,t}^{\text{scene}} = A_{i,c,t}^{\text{learned}} E, \quad \pi_{i,c,t}(a) \propto \exp\{\gamma_T [C^\top A_{i,c,t}^{\text{scene}}(:, a) - \eta_H H(A_{i,c,t}^{\text{scene}}(:, a))]\}, \quad (13)$$

where t denotes pre/post intervention, i a paired agent and c its condition. The GNN declaration sets $\gamma_T = 1.5$, $\eta_H = 0.5$ and $C = (2, 0, -2)$. This source-like policy is stationary within an assessment: the entropy penalty is not called information gain. The subscript distinguishes this weight from the recall learning increment η . Preferences and the common external scene process remain fixed. Pre/post sampled paths share 127 action uniforms after an initial neutral location, giving 128 states per path. Both expected positive-location policy mass and whole-scene predictive loss are retained. Environmental feedback does not update these transferred channels.

A separate sensitivity bridge mixes the fixed source-like positive/negative templates using the independently assessed world-well readout as the mixing weight. It tests a different mapping and consumes assessment information; it is not parameter transfer. The direct bridge is tested for invariance to perturbing that readout. Both mechanisms retain the same agent identities and exposure hashes, and their results are reported separately. The map E , mixture assumption and unchanged scene process are new hypotheses. Neither bridge is claimed to be the authors’ missing after implementation.

Active visual inquiry with revisable structural support

The second gaze extension uses 3 candidate scene representations. The imported Realizing Emptiness admission supplies the restricted prior R , which excludes candidates incompatible with the enforced separation, and the open prior O , which admits the complete menu. The GNN gaze specification declares 3 regional observation channels, a canonical positive-region A tensor, identity transitions within an episode, preferences, geometric coordinates, policy weights, contexts and paired contrasts. Strict admission checks their compatibility with the imported structural family. Native distribution-derived free-energy evaluation is compared with an independent exact evidence sum after every update.

Let m index candidates, a locations and $r(a)$ their region. For posterior $q_t(m)$, the predictive channel is $p_t(o | a) = \sum_m A_{r(a)}(o | m) q_t(m)$. Expected information gain is enumerated exactly:

$$I_t(a) = H[p_t(o | a)] - \sum_m q_t(m) H[A_{r(a)}(o | m)] = \sum_o p_t(o | a) D_{\text{KL}}(q_t(m | o, a) \| q_t(m)). \quad (14)$$

Independent posterior-KL and known-answer tests check Equation (14). Identical candidate channels and a known candidate yield zero information gain. The authored one-step policy separates expected observation value, information gain and movement:

$$\pi_t(a) \propto \exp\left\{\gamma_A \left[\sum_o C_o p_t(o | a) + \lambda_I I_t(a) - \kappa \|x_a - x_{a_{t-1}}\|_2\right]\right\}. \quad (15)$$

The declaration sets $\gamma_A = 4$, $\lambda_I = 4$ and $\kappa = 0.2$ in a unit-square geometry. The information weight λ_I is separate from the practice gain weight λ_g . Utility-only controls set it to zero; a uniform open control samples all 13 locations equally. These

weighted terms define the experiment; we do not claim equivalence to a complete expected-free-energy planner or multi-step optimal policy. Observations are drawn from a shared external regional channel, and the posterior changes only through the selected observations. A zero prior cannot acquire an omitted candidate by ordinary within-family Bayesian updating.

Whole-scene risk prevents an agent from choosing its own favorable evaluation set. Before every observation, we score the predictive distribution against the known synthetic process at every location:

$$L_t = -\frac{1}{13} \sum_{a=1}^{13} \sum_o p_*(o | a) \log p_t(o | a). \quad (16)$$

This is exact population cross-entropy under the declared process, not a held-out human sample or measured calibration. Episode means average this prequential quantity across 48 sampled locations. Stationary windows use candidate sequence (0,0,0) and reversal windows (0,2,0). Each episode resets its admitted prior; candidate identity is stable within the episode. Thus this experiment tests repeated opportunities for structural access under different contexts, not persistent learning of a changing context across windows.

6 conditions separate open access, inquiry and activity: restricted utility, open utility, open inquiry, restricted inquiry, a behaviorally exact restricted sham, and uniform open sampling. Potential action uniforms and time-by-region outcome uniforms are shared across policies. Opening support does not force inquiry, and choosing inquiry cannot recover a candidate that remains excluded. The imported practice model elsewhere makes access itself an enacted decision; this gaze extension conditions on access in order to isolate the sampling consequences.

Scanpaths use the 13 admitted coordinates, normalized to the unit frame and registered to the complete portrait. The licensed museum photograph of Munch’s *The Scream* is credited in each figure (Munch 1893). Straight segments connect sampled locations; circle and square mark the first and last displayed locations. This registration is newly declared, not recovery of the source montage. The painting supplies no pixels to inference, and no gaze coordinates, dwell times or affective ratings were measured. Context changes the synthetic regional channels while the background image remains fixed.

From scene gaze to inquiry into its boundary

A scene can be inspected while its framing assumptions remain unexamined. Conversely, an agent can ask how its observations depend on an action without learning which scene it is viewing. Our intersection experiment makes these complementary operations explicit. It imports the admitted gaze channels and Realizing Emptiness support, then adds a separately declared variable for a predictive boundary/reference encoding. This is joint finite structural inquiry. Calling its question selection *metacognitive gaze* is an operational metaphor, not a claim that ordinary inference has become human meta-awareness.

Nelson and Narens distinguish monitoring of an object-level process from control that changes that process (Nelson and Narens 1990). Their framework motivates recording an encoding posterior separately from the policy that selects evidence. Sandved-Smith and colleagues implement a deeper precision-control and meta-awareness architecture while explicitly setting aside selective epistemic foraging (Sandved-Smith et al. 2021). Our 9-state model is a simpler, different construction. Neither its posterior entropy nor its confidence is a measure of metacognitive sensitivity or efficiency in the sense reviewed by Fleming and Lau (Fleming and Lau 2014).

Let Z denote a reference scene and H a boundary/reference encoding, each with 3 values. Under the admitted independent uniform prior, the diagnostic regime displays appearance W :

$$p(Z = z, H = h) = \frac{1}{9}, \quad W = T_h(Z) = (Z + H) \bmod 3. \quad (17)$$

The 13 visual queries use the existing regional channels conditional on W . They can identify appearance without identifying H . Marginalizing the uniform Z leaves the same visual likelihood for each encoding; this result depends on the admitted prior and cyclic maps, not on a general principle that vision cannot inform a boundary. The restricted condition replaces only the H prior by the actual upstream-admitted support (0.5, 0.5, 0).

A boundary question instead reveals a paired randomized action v and response r . The action is uniform and independent. Query j has declared reliability ρ_j :

$$p_j(v, r | h) = \frac{1}{3} \left[\rho_j \mathbf{1}\{r = (v + h) \bmod 3\} + \frac{1 - \rho_j}{2} \mathbf{1}\{r \neq (v + h) \bmod 3\} \right]. \quad (18)$$

Both individual marginals are uniform; the relation carries information about H . The 3 reliabilities are (0.65, 0.8, 0.95) and costs (0.02, 0.05, 0.1). Visual queries also use a 9-outcome alphabet, pairing their valence observation with an independent uniform nuisance variable. Every query is one declared computational decision; this does not establish equal human perceptual effort, duration or number of fixations. In the separate label-only regime, visual channels depend only on Z and all boundary-question outcomes are uniform. Its labels have no distinguishing observation law.

The joint posterior updates only through the selected query and supplied observation. Zero prior support remains zero. The integrated policy applies Equation (15) over 16 queries, with information evaluated once over the joint state:

$$I_t(j) = \sum_{z,h,o} q_t(z,h) A_{o,z,h}^{(j)} \log \frac{A_{o,z,h}^{(j)}}{\sum_{z',h'} q_t(z',h') A_{o,z',h'}^{(j)}}. \quad (19)$$

Visual utility, action precision, information weight and spatial costs are imported from the actual gaze declaration. Boundary questions have zero immediate visual utility and their separately declared costs. Their selection can therefore trade current-scene sampling against encoding information. Visual-only sampling exactly reduces to the imported gaze policy over appearance; boundary-only sampling admits only the 3 questions; uniform sampling admits all 16 equally. The restricted policy uses the same choices while excluding H2. No gate requires the integrated policy to win.

The matched-query condition is a nonautonomous sham: it follows the paired integrated agent’s schedule and observes the same outcomes, but discards the randomized-action component when updating from a boundary question. The remaining response marginal supplies no encoding update. Its updater deliberately ignores information potentially conveyed by the demonstrator’s adaptive schedule. It is therefore a precisely specified erasure comparator, not a proof that the entire schedule is information-free or a simulation of a person’s unawareness.

In the label-only regime, all query histories preserve the prior odds of encodings:

$$\frac{q_t(H=h)}{q_t(H=h')} = \frac{D_H(h)}{D_H(h')}, \quad D_H(h') > 0. \quad (20)$$

More independent replicates cannot resolve that equality. In the diagnostic regime, scene-only and boundary-only data each leave Z uniform for the open agent, whereas their joint use can resolve the reference scene. This provides an executable intersection without equating identification of a stipulated encoding with realization of emptiness.

Reference-frame transfer and proper boundary scores

The common transfer target explicitly anchors the reference encoding at $H = 0$. It tests whether inferred appearance and encoding jointly support prediction in that reference condition. It is an intervention on the reference frame, not ordinary future prediction under the original unchanged scene process. With final scene marginal q_Z and true reference scene z_* , excess expected log loss is

$$\mathcal{R}_0(q_Z; z_*) = \frac{1}{13} \sum_{a=1}^{13} \sum_o A_{r(a)}(o | z_*) \log \frac{A_{r(a)}(o | z_*)}{\sum_z A_{r(a)}(o | z) q_Z(z)}. \quad (21)$$

This exact population KL score is zero for the true predictive channel. The second endpoint is multicategorical boundary Brier loss, $\sum_h (q_H(h) - \mathbf{1}\{h = h_*\})^2$. It stays finite when the true encoding is excluded; no infinite log loss is silently clipped. Current-scene prequential regret, boundary-question allocation, declared expenditure and posterior entropy remain separate diagnostics. The anchored target assumes an available reference intervention and known channels; those assumptions need independent justification in any empirical translation.

S3. Experimental design, estimands and uncertainty procedures

Paired exposure and matched recall dose

The main design uses 5 fixed luckiness strata and 1024 independent seeded replicates per stratum. The resulting design contains 5120 independently seeded agent-world units, each evaluated in 4 paired conditions. One baseline is trained for each unit and copied into no recall, positive recall, dose-matched random recall, and absence appreciation. Every condition receives the same intervention-phase environmental states and observations. Their shared exposure digest records that identity. Evaluation data and optimism-task priors are independently generated from the training stream but reused across paired conditions and pre/post assessment.

No recall continues ordinary learning. Positive and random recall differ in selected event content while sharing dose and replay precision. Absence appreciation follows exactly the no-recall inference and learning path, then assigns a separate value to designated uneventful observations. This condition supplies an intentionally exact belief null. Any belief difference between it and no recall would indicate an implementation or contract failure; agreement is not evidence that human gratitude can be increased without changing beliefs.

The design is paired within each seeded unit and stratified across a fixed grid. It does not randomly sample people or environments from a population. Its units are independent simulation streams, not independent time steps within a trajectory. Daily trajectory plots retain only replicate zero in each stratum and condition. They illustrate paths through the implementation and must not be read as population means or uncertainty estimates.

Independent predictive evaluation and the optimism readout

Assessment freezes the learned likelihood and evaluates independent observations using the declared prior. It does not continue learning during scoring. For the true binary hidden state y_i and inferred going-well probability q_i , the latent-state Brier error is

$$\text{BS} = \frac{1}{N} \sum_{i=1}^N (q_i - y_i)^2. \quad (22)$$

Brier loss scores predictive probabilities; it does not isolate calibration from discrimination or resolution. We therefore describe Brier contrasts as predictive-loss results. Statistical calibration of an interval procedure is a separate question. The Brier definition follows the probability-scoring lineage (BRIER 1950).

Observation predictive log loss instead evaluates the probability of what was observed:

$$\text{NLL} = -\frac{1}{N} \sum_{i=1}^N \log \left(\sum_s A_{o_i s} D_s \right). \quad (23)$$

These quantities answer different questions. Brier error evaluates posterior discrimination and probability accuracy against the simulated hidden state. Log loss evaluates the observation distribution under the frozen model and prior. Lower values are better for both scores, and neither is a well-being scale. Posterior entropy is reported as uncertainty in natural-log units; lower entropy alone is not a success criterion because an agent can be confidently wrong. Proper scoring provides the relevant methodological context (BRIER 1950; Gneiting and Raftery 2007), while the finite formulas state precisely what is computed here.

Strict propriety concerns expected performance: reporting the data-generating distribution uniquely optimizes the expected score, with our losses using the opposite sign from the reward convention in Gneiting and Raftery (Gneiting and Raftery 2007, secs. 1, 3). It does not guarantee that a fitted forecast wins on every finite assessment sample. Forecast calibration concerns agreement between issued probabilities and outcome frequencies; sharpness concerns concentration of the predictive distributions. An aggregate loss comparison does not separately identify either property, and an entropy decrease cannot substitute for evaluating outcomes. These forecast properties also differ from the repeated-sampling coverage of the interval procedures in Figure 40.

Model evidence and predictive scoring are related but evaluate different uses of observations. Integrated evidence averages a calibration sequence’s likelihood under a declared prior; it can also be factored into sequential predictive probabilities (Gneiting and Raftery 2007, sec. 7.1, equations 50–53). Here that evidence selects a working family, whose fitted forecast is then frozen and evaluated on new observations. Consequently, an evidence preference on the calibration batch does not establish lower future loss. The exhaustive practice calculation evaluates both the process probability of each selection and its expected consequence for a new outcome; the enacted experiment additionally includes action, access and expenditure.

The optimism task uses 2 declared binary news likelihoods, one for each world-belief mode. For each shared initial good-outcome probability p , inference is performed separately after good and bad news in each mode. The resulting posteriors are mixed using the agent’s mean assessed going-well probability w . If these mixed probabilities are q_g and q_b , the readout is

$$O = \text{mean}_p [(q_g - p) + (q_b - p)]. \quad (24)$$

When bad news decreases the good-outcome probability, the second term is negative; the sum contrasts upward good-news updating with downward bad-news updating. Both news conditions use the same initial probability rather than independently resampled priors. The readout is an explicitly constructed function of the world-belief summary and news matrices. It is not an independently validated psychological outcome, and covariance with world belief is partly built into its definition.

Estimands, intervals and Monte Carlo decisions

Let D_{hr} be the difference between two conditions’ pre/post changes for paired unit r in fixed world h . The primary target is the equally weighted mean of world-specific expectations,

$$\theta = \frac{1}{H} \sum_{h=1}^H E[D_h], \quad \hat{\theta} = \frac{1}{H} \sum_{h=1}^H \bar{D}_h. \quad (25)$$

The 8 primary hypotheses combine optimism change and Brier-loss change with positive versus no recall, random versus no recall, absence value versus no recall, and positive versus random recall. This revision retains that family and its data. Additional world-specific and observation-log-loss analyses are exploratory, undertaken after inspecting the original results; they are not retroactively preregistered hypotheses.

Bootstrap resampling draws complete paired units with replacement separately within each fixed world, preserving its weight. It does not resample world labels or count time steps as independent units. The original percentile interval uses the 0.025 and 0.975 quantiles of bootstrap means (Efron 1979). A studentized comparison uses each resample’s own within-world variance, as well as the original estimated standard error (boot package documentation, n.d.):

$$\widehat{\text{SE}}^2(\hat{\theta}) = \frac{1}{H^2} \sum_h \frac{s_h^2}{R}, \quad T^* = \frac{\hat{\theta}^* - \hat{\theta}}{\widehat{\text{SE}}^*}. \quad (26)$$

If q_α^* denotes a quantile of that root, the bootstrap t interval is $[\hat{\theta} - q_{0.975}^* \widehat{\text{SE}}, \hat{\theta} - q_{0.025}^* \widehat{\text{SE}}]$. Studentization is not a universal repair for percentile undercoverage; distributional conditions still matter (Hall 1988). If the original variance or any resampled variance is zero, the studentized interval is unavailable. No problematic draw is dropped, no finite interval is silently substituted, and the degenerate-draw count is retained. In particular, a deliberately constant software null is justified by its mechanism rather than a fabricated nondegenerate interval.

The numerical implementation subtracts each stratum’s first observation before computing its variance and bootstrap roots, and subtracts each resample’s own first observation before computing its variance. These translations preserve the mathematical variances while preventing constant decimal-valued samples from acquiring a small positive variance through floating-point mean rounding. Exact-range controls independently verify degenerate-resample counts; no numerical tolerance or artificial noise turns an unavailable interval into an available one.

Exact availability for the discrete stressor The empirical bootstrap draws independently with replacement from the observed sample (Efron 1979, sec. 2). For our single-stratum centered Bernoulli stressor, centering leaves the variance unchanged. Let $K \sim \text{Binomial}(n, p)$ be the original number of successes. Conditional on $K = k$, a size- n resample is constant with probability $q_k = (k/n)^n + (1 - k/n)^n$. Independence of the B resamples therefore gives the following project-derived availability probability under the stated strict failure rule:

$$A(n, p, B) = \sum_{k=1}^{n-1} \binom{n}{k} p^k (1-p)^{n-k} (1-q_k)^B. \quad (27)$$

Original constant samples contribute zero. The calculation uses the admitted sample size and draw count, and the independent fixture’s generating probability; it neither estimates parameters from the observed availability nor fits a coverage curve. Exhaustive enumeration of original binary samples and all small-sample resampling sequences independently checks the implementation. For fixed finite n and interior p , every interior q_k is positive: A decreases as B increases and tends to zero. This behavior follows from our all-resamples-positive-variance policy, not from a general assertion that more bootstrap draws worsen inference. Dropping undefined roots or substituting a different interval would define a different procedure.

Writing V for interval availability and C for coverage, Figure 40 separately reports $P(V)$, $P(V \cap C)$ and $P(C | V)$. The coverage plotted across all experiments is $P(V \cap C)$ and cannot exceed availability. Conditional coverage describes the selected subset that returned intervals; it is undefined if none did. The exact formula benchmarks $P(V)$ alone and supplies no coverage guarantee. Reporting failures and their data dependence follows the simulation-study guidance of Morris and colleagues (Morris et al. 2019, secs. 3.4, 5.1).

An analytic robustness comparator uses the same mean and within-world variance. With $v_h = s_h^2/R$, its Welch-Satterthwaite degrees of freedom are $(\sum_h v_h)^2 / \sum_h [v_h^2 / (R - 1)]$ (Welch 1947; Satterthwaite 1946). The resulting Student interval is exact for a single Gaussian stratum and approximate for general finite-model contrasts. Percentile, bootstrap t and analytic intervals target the same quantity and remain reported regardless of which appears most favorable.

The two-sided sign-flip test retains the original absolute mean statistic. Its null requires symmetric or appropriately exchangeable paired differences; a zero mean alone is insufficient. Finite Monte Carlo p-values use the plus-one convention (Phipson and Smyth 2010). Holm correction applies to all 8 primary tests (Holm 1979). Exact Clopper-Pearson intervals quantify tail-probability simulation error with a Bonferroni allocation across that family (Clopper and Pearson 1934). A Holm decision is reject only when its adjusted upper Monte Carlo bound is below 0.05, retain only when its lower bound exceeds 0.05, and unresolved otherwise. Retaining a null does not establish equivalence. An independently enumerated small-sample sign oracle checks tail counting and ties.

Inconclusive contrasts and equivalence claims

An interval spanning zero can leave scientifically meaningful effects unresolved. Lakens distinguishes failure to reject a difference from evidence that an effect lies within justified equivalence bounds (Lakens 2017). A future human study would need to define those bounds from its substantive target before using an appropriate equivalence analysis. We have not fitted bounds to these results or applied a simple-test rule to the joint bootstrap rectangles. Algebraic mechanism identities have a different warrant: they follow from the declared operations, rather than from statistical nonrejection. The complete tables therefore retain their exact, stochastic and unavailable statuses.

Joint uncertainty for the exploratory gaze family

The gaze analyses were added after the recall results, and this joint analysis was designed after examining version 0.5.0. Multiplicity adjustment does not make them preregistered. The original 8-test Holm family is unchanged. A separate, explicitly exploratory rectangle covers all 16 transfer and 42 active-policy contrasts in one joint resampling operation.

Transfer resampling draws whole paired-agent vectors independently within each of the 5 fixed worlds; the same selected indices serve every bridge, condition and endpoint. Active resampling first averages all 3 windows within each replicate, shares its indices

across that context’s contrasts, and draws independently between contexts. Matching replicate numbers across separately seeded contexts does not make them paired units. The critical maximum therefore preserves dependence within a block and combines all 3 independent blocks in each draw.

For contrast j , let $\hat{s}_j$ be the original standard error and let $\hat{\theta}_j^{*(b)}$ be its jointly resampled mean. The declared single-step procedure is

$$M^{*(b)} = \max_{j: \hat{s}_j > 0} \left| \frac{\hat{\theta}_j^{*(b)} - \hat{\theta}_j}{\hat{s}_j} \right|, \quad (28)$$

$$\hat{c} = \widehat{Q}_{0.95}(M^*), \quad \mathcal{J}_j = [\hat{\theta}_j - \hat{c}\hat{s}_j, \hat{\theta}_j + \hat{c}\hat{s}_j].$$

Each draw uses the **original** standard errors; this is a standardized empirical-bootstrap maximum, not resample-studentized bootstrap t or a stepdown algorithm. The finite-draw quantile is the order statistic at rank $\lceil 0.95B \rceil$, with $B = 16383$. A separate 99% binomial order-statistic bracket describes Monte Carlo error in that conditional quantile, including tied roots and unbounded finite-rank endpoints. It does not measure the rectangle’s sampling coverage.

The joint-bootstrap lineage requires a consistent joint approximation and appropriate nondegenerate limiting distributions (Romano and Wolf 2005). Covariance-aware simultaneous inference and positive-variance requirements supply related context (Hothorn et al. 2008); the inspected Hothorn source is its author technical report, with journal identity verified separately. Neither citation establishes finite-sample coverage here. We retain all original pointwise procedures and measure all-target coverage on independent Gaussian, correlated Gaussian, unequal-variance fixed-stratum, centered-exponential and rare Bernoulli fixtures. No seed or coverage threshold is chosen to force success.

Valuation and sham zeros receive exact bounds only after their named intervention identities are verified against the complete records. An accidentally constant sample receives no stochastic interval. If any such column is unavailable, the full-family status is partial or unavailable and the supported subfamily is listed; it cannot be advertised as a complete 58-dimensional region. Duplicated or opposite columns retain their dependence without covariance inversion or numerical jitter.

Independent sample expansion and intersection uncertainty

The corrected learning cohort has 1024 independent paired agents in each of 5 fixed worlds (5120 units); active gaze has 512 paired replicates per context; enacted practice has 512 per context. Sensitivity cells use 128 agents per world. Known-answer operating checks use 4096 independent experiments with 64 units per fixed stratum. These independent units are distinct from repeated observations, windows, conditions and bootstrap draws. Original source reconstructions retain 1000 agents and 5 seed-42 gaze runs.

The intersection has 2 regimes, 3 fixed scene values and 3 fixed encodings, each with 512 independent paired replicates and 24 decisions. Thus 9216 independent units generate 55296 6-policy episodes. Counterfactual query outcomes and action-selection uniforms are shared within each unit. Query and observation sequences remain dependent within an episode. Nested prefixes of 16, 32, 64, 128, 256, 512 replicates are fixed in the GNN declaration; increasing N appends independent identities while preserving earlier potential outcomes. Prefix stability is a Monte Carlo diagnostic, not an adaptive stopping criterion or a human-power calculation.

The 7 policy comparisons, 2 endpoints, 2 regimes and 3 encoding contexts define a separate 84-member exploratory family. Of these, 5 comparisons use visual-only as reference; integrated sampling is also compared directly with uniform and boundary-only sampling. Each regime/H block equally averages 3 fixed scene strata and resamples complete paired replicate vectors within those strata. Blocks use independent bootstrap streams, with a shared maximum across all members. The joint bootstrap method and its conditional quantile bracket follow the earlier joint contract; the intersection is not added retrospectively to the 8-test recall family or the 58-member gaze family. Fixed stratum origins are removed before variance and bootstrap centering to avoid floating-point cancellation; a representable-translation negative control reproduces and rejects the former failure. The active-comparator closure was added during development after the first numerical diagnostic run; all intersection analyses are exploratory, not preregistered confirmatory tests.

Mechanism-exact constants require the admitted channel/prior symmetries and complete paired-record checks. Label-only Brier contrasts and the diagnostic sham Brier null are exact; diagnostic boundary-only and sham anchored contrasts are also exact because their scene marginals remain uniform. Restricted label-only Brier offsets are nonzero and can be adverse. Proven offsets are subtracted for exact-zero admission and restored afterward. Other observed constants remain unavailable, and nearly deterministic stochastic endpoints do not inherit an exactness license or a coverage guarantee. All 84 bounds and every nested-prefix estimate are retained in the supplement and data record.

Simulation precision and crossed-design scope

Replication budgets, random-stream offsets, confidence level and diagnostic prefixes are declared model inputs. The analysis protocol is preserved alongside the generating models. Larger replication budgets improve precision conditional on those models; they do not broaden the sampled population or establish human power. Following simulation-study reporting guidance, independent experiments, observations within an experiment and bootstrap resamples retain separate denominators (Morris et al. 2019; Koehler et al. 2009).

For each displayed contrast, complete paired units are averaged within a fixed stratum. The mean across equally weighted strata has estimated Monte Carlo standard error $\sqrt{\sum_h s_h^2 / R / H}$, using the same within-stratum variance as the analytic interval. Repeated practice windows are averaged within their replicate before this calculation. The [Figure 14](#) uses nested prefixes of the same generated units. Those prefixes are correlated, and their stabilization is neither an independent replication nor a stopping rule. The inverse-square-root guide is a precision reference conditional on finite variance, not a fitted law or a criterion for selecting a favorable result ([Koehler et al. 2009](#)).

The recall sensitivity design crosses every declared prior concentration with every declared recall precision. Crossing factors can reveal response changes that a single-factor perturbation misses, but this finite grid neither samples all plausible generating mechanisms nor supplies a global variance decomposition ([Saltelli and Annoni 2010](#)). The grid, world-weight scenarios and operating-characteristic controls therefore answer distinct questions: sensitivity to selected model settings, sensitivity to the estimand’s environmental mixture and performance under known mathematical distributions.

World weighting and distribution-sensitive checks

For each positive-recall comparison, exploratory diagnostics report paired means, pointwise intervals, medians, quartiles and the positive-difference fraction within each world. Brier and observation log loss are losses, so a positive difference is adverse predictive performance for the first condition. An optimism-readout difference has a different interpretation. Unit quartiles describe heterogeneity; they are not uncertainty intervals for a mean. Unadjusted world-specific intervals do not license subgroup significance claims.

For nonnegative world weights summing to one, the target changes to $\theta(w) = \sum_h w_h E[D_h]$ ([Morris et al. 2019](#)). Holding the estimated world means fixed, the range over all such weights is their minimum to maximum. We also omit one world at a time and equally weight the remaining worlds. These are conditional sensitivity ranges, not confidence intervals, new observations or evidence of transport to a human population. They expose assumptions that an equally weighted synthetic average can otherwise hide.

Statistical operating checks follow the distinction between estimand, procedure and data-generating distribution ([Morris et al. 2019](#); [White et al. 2024](#)). 6 independent mathematical fixtures span a standard Gaussian null, a +1 SD Gaussian shift, a variance-standardized Student-t null with 3 degrees of freedom, a centered exponential null, 5 zero-mean Gaussian strata with standard deviations 0.25, 0.5, 1, 2, 4, and a centered Bernoulli(0.1) null. Distribution shapes are known-answer controls, independent of gratitude-model parameters; the GNN design declares replication, sample-size and resampling budgets.

All interval methods evaluate the same generated samples. Skewed exponential and Bernoulli differences intentionally violate sign symmetry. Coverage success is counted over every experiment, with unavailable intervals counted as noncoverage and availability reported separately. Exact binomial intervals and Monte Carlo standard errors accompany rate estimates. Failures and undercoverage remain visible; there is no automatic backup procedure or seed selection ([White et al. 2024](#)). These checks characterize tested mathematical conditions, not empirical power for gratitude or a guarantee of coverage on every project contrast.

Source custody and replay scope

The archived preprint and two preserved notebooks are digest-bound reconstruction inputs, not a moving external dependency. Their selected numerical behavior is reported separately from the corrected experiments. The supplementary reconstruction and attention sections retain the adapter scope and source findings. Every numerical producer consumes the preserved GNN declarations; deterministic replay tests software identity, while independent assessment tests a specified predictive consequence.

S4. Complete results, including adverse and unavailable estimates

Precision across independent sample sizes

The diagnostic retains the signed contrast and its estimated Monte Carlo error at every declared prefix. The recall, reversal-practice, reversal-gaze and diagnostic-intersection comparisons use their own native loss units and fixed-stratum definitions. Their magnitudes should not be compared as a common psychological scale. Quartiles and positive-difference fractions are retained in the machine-readable diagnostic even when the plotted mean is close to zero.

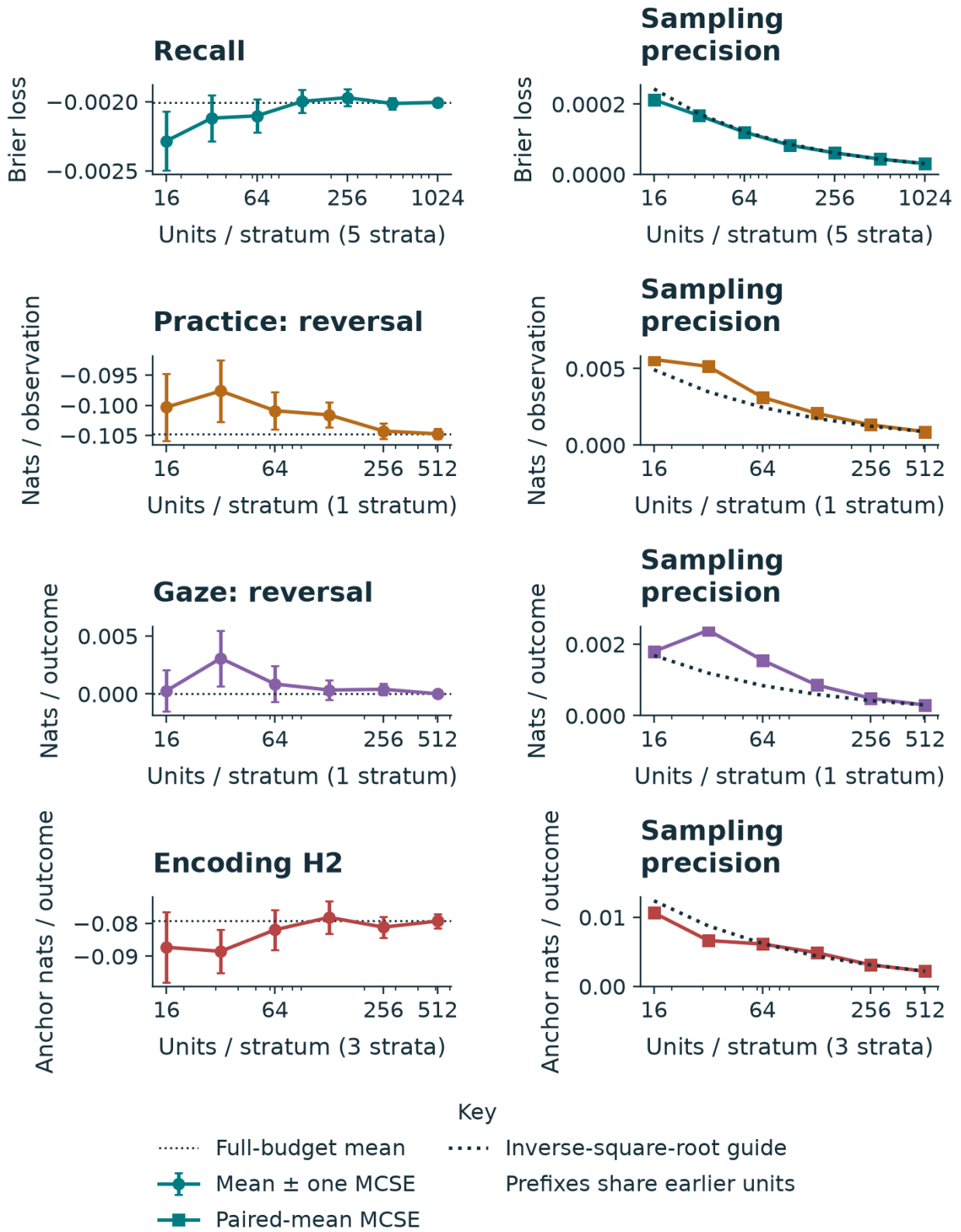


Figure 14: More independent units improve numerical precision within a fixed model. Rows compare positive-minus-random recall Brier change, reversal inquiry-minus-rigid practice loss, reversal inquiry-minus-open-utility gaze loss, and integrated-minus-visual anchored loss for the final declared encoding. Left panels give equal-stratum paired means with one-MCSE whiskers and a dotted full-budget reference. Right panels show the Monte Carlo standard error and an inverse-square-root guide anchored at the full budget. Each row retains its own native units. Prefixes retain whole independent paired units within fixed strata; repeated practice and gaze windows are averaged first. Prefixes share earlier units and are correlated. One-MCSE whiskers do not claim nominal confidence coverage, and the scaling guide is not a guaranteed error rate. Stabilization and larger samples do not validate the generating model, repair observation equivalence or establish human efficacy. An empirical constant is not automatically a mechanism identity.

Exact conditional decisions and rare selections

The exhaustive analysis covered all 25 counts in a window of 24 observations. The open family was selected for channel-0 counts 0, 1, 2, 3, 4, 5, 6, 7, 8; exact ties retain the restriction.

Process probability	P(select open)	Expected loss difference
0.15	0.994126513	-0.341492404
0.85	1.43362441e-08	+4.59820882e-09

These are conditional on an effective inquiry with the stated calibration size. The loss difference is evidence-selected minus restricted prediction in nats per future observation, averaged over every possible calibration count. Probability masses and selections use exact rational arithmetic; logarithms and loss values are evaluated numerically. There is no Monte Carlo interval. Availability, policy history, action cost and human parameter uncertainty are outside this calculation.

The declared process channels (0.15, 0.85) assign different calibration mass to the open-family region: selection is common under the contradicted process, while a useful restriction retains a small but nonzero adverse-selection tail. Selection on evidence can therefore help one declared process and slightly worsen another even before inquiry costs are counted. This is a result of the stated finite repertoire and batch size, not a universal property of openness.

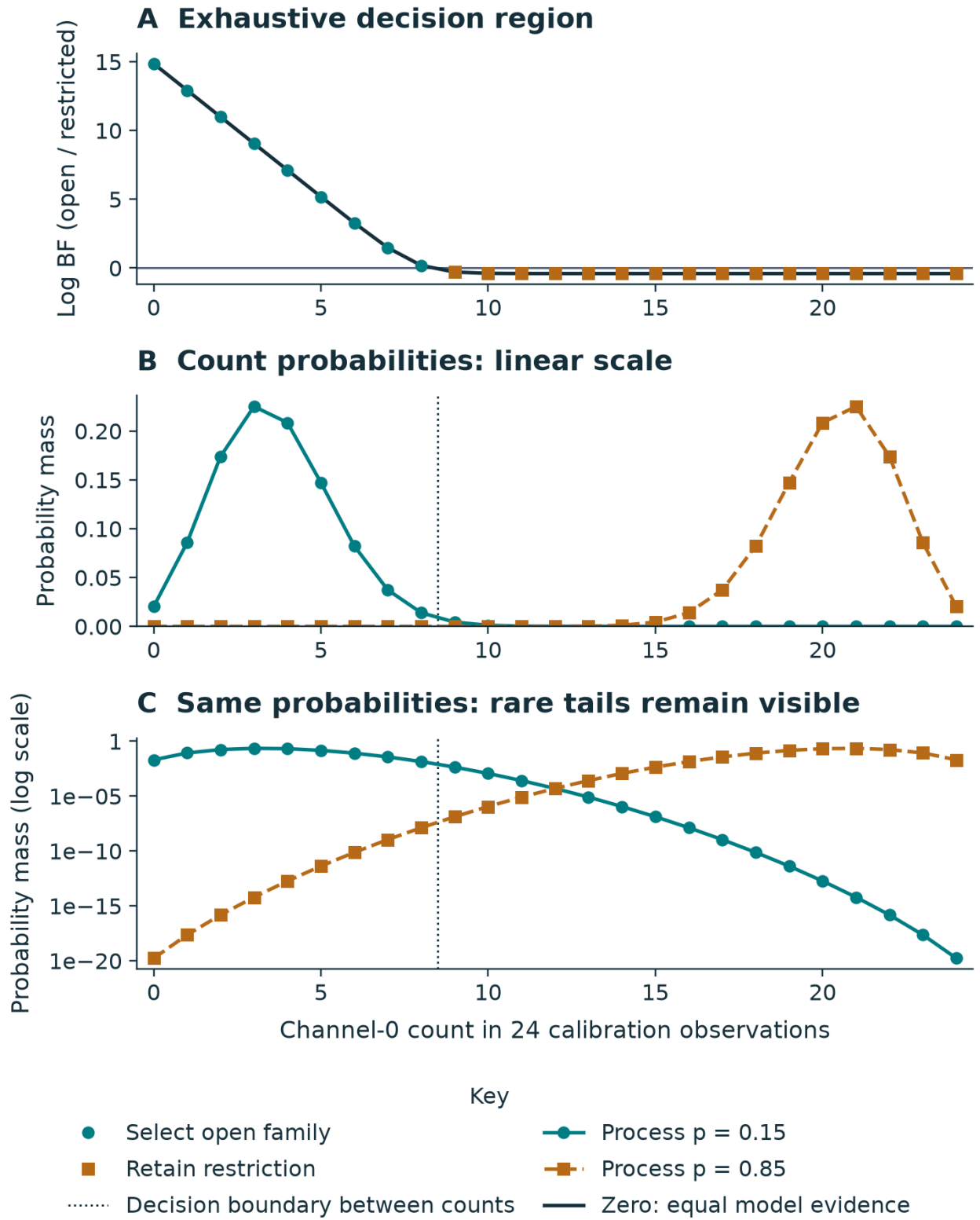


Figure 15: Exhaustive selection exposes both useful and rare adverse decisions. Every possible channel-zero calibration count is evaluated under the same admitted restricted and open families. Panel A shows log Bayes factors: circles select open and squares retain restricted, with zero marking equal evidence. Panels B and C show the same count probabilities on linear and logarithmic scales. Dotted vertical boundaries mark a change in the selected family between adjacent counts. The log view exposes nonzero tails that visually merge with zero in the linear view. Count masses are exact finite binomial probabilities evaluated numerically; they are distributions, not confidence bands, posterior family probabilities or repeated-policy activity rates. Connecting segments guide the eye between discrete counts. This conditional one-window oracle is independent of the seeded trajectory sample and does not establish that unrestricted prediction is universally better.

Figure 15 follows Equation (6); Equation (7) and Equation (8) average over count outcomes rather than sampled replicates.

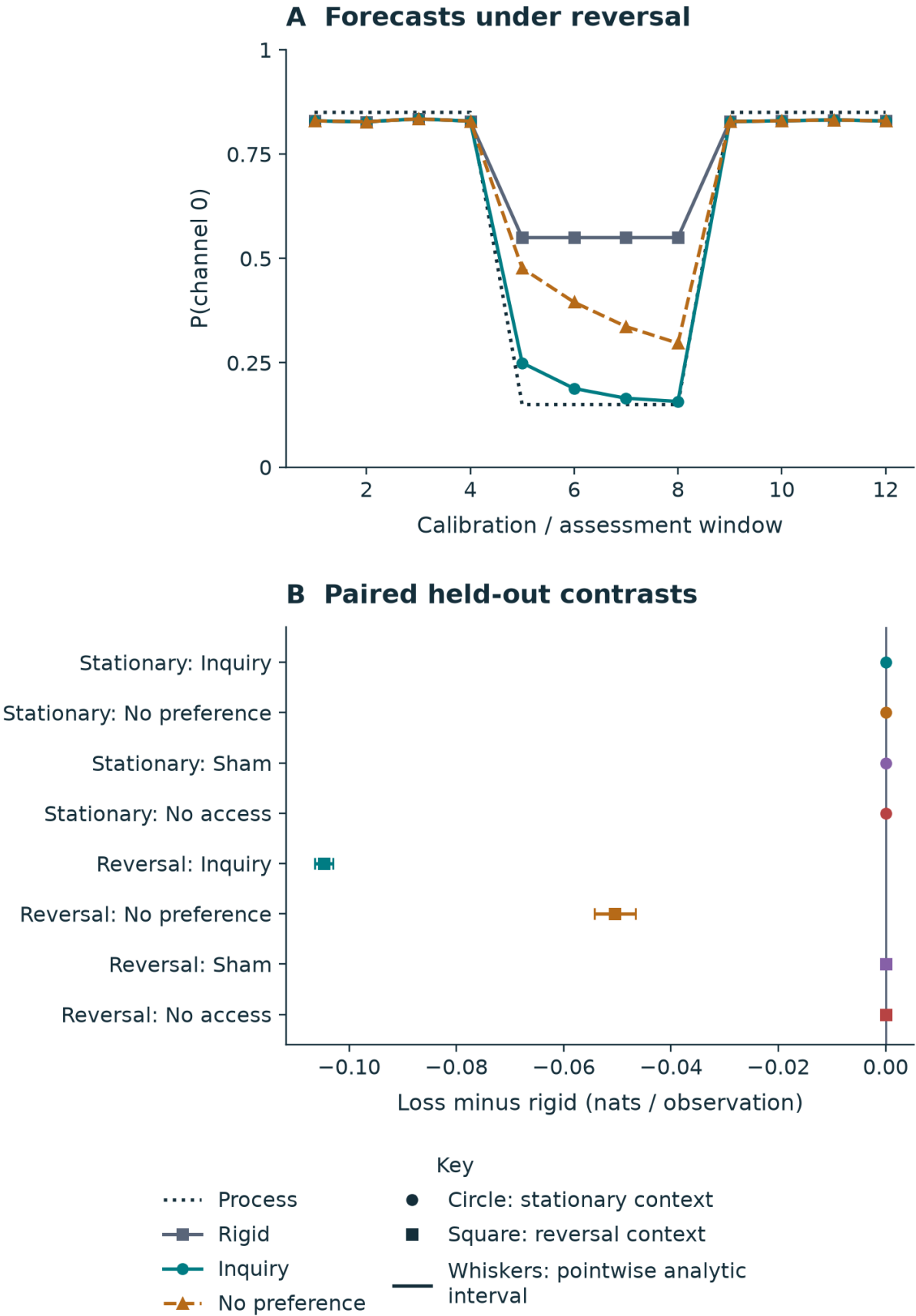


Figure 16: Conditional revision must face new observations. Policies choose and fit a working family from calibration evidence before assessment. The forecast panel compares process probabilities with mean frozen forecasts in reversal windows. Sham and no-access forecasts equal rigid forecasts exactly and are not overplotted. The comparison panel shows policy-minus-rigid held-out log loss. Whiskers are pointwise analytic 95% intervals over independent paired replicate averages; curves are descriptive window means. Negative differences favor the policy in nats per outcome. Inquiry costs are retained separately, and forecast overlap is a mechanism identity rather than missing data.

The separate practice extension used 512 independent paired replicates per context and 12 repeated windows per replicate. Each window contained 24 calibration observations and 512 independently generated assessment observations. Intervals resample replicate averages; repeated windows are not treated as independent units.

Context / policy	Chosen activity	Effective access	Cost / window
Stationary / Inquiry	0.768	0.768	0.154
Stationary / No preference	0.225	0.225	0.045
Stationary / Sham	0.768	0.000	0.154
Stationary / No access	0.768	0.000	0.154
Reversal / Inquiry	0.834	0.834	0.167
Reversal / No preference	0.340	0.340	0.068
Reversal / Sham	0.760	0.000	0.152
Reversal / No access	0.760	0.000	0.152

Context / policy	Loss difference	Analytic t 95% interval
Stationary / Inquiry	+0.0000	[+0.0000, +0.0000]
Stationary / No preference	+0.0000	[+0.0000, +0.0000]
Stationary / Sham	+0.0000	[+0.0000, +0.0000]
Stationary / No access	+0.0000	[+0.0000, +0.0000]
Reversal / Inquiry	-0.1047	[-0.1064, -0.1030]
Reversal / No preference	-0.0505	[-0.0543, -0.0466]
Reversal / Sham	+0.0000	[+0.0000, +0.0000]
Reversal / No access	+0.0000	[+0.0000, +0.0000]

Loss differences are policy minus rigid, in nats per held-out observation; negative values favor the policy. All intervals are pointwise and exploratory, outside the original primary Holm family. Cost is a declared utility charge, reported separately rather than subtracted from predictive loss. The supplement reports percentile and studentized bootstrap intervals for all 8 contrasts.

All finite mathematical and semantic controls passed.

The policy results make the distinction between aspiration and enactment visible. Sham and no-access agents can value inquiry and spend its cost while retaining the rigid forecasts exactly. The no-preference condition can still inquire stochastically and reinforce an observed gain, so its comparison tests the added declared value of examinability rather than all inquiry versus none. Activity rates, effective access, predictive loss and cost must therefore be read together.

The stationary and reversal contexts also distinguish useful retention from conditional revision. The restricted repertoire contains good predictors for the stationary process. The reversal interval supplies evidence that its omitted candidate may matter, and the final interval permits a return to a useful restricted family. These are deliberately diagnostic contexts. Forecast trajectories are averages of simulated replicates; a trajectory returning toward a channel probability is not a trajectory of contemplative attainment.

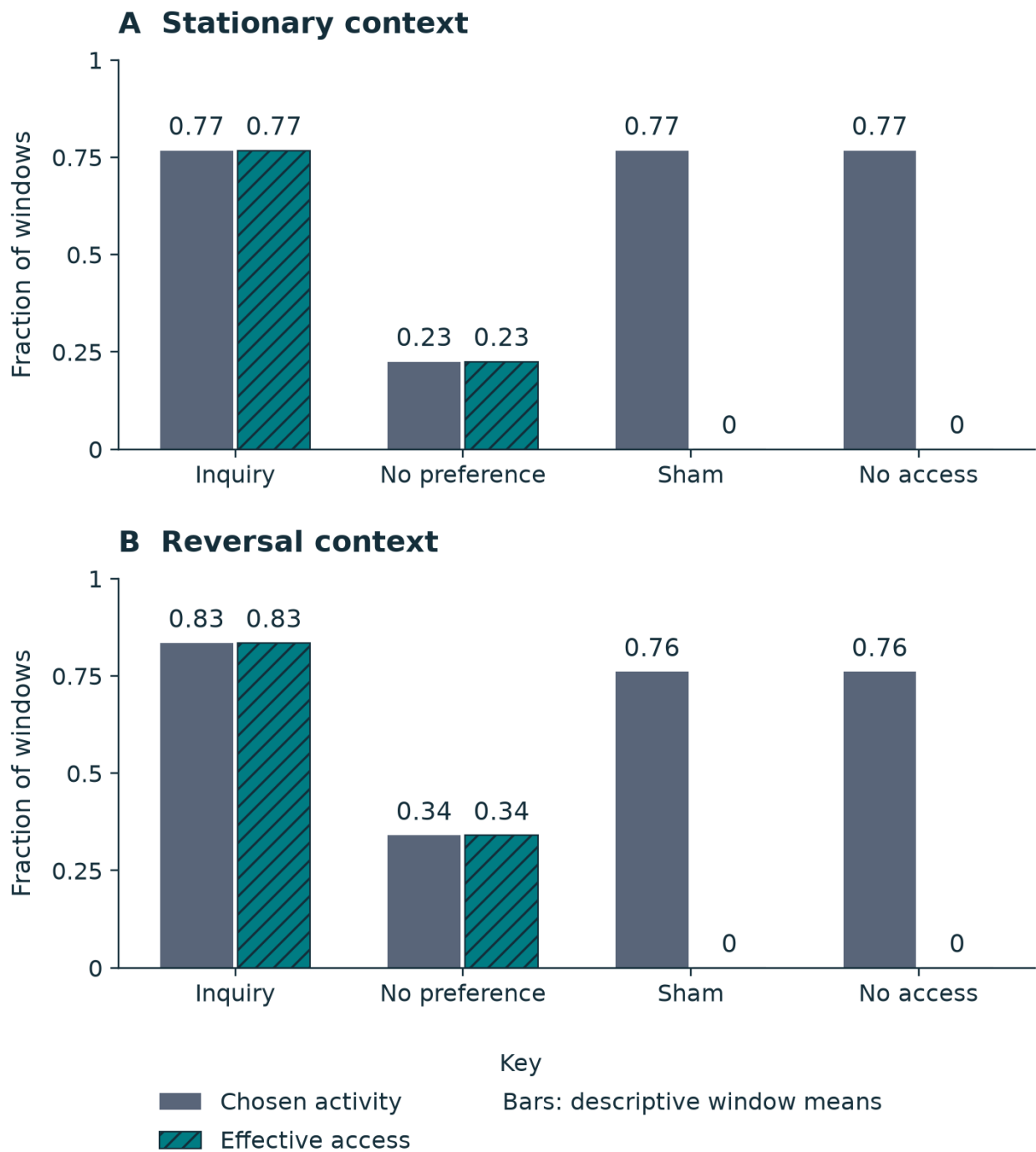


Figure 17: Choosing an activity does not guarantee access to alternatives. Matched policy conditions separate declared motivation, chosen activity, effective access and expenditure. Within each pair, the solid gray bar on the left shows the fraction of windows with chosen activity; the hatched teal bar on the right shows effective access. Labels give means to two significant digits; a visible zero marks no access. Each context uses 512 independent paired replicates and 12 repeated windows. Bars are descriptive finite means without inferential whiskers; windows are not additional independent agents. Sham and blocked access can incur activity costs while providing no access. These rates describe software choices, not human adherence or insight.

The upstream interoperability receipt admits 3 practice identifiers and passes the pinned ongoing-revision controls. This is source-qualified software interoperability, not an independent held-out result for the upstream diagnostic. The new assessment split, action policy, acquisition/activity charge and blocked-access controls belong to Thanks for Nothing. Full numerical receipts retain the exact model bytes and all comparisons, including nulls and any unfavorable scores.

The exhaustive stationary probability makes the seeded exact null interpretable: no open-family selection in a particular stationary run does not imply that such selection is mathematically impossible. Its probability is small but positive for this batch size. The conditional oracle and repeated-policy estimates have different denominators and should not be substituted for one another.

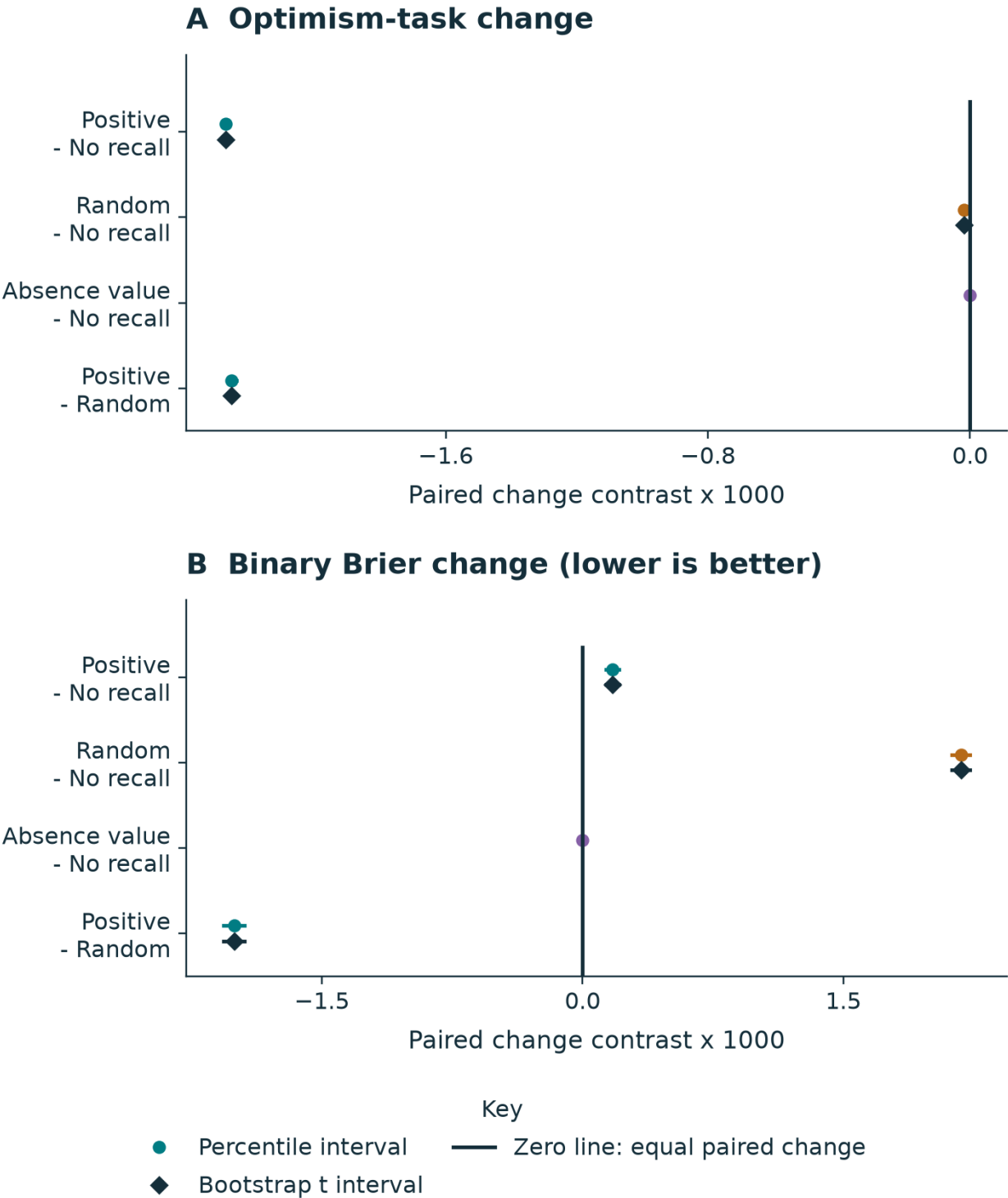


Figure 18: Complete primary recall contrasts on the fixed world grid. Each target is an equally weighted mean of paired post-minus-pre changes across the declared worlds. Circles show percentile intervals and diamonds studentized bootstrap intervals; horizontal values are multiplied by 1000. Negative Brier contrasts favor the first condition, while the optimism readout has a separate interpretation. Both interval types are pointwise 95% intervals over complete agents within fixed strata. The exact absence-value null has no studentized interval because variance is zero. Interval overlap does not encode the primary Holm decision; sign-flip assumptions and Monte Carlo qualification remain separate.

The paired comparisons distinguish three questions: what additional rehearsal does relative to continued learning, what positive selection does beyond an equal replay dose, and whether a separate valuation changes inference. Changes in the declared optimism readout are interpreted alongside latent-state Brier error and observation log loss. Agreement among these measures is possible, but is not imposed as an acceptance criterion. Their disagreement would be informative because each scores a different property of the model.

Optimism-task change. Estimates and interval bounds are multiplied by 1000.

Contrast	Mean	Bootstrap t 95% interval	Holm p; MC decision
Positive - None	-2.270	[-2.278, -2.262]	0.00098; reject
Random - None	-0.017	[-0.031, -0.004]	0.23535; retain
Absence - None	0.000	Unavailable	1.00000; retain
Positive - Random	-2.253	[-2.267, -2.239]	0.00098; reject

Brier-loss change. Estimates and interval bounds are multiplied by 1000.

Contrast	Mean	Bootstrap t 95% interval	Holm p; MC decision
Positive - None	0.171	[0.135, 0.207]	0.44312; retain
Random - None	2.177	[2.125, 2.229]	0.00098; reject
Absence - None	0.000	Unavailable	1.00000; retain
Positive - Random	-2.005	[-2.066, -1.946]	0.00098; reject

Intervals in this primary table and its accompanying figures are pointwise and unadjusted; the 8-test Holm family retains simultaneous Monte Carlo tail bounds. The separate gaze rectangle is defined in [the joint method](#). Retain does not establish equivalence. Bootstrap t is unavailable for the exactly constant valuation null; its zero difference follows from the mechanism. The interval comparison appendix retains percentile and analytic t intervals. Lower Brier loss is better predictive performance, not a calibration-only diagnosis.

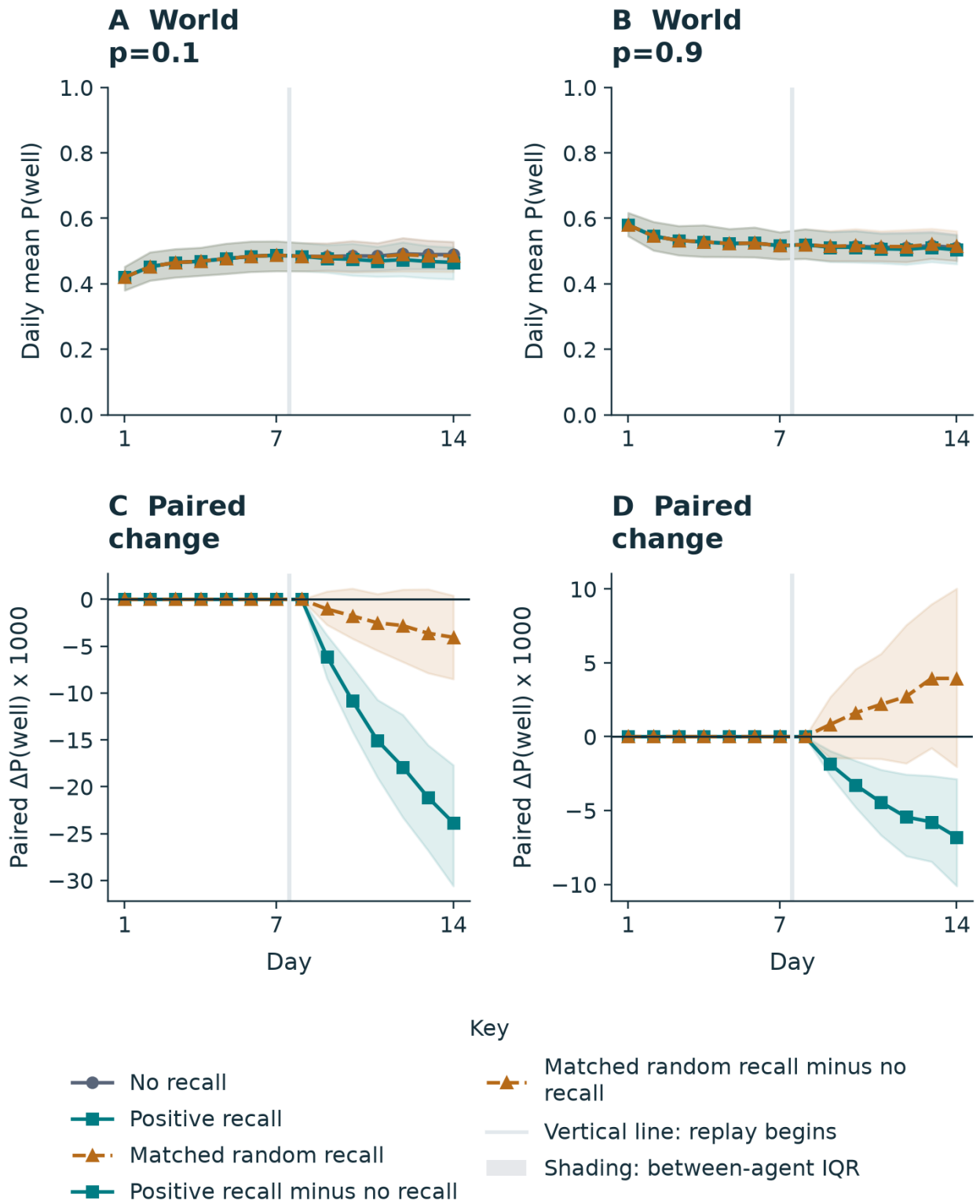
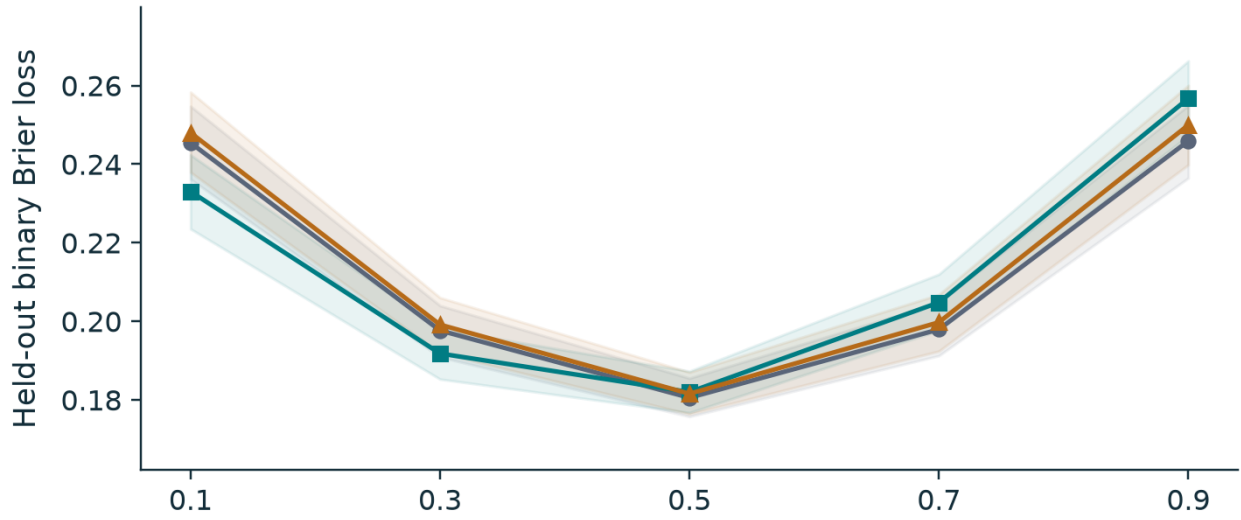


Figure 19: Ensemble means and paired changes answer different trajectory questions. The extreme admitted worlds retain all 1024 independent paired agents per world. Upper panels show mean daily beliefs; lower panels subtract each agent’s no-recall trajectory before averaging and multiply differences by 1000. Markers identify conditions, and the vertical line starts rehearsal. Shaded bands are between-agent interquartile ranges, including ranges of paired changes below; they are not confidence intervals. Days and decisions repeat within agents. Shared baselines produce exact paired zeros before rehearsal. Independent predictive assessments, rather than a smooth belief trajectory, evaluate predictive consequences.

A Means and unit IQRs (n=1024/world)



B Pointwise analytic t intervals

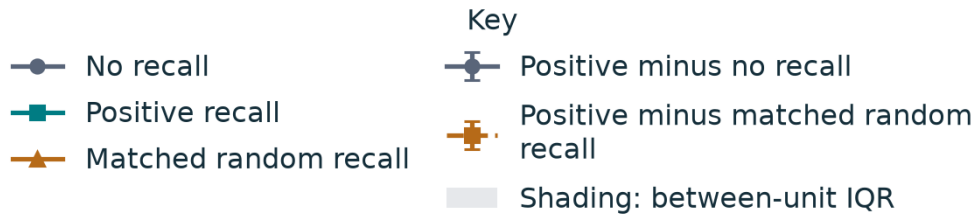
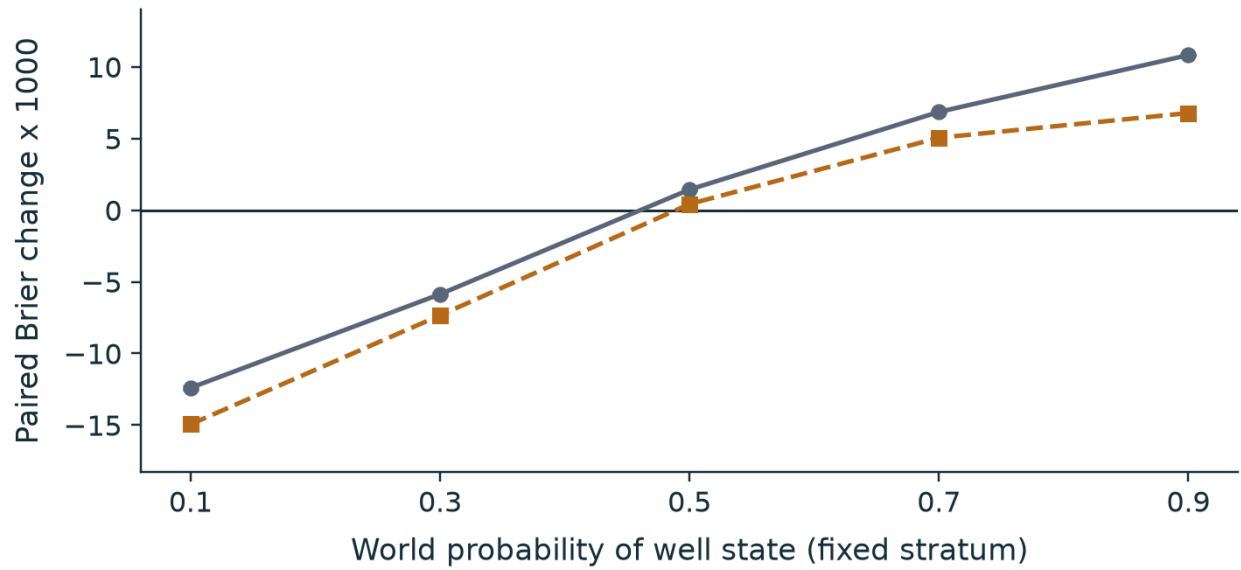


Figure 20: Predictive loss depends on the world being modeled. The display separates absolute post-learning Brier performance from paired changes caused by replay. The absolute-loss panel shows means and between-unit interquartile ranges; the contrast panel compares positive recall with its named references at each fixed world probability. Contrast whiskers are pointwise analytic 95% intervals over 1024 independent paired units per world; interquartile ranges describe spread and are not confidence intervals. Brier is a loss, so negative contrast values favor positive recall. These exploratory world comparisons do not establish subgroup significance or isolate calibration from other score components.

World heterogeneity and mixture sensitivity

For positive minus no recall, Brier-loss change had an equal-grid mean of 0.000171; 2 world strata had negative means and 3 had positive means. Holding stratum estimates fixed, arbitrary nonnegative world weights yielded $[-0.012401, 0.010826]$. The leave-one-world-out range was $[-0.002493, 0.003314]$.

For positive minus matched random recall, observation log-loss change had an equal-grid mean of 0.002885; 2 world strata had negative means and 3 had positive means. Holding stratum estimates fixed, arbitrary nonnegative world weights yielded $[-0.000541, 0.009078]$. The leave-one-world-out range was $[0.001336, 0.003741]$.

These are exploratory descriptions. Pointwise intervals are not adjusted for subgroup selection, and weighting ranges are sensitivity scenarios rather than confidence intervals. A different world mixture changes the target; opposite signs do not establish individual human benefit or harm.

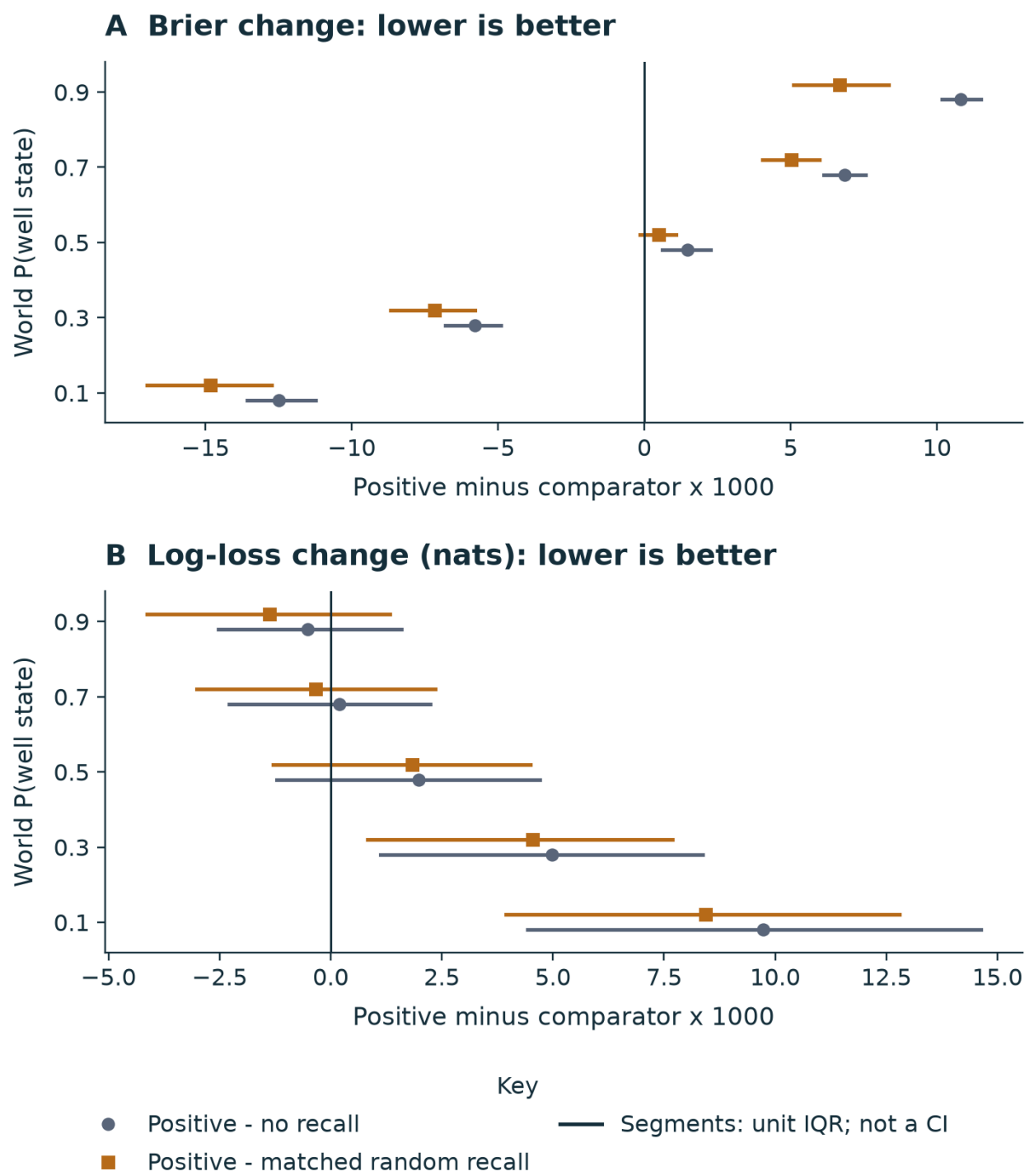


Figure 21: A mean effect can conceal variation between paired agents. The display examines the distribution of predictive-loss differences within each fixed world. Circles and squares mark medians for the named positive-recall comparisons; horizontal segments show the interquartile range of complete paired units. The vertical axis is quantitative world probability. The segments are distribution summaries, not confidence intervals; repeated decisions do not enlarge the sample. Negative loss differences favor positive recall. These exploratory spreads complement, rather than replace, the equally weighted primary target.

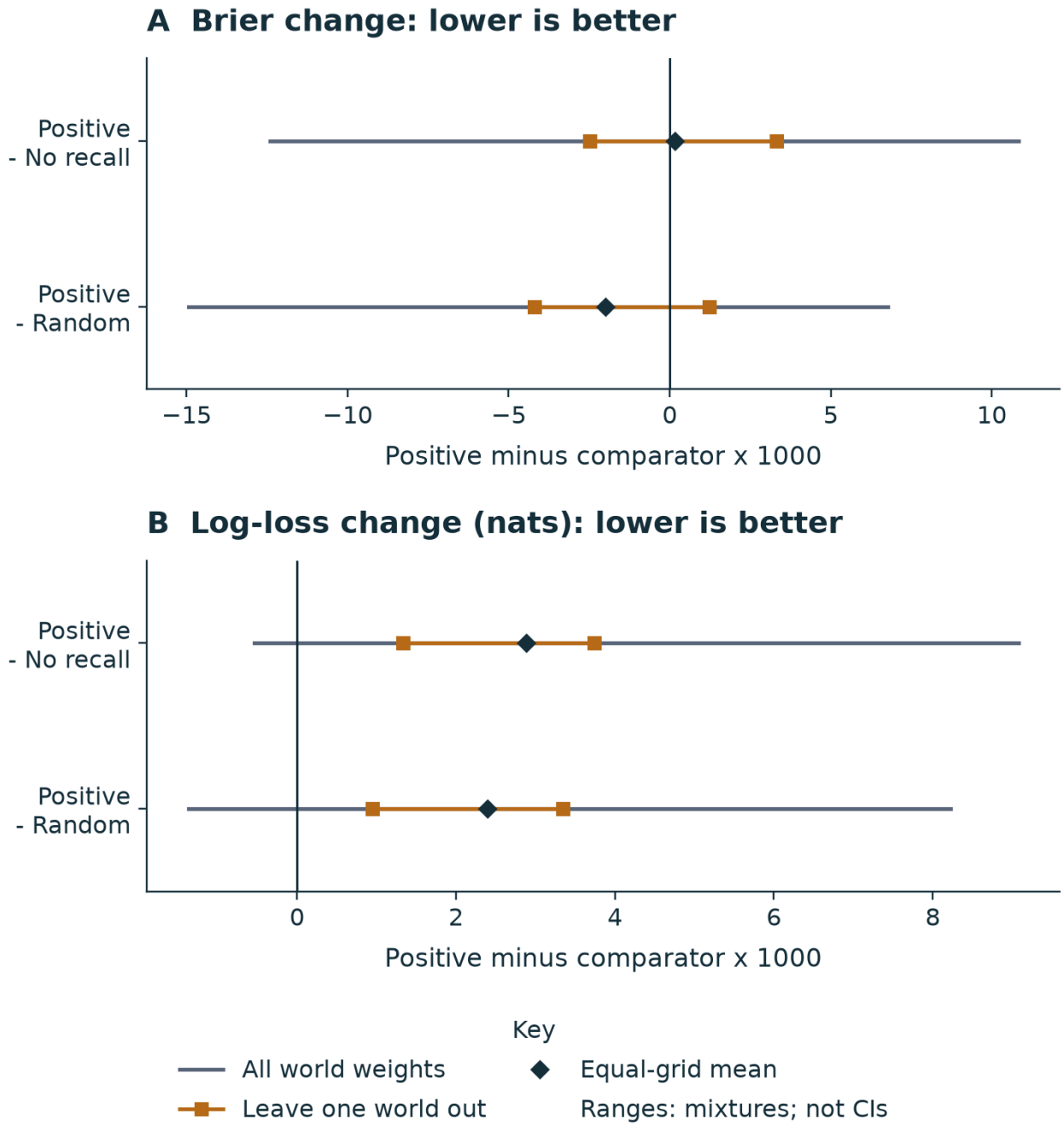


Figure 22: Changing world weights changes the estimand. The analysis reweights the same estimated world means without generating new agents. Gray segments span all convex world mixtures, orange segments span leave-one-world-out estimates, and diamonds mark equal weighting. Horizontal values use the multiplier 1000. These are ranges conditional on fixed estimated means, with no sampling-confidence interpretation. A different population mixture can change the aggregate conclusion. The ranges quantify that dependence rather than adding independent evidence.

Observed absence and the valuation-only null

Missingness preserved the prior with KL divergence 0.0000 bits; an observed absence produced 0.2577 bits. The imported finite Blackwell comparator returned Bayes risk 0.2000 for its original channel and 0.5000 under erasure. These are separate declared finite comparators, not measured psychological or contemplative changes.

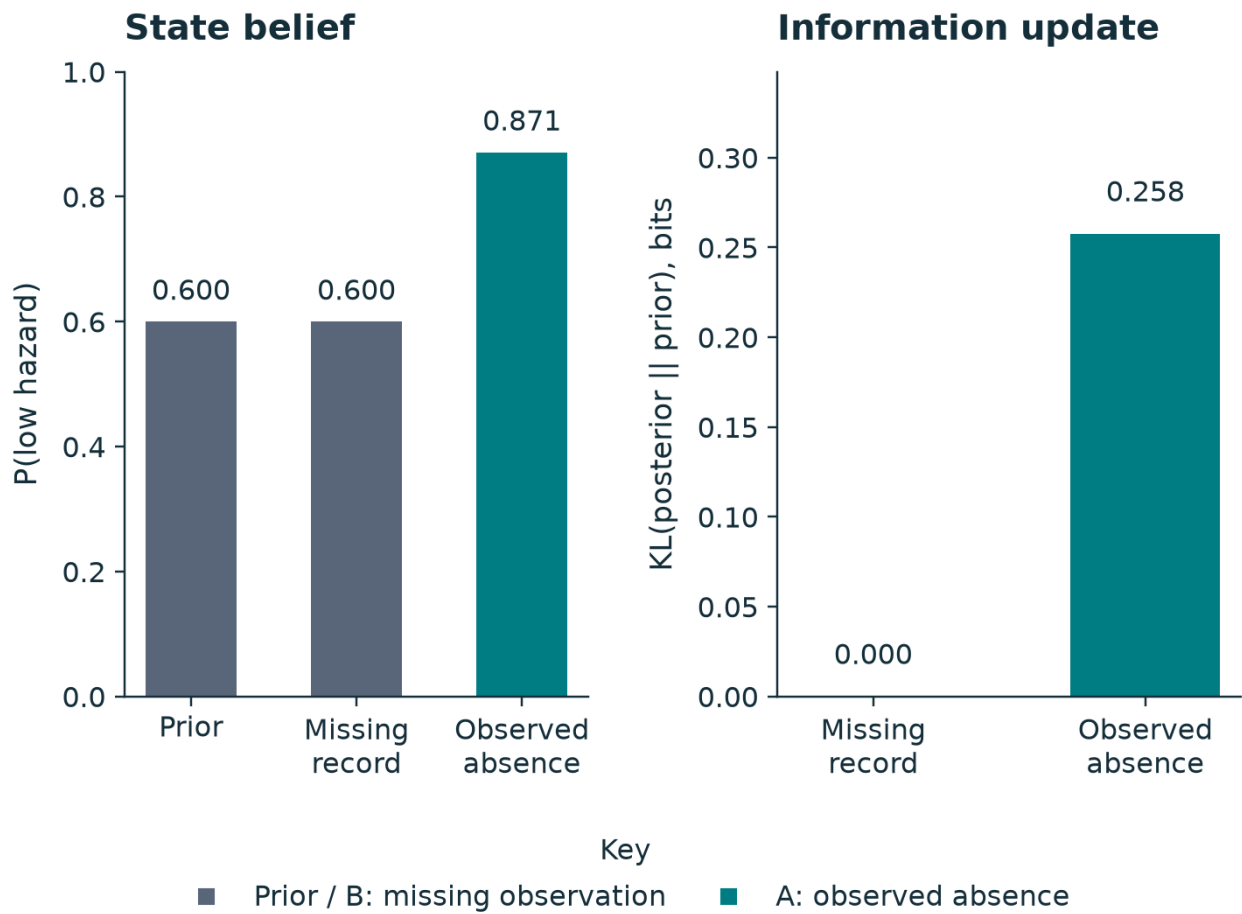


Figure 23: A recorded absence can inform; an unavailable record need not. The fixture compares the same hazard prior under a functioning detector’s no-event result and the declared missing-observation rule. Bars show state probabilities; annotations report KL information in bits for A: Observed absence and B: Missing observation. These are exact admitted finite calculations with no sampling intervals. The no-update missingness rule is stipulated. This control distinguishes informational semantics and does not measure gratitude or contemplative insight.

Absence appreciation and no recall have the same belief-update path by construction. Their equality is an executable null, and any assigned appreciation is a reference-dependent valuation. The separate absence audit distinguishes an informative absence observation from an uninformative observation and from missing evidence. This distinction is a property of the admitted likelihood and missingness convention, not a claim that every real absence is informative or that an undetected harm is safe.

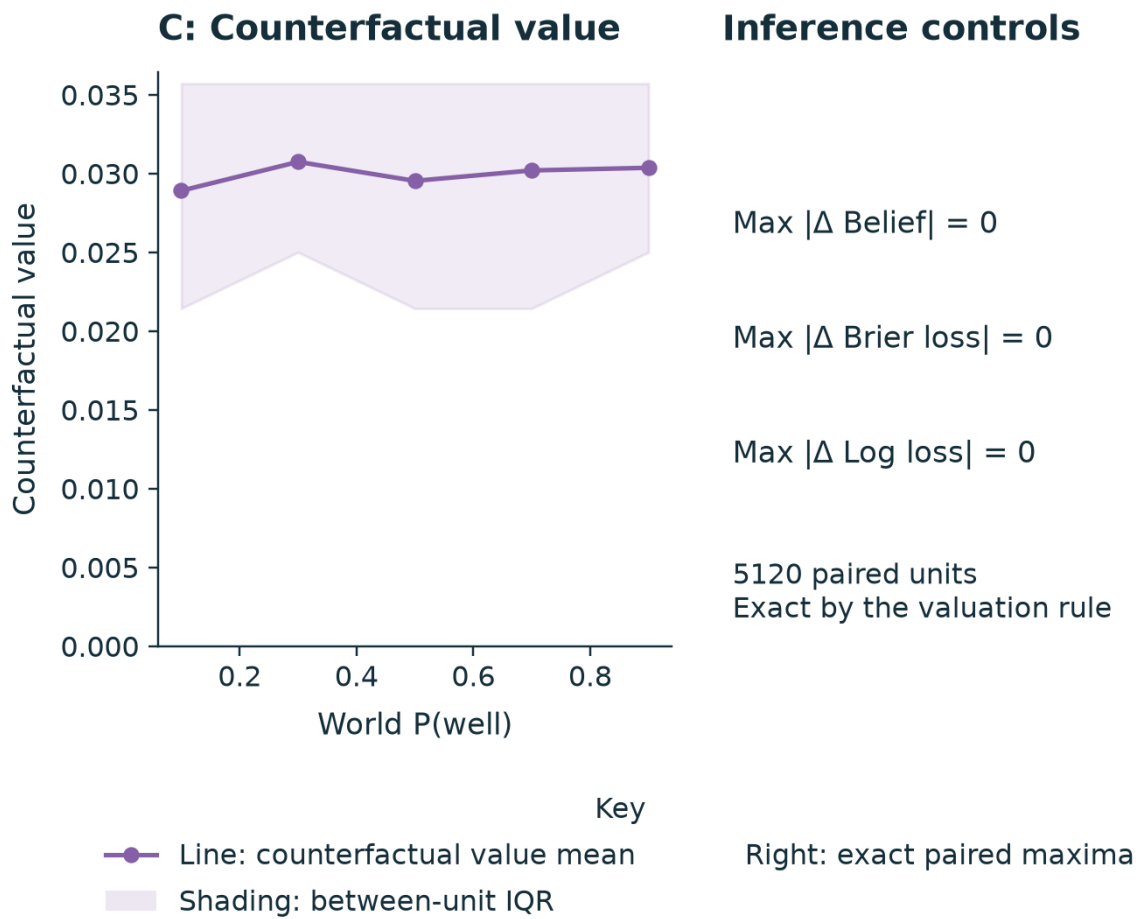


Figure 24: Counterfactual value can change without new belief evidence. C: Counterfactual value evaluates an outcome against a specified alternative while preserving the no-recall inference path. The value panel shows valuation means and between-unit interquartile ranges. The equality panel reports maximum paired differences in belief, Brier loss and observation log loss. Interquartile ranges describe unit spread; the paired maxima are exact numerical control quantities rather than estimated confidence bounds. Equality is imposed by the valuation mechanism. It is a semantic null, not empirical evidence that human appreciation is independent of belief.

Reconstructing and extending the gaze experiment

The exact 5-pair source run reproduces the published before bars to one decimal place:

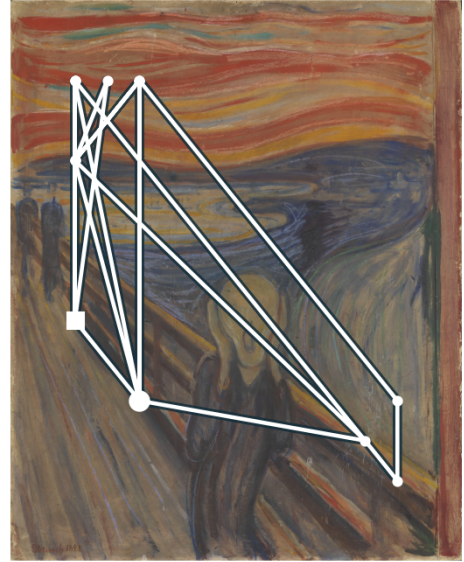
Source profile	Positive (%)	Negative (%)	Neutral (%)
Positive template	82.14	10.98	6.88
Negative template	54.62	11.80	33.58

The paper's after bars are visible as 90.1/6.1/3.8 and 90.6/2.3/7.1 percent in positive/negative/neutral order. They are transcribed observations about its figure, not reproduced values. The preserved executable notebook supplies no corresponding intervention. **Figure 25** therefore shows only the admitted before comparison, with original numeric paths on a newly registered licensed background.

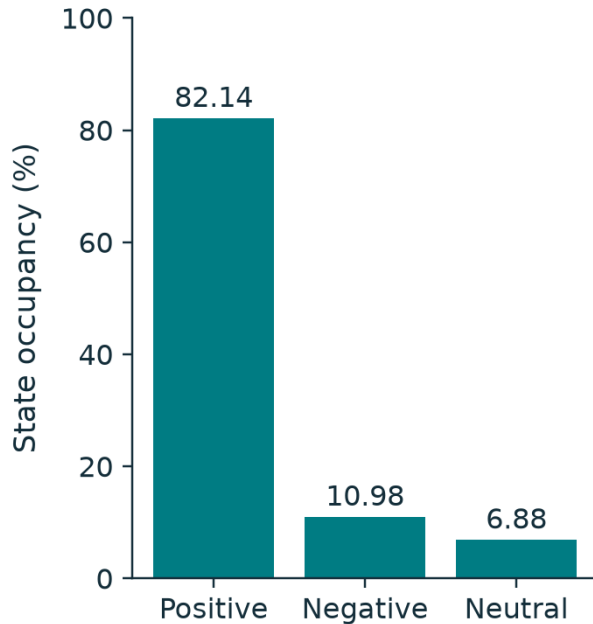
A Source positive



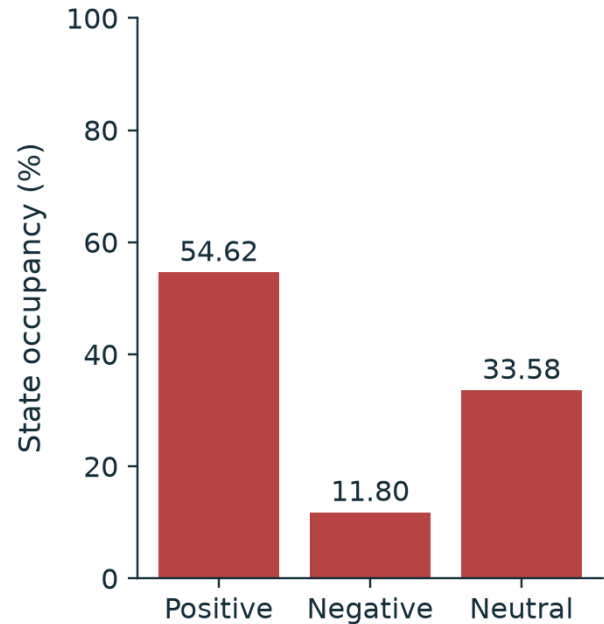
B Source negative



C Source positive



D Source negative



Munch (1893); photo Nasjonalmuseet / Børre Høstland (CC BY 4.0).
Rescaled background with model paths; no measured eye movements.

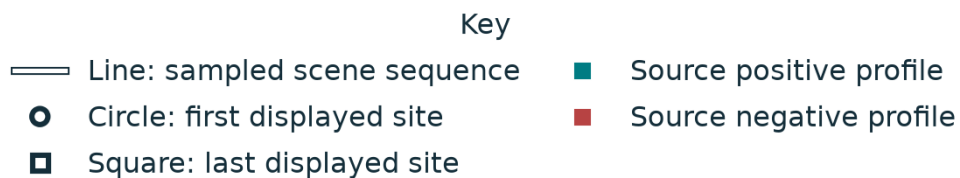
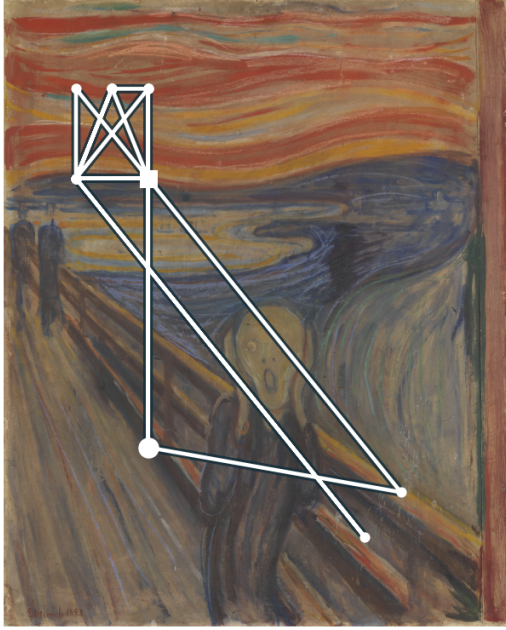


Figure 25: The preserved notebook reproduces source-before occupancy only. The same seed 42 and alternating run order drive the source adapter for 5 runs. Paths show the first 20 locations of the first positive and negative runs; occupancy summaries use every source run. Lines connect discrete registered model sites. Paths are illustrative and occupancy summaries descriptive; no executable source-after intervention or uncertainty interval is supplied. This is a new full-frame registration, not recovered eye tracking. Artwork: Edvard Munch, The Scream (1893), NG.M.00939; photograph Nasjonalmuseet/Børre Høstland; CC BY 4.0; rescaled with model overlays.

Direct likelihood transfer gave a positive-recall-minus-no-recall change in positive-location probability of $+0.038492$ [$+0.038312$, $+0.038670$]; multiplying by 100 gives 3.849 percentage points. The corresponding scene-risk change was $+0.002362$ [$+0.002175$, $+0.002557$] nats per outcome, indicating worse prediction under this bridge. The readout bridge instead gave -0.004677 [-0.004693 , -0.004661] probability units. All brackets are pointwise bootstrap t intervals, not multiplicity-adjusted tests. Open inquiry minus

restricted utility changed scene-wide prequential log loss by $+0.009639$ $[+0.008892, +0.010535]$ in the stationary context and -0.115884 $[-0.116824, -0.114765]$ in the reversal context. Access to the omitted candidate explains the reversal advantage: open utility and uniform open sampling also improve over the restricted reference. The full table retains the direct inquiry-versus-open-utility and inquiry-versus-uniform comparisons; their risk intervals include zero.

A $p = 0.3$: before recall**B $p = 0.3$: after recall****C $p = 0.7$: before recall****D $p = 0.7$: after recall**

Munch (1893); photo Nasjonalmuseet / Børre Høstland (CC BY 4.0).
Rescaled with model overlays; replicate zero is illustrative, not a mean.

Key

- Line: sampled scene sequence ■ Square: last displayed site
- Circle: first displayed site

Figure 26: A learned channel is transferred through an explicit bridge. Frozen pre/post likelihoods from the same paired agent drive a distinct scene-sampling task. The positive-recall examples at world settings 0.3, 0.7 show the first 20 of 128 states. The process, preferences and action uniforms remain paired. These first-agent paths are illustrations without inferential intervals. The display uses the likelihood embedding, not the alternative readout projection or unavailable source-after code. Artwork: Edvard Munch, The Scream (1893), NG.M.00939; photograph Nasjonalmuseet/Børre Høstland; CC BY 4.0; rescaled with registered model overlays.

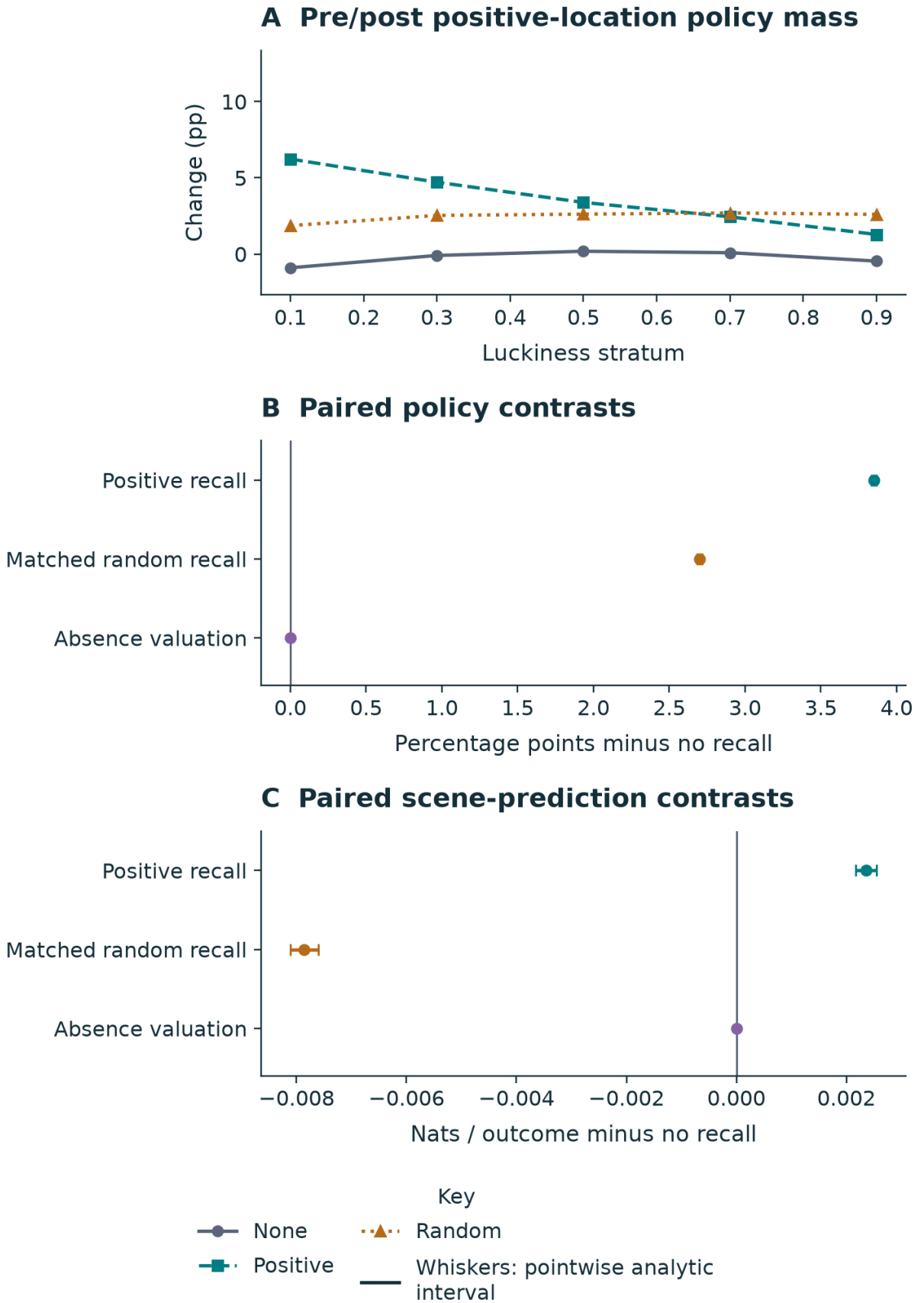


Figure 27: An attention shift can coexist with worse prediction. The likelihood bridge transports actual frozen learned channels into a separate scene task. The upper panels show descriptive pre/post positive-site policy changes by world. The lower panels show equally weighted paired differences from no recall, keeping attention and scene-loss units separate. Whiskers are pointwise analytic 95% intervals over complete paired agents. Exact nulls have zero-width analytic intervals and unavailable bootstrap t intervals. More positive-location policy mass is not necessarily better prediction. The full tables preserve both bridges and every contrast.

The transfer design contains 5 fixed strata with 1024 paired agents each and 4 conditions, evaluated under 2 bridges. The resulting 40960 records are repeated conditions and mechanisms for 5120 independent agent-world units, not 40960 independent participants. We resample whole paired agents within strata and report 16 exploratory transfer contrasts: 4 comparisons, 2

endpoints and 2 bridges.

The active experiment contains 512 paired replicates per context, 3 windows and 6 policies: 18432 episodes. The 48 observations per episode are dependent measurements, not independent units. Windows are averaged within each replicate before resampling. The 7 policy comparisons, 3 endpoints and 2 contexts yield 42 exploratory active-policy contrasts. Every gaze contrast uses 4095 bootstrap draws, with percentile, bootstrap t and analytic t intervals retained. These 58 pointwise intervals do not extend the original 8-test Holm family and carry no confirmatory p-value claim.

A Stationary: restricted



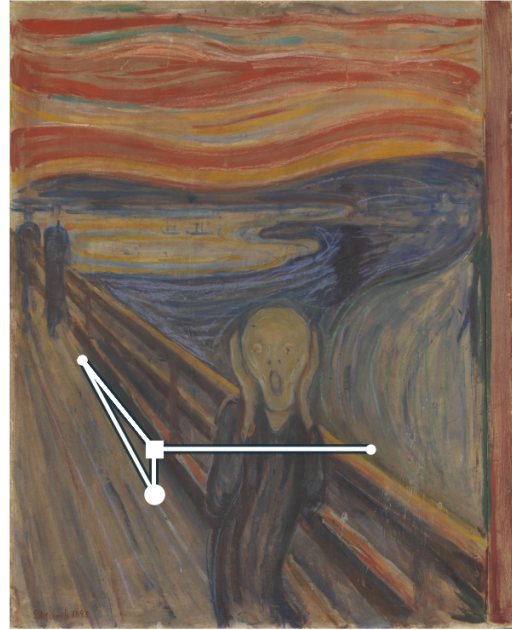
B Stationary: inquiry



C Reversal: restricted



D Reversal: inquiry



Munch (1893); photo Nasjonalmuseet / Børre Høstland (CC BY 4.0).
Fixed background; GNN context channels change, not image pixels.

Key

- Line: sampled scene sequence
- ◻ Square: last displayed site
- Circle: first displayed site

Figure 28: Access and policy are distinct in active scene sampling. Restricted utility and open inquiry see matched potential observations in stationary and reversal contexts. Each panel shows 20 samples from the first paired replicate and declared example window. The background image stays fixed while the supplied regional likelihood changes. These paths illustrate model decisions without uncertainty bands; the prior resets for each episode. A different scanpath does not identify a contemplative mechanism. Artwork: Edvard Munch, *The Scream* (1893), NG.M.00939; photograph Nasjonalmuseet/Børre Høstland; CC BY 4.0; the registration and overlays are authored displays.

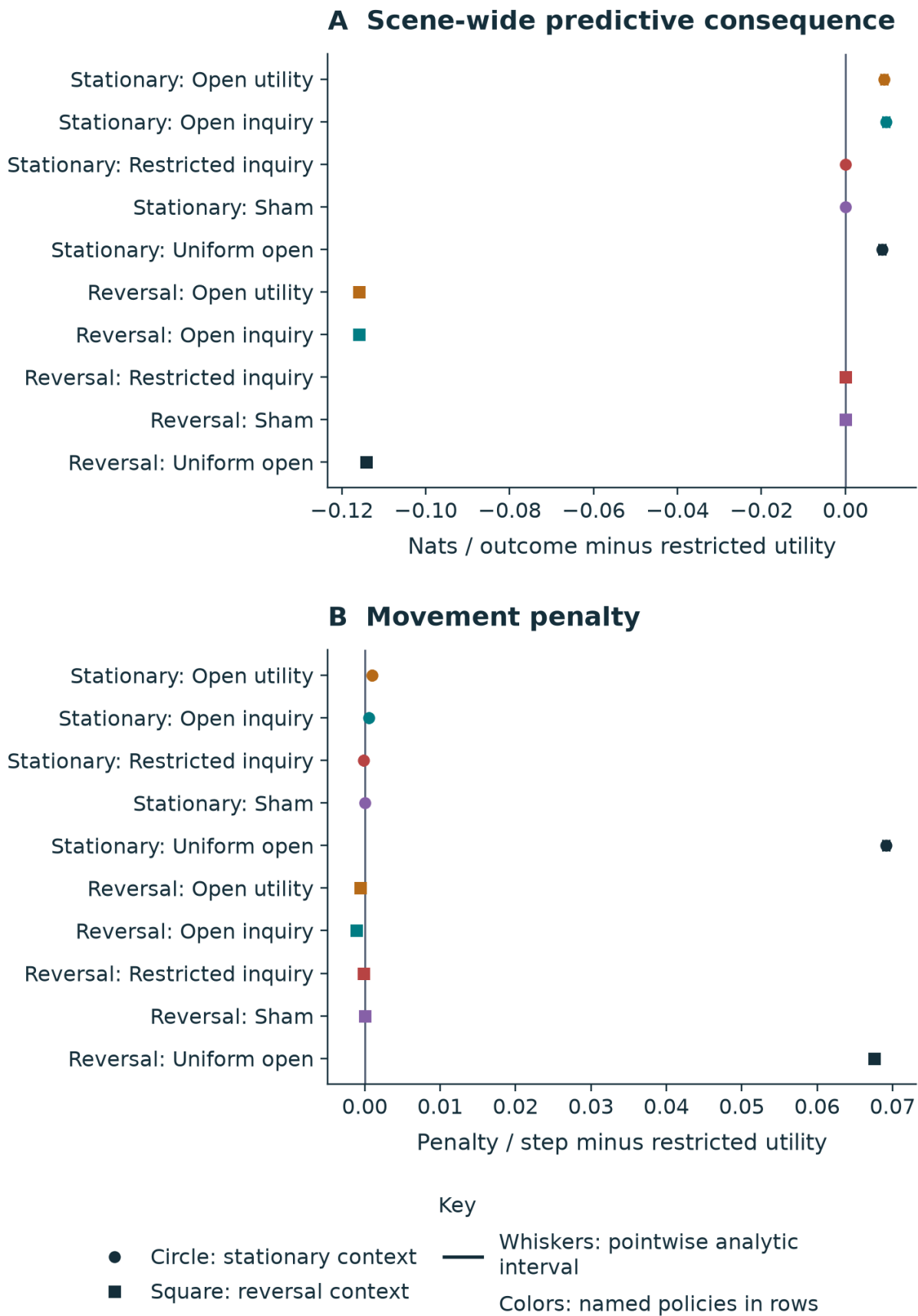


Figure 29: Predictive benefit and movement cost require separate comparisons. Open and restricted policies share a decision budget and complete paired potential outcomes. Estimates compare each policy’s 3-window average with restricted utility. Prediction and movement are separate endpoints; uniform open sampling is a same-budget comparator. Whiskers are pointwise analytic 95% intervals over independent complete replicate averages; proved constant contrasts have zero width. Open access can help reversal prediction while incurring stationary costs. An inquiry-specific claim additionally requires direct comparison with already-open utility and uniform sampling, retained in the complete tables.

The strongest supported distinction is structural access under a contradicted restriction. These settings do not establish an additional scene-risk advantage from the information bonus over open utility or uniform open sampling. In the stationary setting, a correct restriction remains useful. More visits to a positive region, more uncertainty, a longer path and a lower scene-

wide loss are different outcomes. Descriptive coverage, revisits, transition entropy and geometric step lengths are available in the executable gaze companion; none is interpreted as human fixation duration or contemplative attainment.

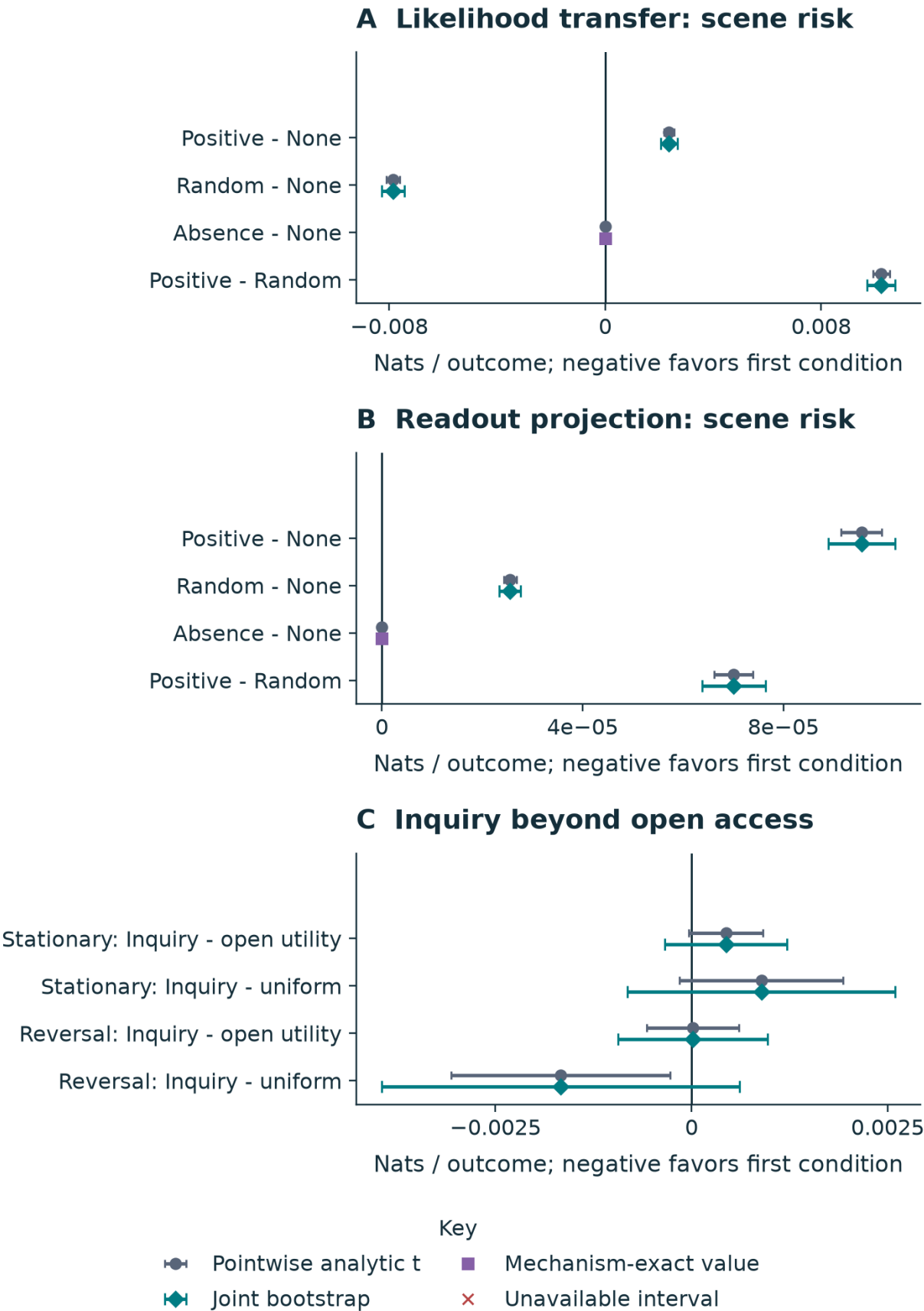


Figure 30: Selected gaze differences retain their full-family uncertainty. The display compares likelihood transfer, readout projection, and inquiry against already-open comparators without pooling those mechanisms. Gray circles show pointwise analytic t bounds; teal diamonds use the single exploratory 58-contrast standardized bootstrap rectangle; purple squares retain mechanism-exact valuation zeros. Each panel uses its native loss scale. The joint region resamples complete paired vectors within fixed independent blocks; its nominal level is 95%. The selected display does not shrink the tested family. Bounds neither establish an extra inquiry-specific benefit nor guarantee finite-sample coverage.

The exploratory joint region has status **complete** across 58 contrasts: 48 estimated, 10 mechanism-exact zeros and 0 unavailable sample constants. Its original-SE-standardized maximum uses 16,383 joint draws and critical value 3.2099. The separate 99% Monte Carlo bracket for that conditional critical quantile is [3.182691, 3.247739]. This is nominal joint uncertainty under bootstrap approximation, not measured coverage for the gaze experiment or a new confirmatory testing family. Mechanism identities supply their own zero contrasts; an accidentally constant sample would remain unavailable.

The joint analysis leaves the point estimates and independent units unchanged. It makes the scope of simultaneous interpretation explicit, rather than reading a collection of pointwise exclusions as family-level evidence. The direct inquiry-versus-open-comparator scene-risk contrasts remain compatible with zero under the larger joint rectangle. Access and information-seeking policy remain different explanations.

Each fixture uses 4096 independent experiments, 64 units per stratum, 6 zero-mean contrasts and 1,023 joint bootstrap draws per experiment. Coverage requires every target and interval; unavailable cases count as noncoverage. Both coverage columns give percentages and exact binomial 95% intervals.

Fixture	Joint region (%)	Pointwise t, complete family (%)	Available (%)
independent gaussian	93.5 [92.7, 94.2]	73.4 [72.0, 74.7]	100.0
correlated gaussian	93.6 [92.8, 94.3]	85.6 [84.5, 86.7]	100.0
heterogeneous gaussian	93.6 [92.8, 94.3]	73.2 [71.8, 74.5]	100.0
centered exponential	88.5 [87.5, 89.5]	69.0 [67.6, 70.5]	100.0
rare bernoulli	1.3 [0.9, 1.7]	1.3 [0.9, 1.7]	1.3

These are operating characteristics of the stated fixtures, not human power or universal coverage. Rare-event sample constants are not licensed as exact population zeros.

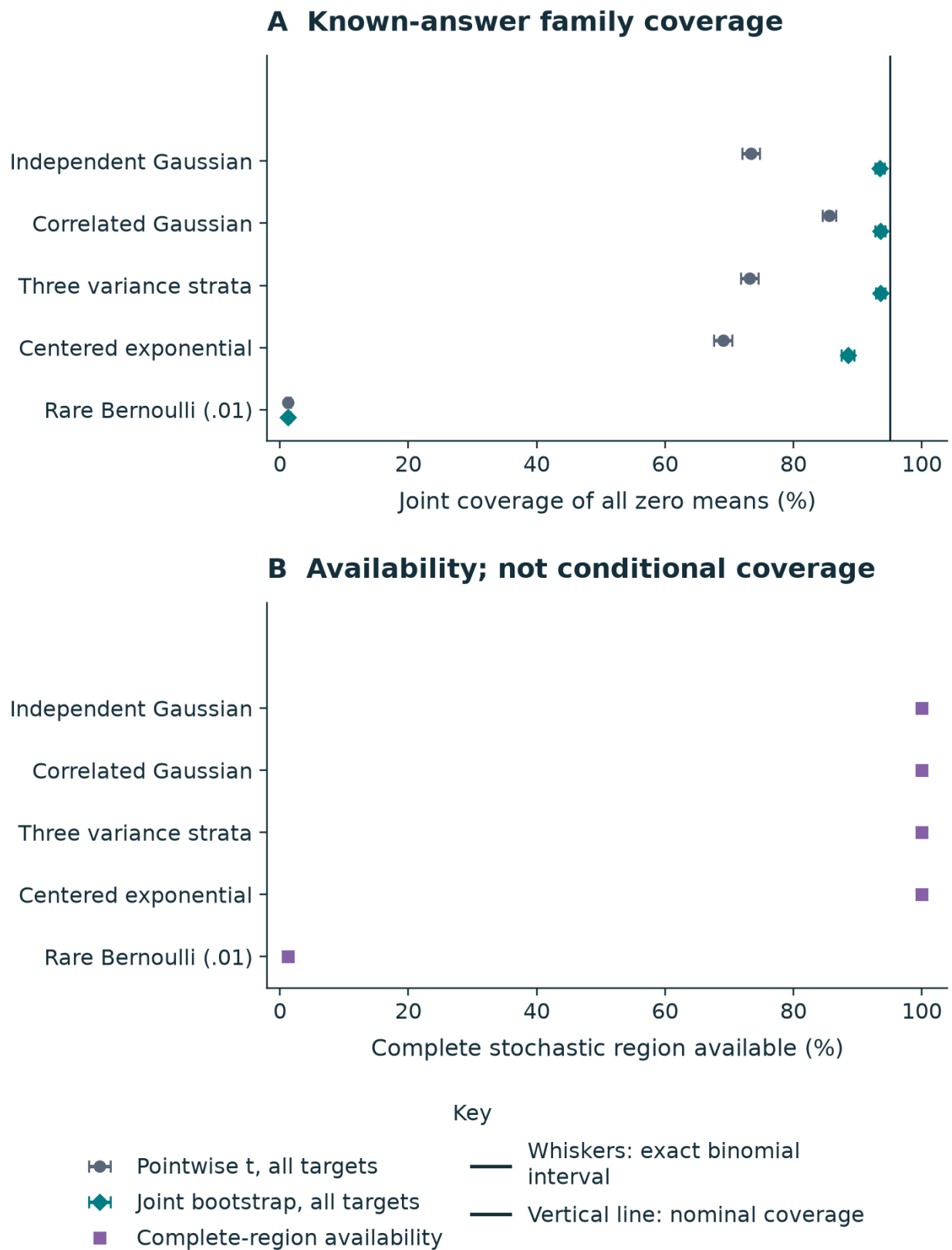


Figure 31: Nominal family confidence needs known-answer checks. Each independent repetition generates a family whose true contrasts are all zero; successful coverage must include every target. Circles apply unadjusted analytic t intervals and diamonds the joint bootstrap rectangle. A separate panel records complete-region availability. Whiskers are exact binomial 95% intervals over independent experiments. Unavailable regions count as noncoverage; the nominal 95% line is a reference, not a tuned acceptance rule. Skew-related undercoverage and unavailable rare-event families remain findings. This is statistical software calibration, not empirical power for gratitude.

These controls distinguish family adjustment from distributional adequacy. Near-nominal performance in a Gaussian fixture cannot repair the substantial undercoverage in the centered-exponential fixture. The rare-event fixture can produce sample constants even though its population mean is known; reporting an unavailable region is more informative than inventing certainty from the sample. No simulated coverage rate licenses a claim about human gratitude or realization.

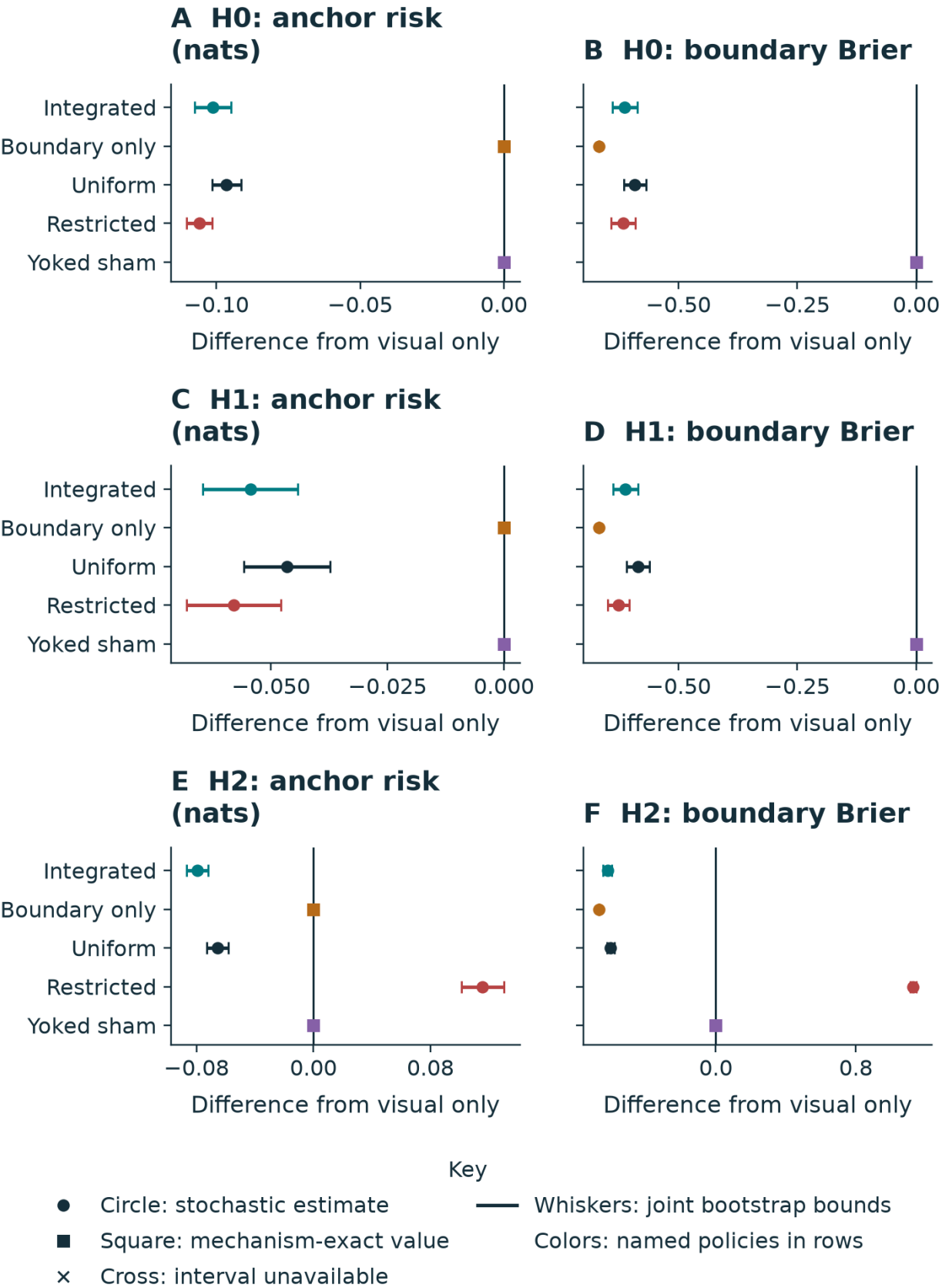


Figure 32: Encoding recovery and reference transfer can diverge. Diagnostic-regime policies are compared with visual-only sampling for each true encoding: H0, H1, H2. Columns separate anchored excess log loss from encoding Brier loss. Negative values favor the named policy; squares mark proved constants and crosses unavailable intervals. Bounds use the global 84-member exploratory bootstrap maximum, including label-only contrasts. Independent paired replicates are resampled within fixed scene strata. The restricted prior excludes H2. Better encoding Brier loss alone does not establish useful reference prediction.

The joint experiment used 9216 independent paired units across 18 fixed regime/scene/encoding cells, with 512 replicates per cell and 24 decisions per episode. The 55296 policy episodes are repeated conditions on those units, not additional independent agents.

Diagnostic-regime anchored excess-log-loss differences (first policy minus visual-only; nats per outcome) were:

Integrated: H0: -0.10122 [-0.10745, -0.09499]; H1: -0.05443 [-0.06459, -0.04426]; H2: -0.07931 [-0.08658, -0.07204].

Boundary only: H0: 0.00000 [0.00000, 0.00000]; H1: 0.00000 [0.00000, 0.00000]; H2: 0.00000 [0.00000, 0.00000].

Uniform: H0: -0.09643 [-0.10136, -0.09150]; H1: -0.04655 [-0.05581, -0.03729]; H2: -0.06559 [-0.07283, -0.05835].

Restricted: H0: -0.10588 [-0.11029, -0.10147]; H1: -0.05799 [-0.06805, -0.04792]; H2: 0.11568 [0.10111, 0.13025].

Yoked sham: H0: 0.00000 [0.00000, 0.00000]; H1: 0.00000 [0.00000, 0.00000]; H2: 0.00000 [0.00000, 0.00000].

Integrated minus Uniform, anchored regret: H0: -0.00479 [-0.01243, 0.00284]; H1: -0.00788 [-0.02119, 0.00544]; H2: -0.01372 [-0.02399, -0.00345].

Integrated minus Uniform, boundary brier: H0: -0.02105 [-0.05349, 0.01139]; H1: -0.02641 [-0.06071, 0.00788]; H2: -0.01757 [-0.04905, 0.01390].

Integrated minus Boundary only, anchored regret: H0: -0.10122 [-0.10745, -0.09499]; H1: -0.05443 [-0.06459, -0.04426]; H2: -0.07931 [-0.08658, -0.07204].

Integrated minus Boundary only, boundary brier: H0: 0.05410 [0.02825, 0.07994]; H1: 0.05537 [0.02922, 0.08151]; H2: 0.04817 [0.02399, 0.07235].

The separate 84-contrast exploratory family retained 54 stochastic estimates, 30 mechanism-exact constants and 0 unavailable sample constants. Bounds use one 95% original-standard-error-standardized empirical-bootstrap maximum across all 6 independent regime/H blocks, with 16383 resamples. They do not establish finite-sample coverage or human construct validity.

The conditional critical value was 3.3196; its 99% order-statistic Monte Carlo bracket was [3.2749, 3.3612]. This bracket concerns bootstrap simulation error, not an additional effect interval.

In the label-only regime, every boundary posterior retained its admitted prior. The restricted-prior Brier differences were H0: -0.166667; H1: -0.166667; H2: +0.833333, with mechanism-exact bounds; other label-only Brier differences were zero. Boundary-only inquiry retained the uniform scene posterior. These identities distinguish information, support and joint use of scene/encoding evidence; posterior entropy is not metacognitive sensitivity.

All 12 independent mathematical/admission controls passed. The sham is a nonautonomous matched-query updater that discards action identity and ignores adaptive-schedule information. The anchored H=0 score is a declared reference-frame intervention, not ordinary future prediction under the original scene process.

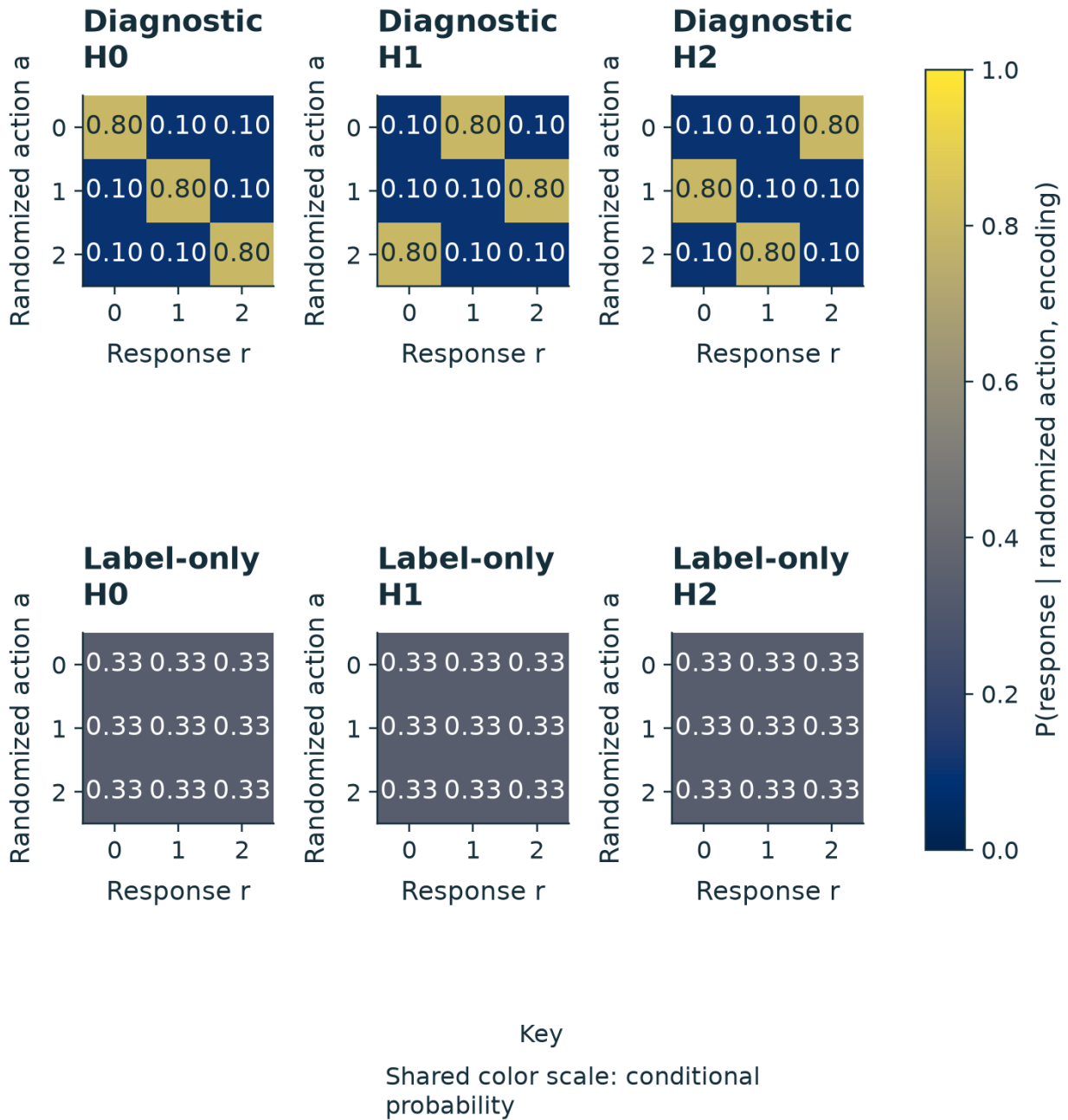
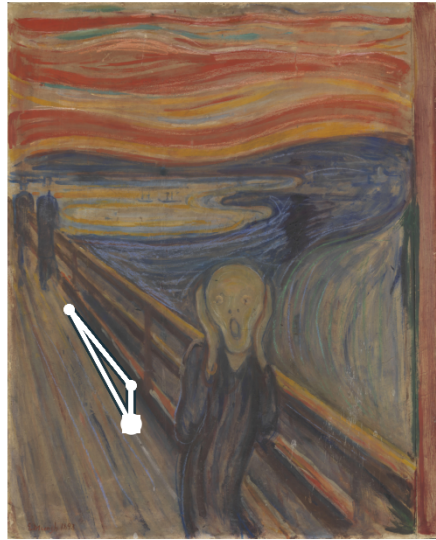
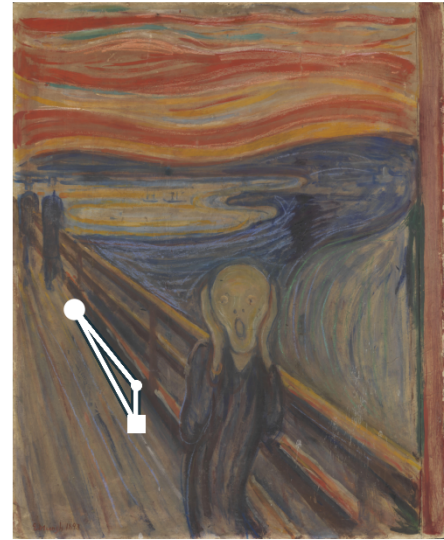


Figure 33: The action-response relation can identify what its marginals cannot. The matrices display declared response probabilities conditional on the randomized action, using the middle probe at reliability 0.8. The diagnostic columns differ by encoding; the label-only channel has no encoding dependence. Shared probability scales and numeric cell labels permit direct comparison. These are exact channel declarations with no sampling intervals; both separate marginals remain uniform. The information lies in the paired relation. These finite matrices are not observed eye movements or measured biological boundary channels.

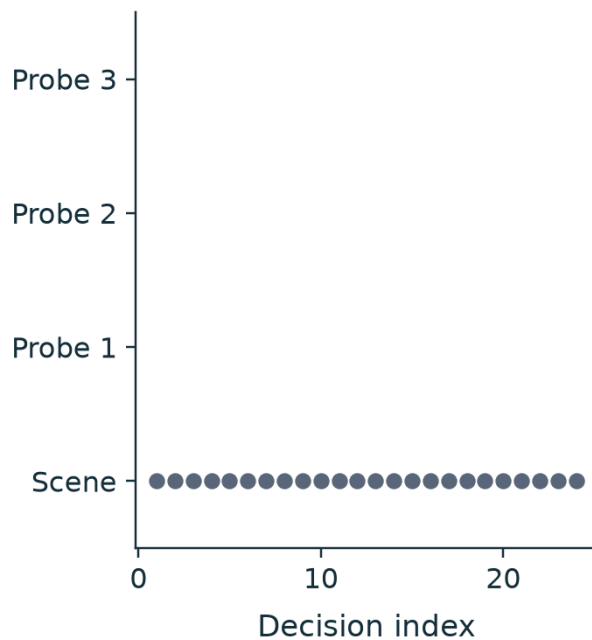
A Visual



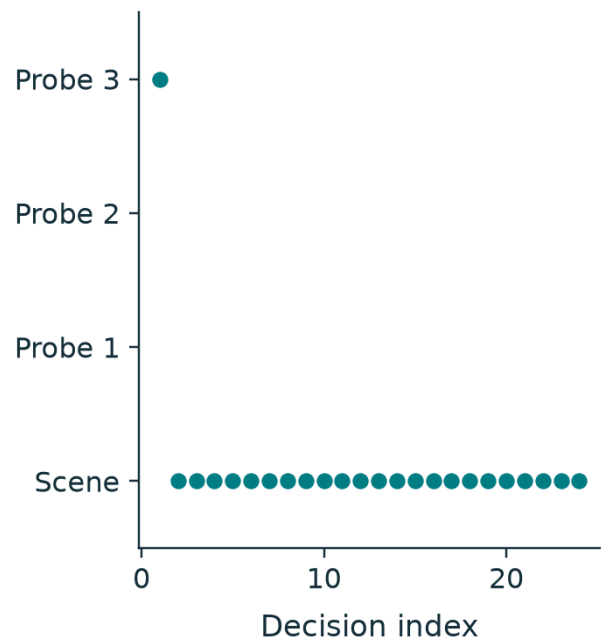
B Integrated



C Sampling schedule



D Sampling schedule



Munch (1893); photo Nasjonalmuseet / Børre Høstland (CC BY 4.0).
Model scene samples above; nonspatial boundary questions below.

Key

- Line: sampled scene sequence
- Circle: first displayed site
- ◻ Square: last displayed site
- Visual-only decisions
- Integrated decisions

Figure 34: Spatial gaze and nonspatial questions share one budget. The preselected example uses the first scene, H2, and the first paired replicate. Upper panels show the first 20 visual samples; lower panels retain all 24 decisions, including contingency questions. Straight lines connect displayed sites across intervening questions. This is an illustrative trajectory, selected by its input indices rather than its result; it carries no inferential interval. A question is not a spatial fixation. Artwork: Edvard Munch, *The Scream* (1893), NG.M.00939; photograph Nasjonalmuseet/Børre Høstland; CC BY 4.0; registration and overlays are authored displays.

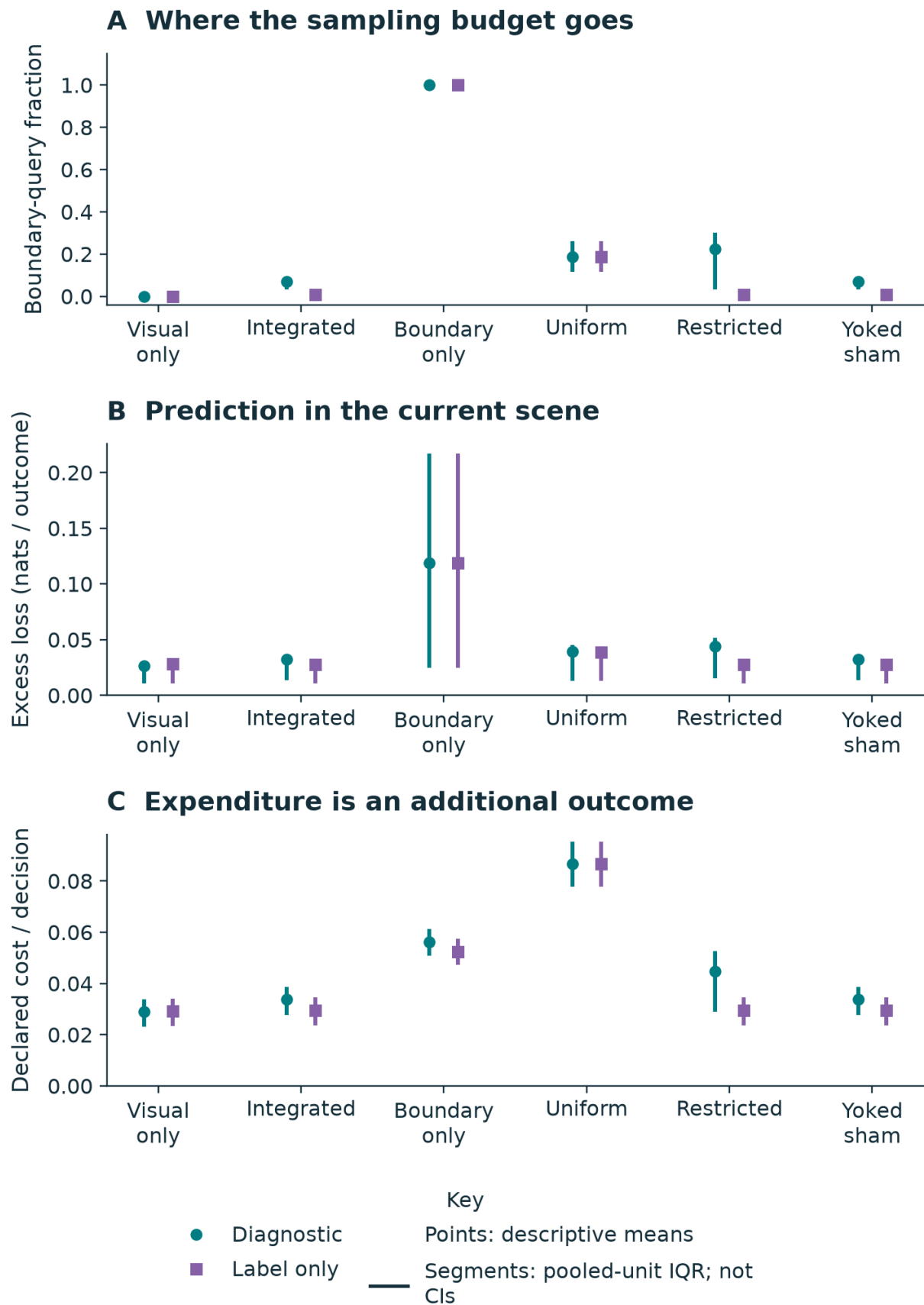


Figure 35: Information, prediction and expenditure are separate outcomes. Policies allocate a common decision budget between scene samples and encoding questions. Points are equal-fixed-stratum means of question allocation, current-scene excess predictive loss and declared cost; vertical segments show pooled whole-episode interquartile ranges. Each panel retains its own units. Segments describe whole-unit spread across the fixed scene/encoding mixture, including between-cell variation; they are not confidence intervals. Repeated decisions are not independent samples. An informative question can reduce scene sampling even when reference transfer improves. Expenditure and encoding knowledge are not interchangeable with predictive success or realization.

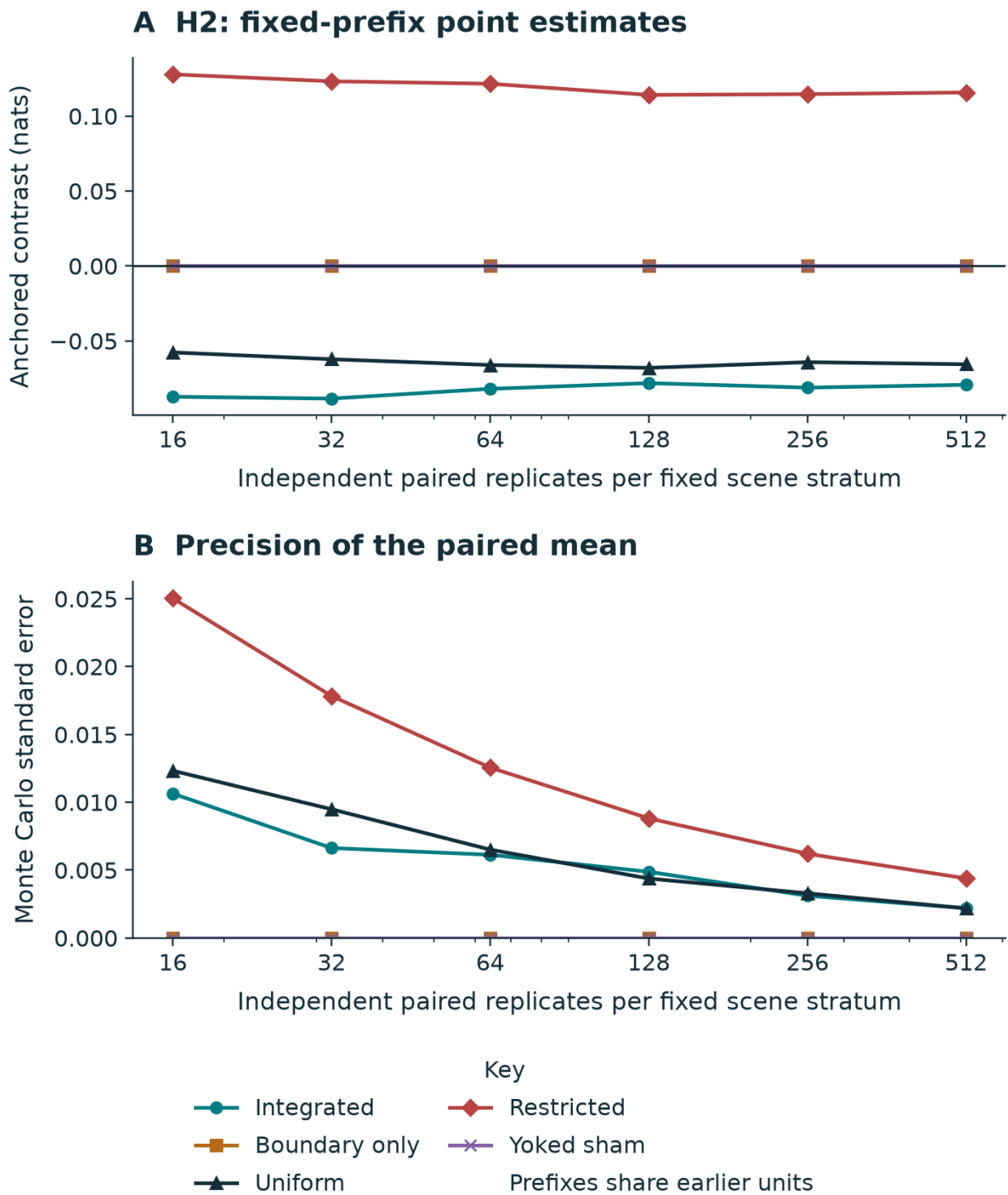


Figure 36: More independent agents reduce Monte Carlo error, not structural ambiguity. The preselected H2 case is evaluated at the declared nested replicate prefixes. Points show paired anchored-risk contrasts averaged over 3 fixed scene strata, alongside their Monte Carlo standard errors. Horizontal axes are logarithmic. Prefixes share earlier units and are correlated; repeated decisions do not enlarge the independent N. Point estimates need not change monotonically. The full convergence record retains both endpoints and all declared contexts. Enlarging N cannot identify observation-equivalent encodings.

What the imported boundary analyses establish

The pinned QRF audit assigned the same complete 64-outcome distribution to 3 label profiles; this equality is constructed, not fitted from observations. A normalized altered distribution was detected independently. In a separate six-variable Gaussian fixture, precision-support F1 was 1.00, partial-correlation best F1 1.00, and the best F1 across every possible marginal-correlation threshold interval was 0.80. Inverting the known full covariance recovered precision to maximum error 4.44e-16. Thus the threshold limitation does not imply general non-identifiability from joint observations. The upstream perturbation audit retained 200/200 positive-definite draws; its recovery fractions are conditional on that admission. These fixtures and the new sampled structural-inquiry experiment test different mathematical claims. All 8 upstream integration and independent mathematical controls passed; actual module, declaration and serialized-API digests are retained.

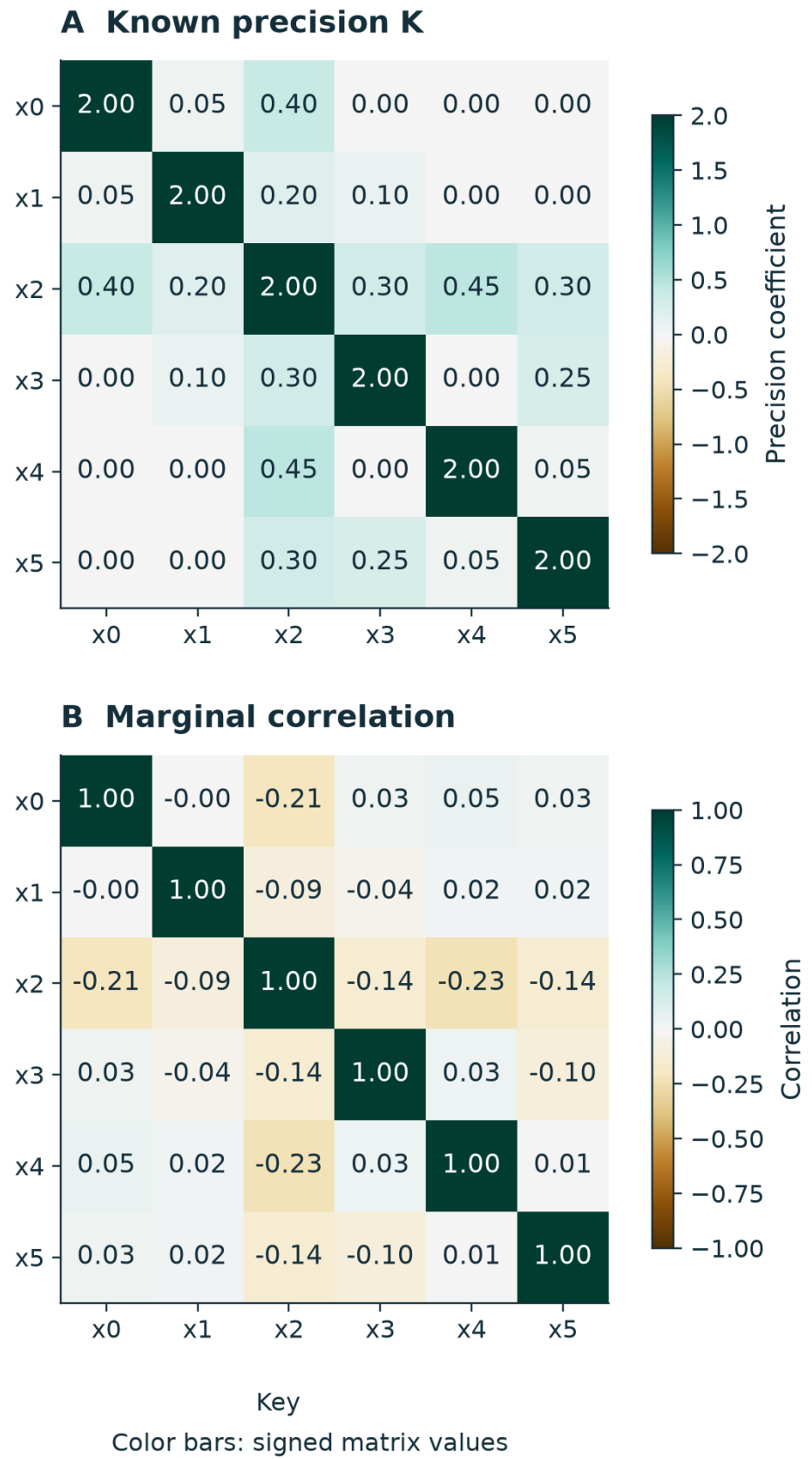


Figure 37: Marginal association and conditional independence answer different questions. This separate pinned Realizing Emptiness Gaussian fixture compares the known precision matrix with full marginal correlations. The matrices describe the same Gaussian law. Exhaustive marginal-threshold intervals fail exact blanket recovery, while inversion of the full covariance recovers precision. Entries are deterministic finite mathematical values, not estimated biological correlations or confidence intervals. Gaussian variables and QRF bit labels name different objects. This comparator is not another participant or observation in the scene/encoding experiment.

S5. Mechanistic interpretation and competing explanations

Gratitude as an orientation toward continued inquiry

The stronger proposal is that appreciation can support the practice of examining what seems compulsory in one's way of construing experience. Gratitude for an opportunity to inquire need not demand a favorable conclusion. A practice can reveal that a distinction is contingent while preserving its everyday usefulness; it can also reveal that a particular alternative does not help. This orientation differs from collecting agreeable evidence. It values the possibility of revising the terms in which evidence is understood, while continuing to answer to consequences.

The finite extension makes this proposal more demanding than a verbal endorsement. Preference must affect action; action must actually provide access; accessible comparison must consume evidence; and a selected working representation must face independent observations. The sham and no-access nulls show where the chain can break. Their role is not to devalue aspiration, but to prevent its presence from being counted as an enacted mechanism. The no-preference condition also shows why instrumental inquiry and appreciation for its possibility need separate identification. Here the latter is declared, not inferred from simulated behavior or validated as human gratitude.

Useful self/world models can persist after their assumed necessity is questioned (Sandved-Smith et al. 2026). Our model reflects that point by allowing a restricted working family to be retained or reused when its evidence is better. Returning to useful distinctions need not reinstate the rule that alternatives are inadmissible: future inquiry remains available. This is the substantive difference between a representation that is employed and a restriction that cannot be examined. The finite family switch supplies a constrained analogy; it does not exhaust the source's physical or contemplative claims.

Practice also has a temporal structure. A moment of insight would not remove changing conditions or the need to examine future constructions. The repeated reversal design therefore asks about continued revisability rather than a one-way progression through named attainments. Its costs matter as well. More inquiry can be predictively useful while consuming resources; if an adequate model already works, the charge can remain without a predictive gain. Our separate cost column makes that possibility visible instead of declaring every relaxation beneficial.

Examinability, instrumental information and welcome answers

The stipulated opportunity value is not the only reason to seek information. In Kobayashi and Hsu's actionable-lottery task, models separated instrumental value, fixed or entropy bonuses and anticipation of future outcomes; the combined anticipatory account improved behavioral prediction within that task (Kobayashi and Hsu 2019). The authors leave room for different motives elsewhere. Their comparison does not identify our information coefficient, and the article's neural results do not supply a neural interpretation of this model.

Information can also be sought without changing a payoff. In van Lieshout and colleagues' second experiment, participants could incur an additional delay to learn an outcome whose payout was already fixed (Lieshout et al. 2018). Their uncertainty predictor multiplied entropy by the difference in possible payoffs; it was not the pure candidate information gain used here. Willingness to incur that delay is a useful behavioral measure, but does not itself establish revisability or appreciation.

Together with the patience result (DeSteno et al. 2014), these studies sharpen the proposed mechanism. Inquiry could increase because information is useful, because good news is anticipated, because delay has become less costly, or because any requested activity is followed more readily. Appreciation of examinability would require discriminating effects beyond these alternatives. In particular, valuing examination should remain compatible with learning an unwelcome result and retaining a useful representation. The current finite models stipulate an orientation and test its enabling consequences; they do not estimate which motive a person has.

Process evaluation adds a complementary distinction between what is delivered, how it produces change and the context in which it operates. Moore and colleagues distinguish fidelity to intended delivery, the amount delivered and the audience reached, while noting disagreement over the boundaries of these implementation components (Moore et al. 2015, MRC process evaluation framework). A human study should therefore record the activity performed, whether the intended opportunity to examine alternatives was available, how participants used it, and the subsequent independent outcome. Context can change effects even when implementation is similar. Our separate activity and access rates illustrate the need for distinct records; they are not validated measures of human fidelity, dose or reach. The guidance motivates a measurement strategy, rather than establishing that this proposed practice works or that an informative alternative constitutes contemplative realization.

Revisability, useful precision and philosophical sufficiency

Maintaining the distinction between a model and its referent is useful even without a philosophical identification. The same prediction can survive consistent relabeling of hidden states; a semantic label is not an additional measurement. Our permutation control checks that prerequisite. The practice extension goes further by making an alternative predictor accessible, but a finite KL divergence still does not measure emptiness. Lower precision need not mean greater wisdom: it can simply mean poorer discrimination. This is why the implementation compares explicit predictive models, refits their posteriors, retains useful structure and scores independent observations rather than rewarding entropy or a label named nondual.

The inspected historical texts encourage care about both existence/non-existence and the residual presence acknowledged when something is absent (Thanissaro Bhikkhu 1997b, 1997a). The catalog distinguishes these textual readings from the more limited access to the philosophical books (Garfield 1995, 2001). They do not provide a shortcut from numerical uncertainty to contemplative realization. Likewise, a computational affect model can offer a formal account while leaving empirical identification

open (Hesp et al. 2021). The present extension is most useful when these gaps remain visible: philosophical reasoning, human experience, and algorithmic invariants answer different questions.

Laukkonen and Slagter offer a different predictive-processing account, proposing reduced temporal and conceptual depth and discussing a radical nondual endpoint while acknowledging differences among traditions (Laukkonen and Slagter 2021). Predictive switching in our finite menu is neither that proposed endpoint nor an agreed definition of emptiness. A system could improve reversal prediction yet reify whichever description currently wins. Revisability therefore supplies a candidate enabling condition, not a sufficient criterion of realizing emptiness. Our positive target is the examinability and conditional use of representations; the stronger philosophical claim still needs argument beyond a score.

Rehearsal and the choice of a benefit criterion

The central methodological issue is not whether positive-event rehearsal can move a probability. Such movement depends on the specified selection rule, likelihood, concentration, and update. The harder question is what that movement accomplishes under an independently defined objective. In this study, the objective is deliberately plural: the optimism readout, latent-state error, observation prediction, and uncertainty are retained together. A model that feels conceptually more hopeful in its labels may become less accurate, while a calibrated model may continue to represent an unfavorable world. Neither result should be concealed by renaming the preferred endpoint well-being.

The source preprint itself supplies an important boundary. It excludes trauma and unstable environments from its initial simulations and warns that optimism can be maladaptive when favorable opportunities are unavailable (Fisher et al. 2025). The corrected stationary worlds do not remove that limitation. They improve causal comparison within a finite experiment by holding exposures fixed, but do not simulate the practical consequences of engagement, withdrawal, protection, grief, or social support. There is no action-outcome welfare measure demonstrating that optimistic engagement benefits an agent in a dangerous environment.

Metacognitive gaze without reifying its target

Task-directed human gaze can change which scene regions are sampled (Castelhano et al. 2009). In the inspected counterbalanced search/memorization experiment, critical search scenes omitted the target; task effects on region selection did not imply the same effects for every aggregate or early-viewing measure. Those observations motivate treating a gaze task as part of a sampling process. They do not license reading gratitude, self-understanding or meta-awareness from an altered scanpath.

The new model extends the question from *where to look* to *which relation would make an encoding examinable*. Its contribution is the joint use of visual evidence and contingent questions, together with exact cases where either operation alone is insufficient. A question can concern the conditions under which a distinction works without discovering a metaphysically privileged boundary. Label-only equality constrains the claim in the opposite direction: no amount of sampling reveals a difference that the observation model does not contain. These are complementary limits, not competing descriptions of one dataset.

The proposal about appreciation remains a further link. A person might value the conditions that permit careful inspection, notice a framing assumption, and choose an informative question. Our integrated agent supplies an explicit finite mechanism for the last operation. It does not infer that motive from scene behavior or simulate a learned appreciation-to-inquiry transition. Nelson and Narens' monitoring/control distinction and Fleming and Lau's measurement distinctions therefore constrain the empirical program (Nelson and Narens 1990; Fleming and Lau 2014). Posterior certainty can concern an incomplete candidate menu, and lower entropy can accompany a wrong model. Neither maximum doubt nor maximum certainty is our criterion for practicing realization of emptiness.

Appreciation without invented evidence

The absence-appreciation condition makes a limited possibility explicit: a value can be assigned relative to a reference without changing the likelihood or posterior used for state inference. This separation prevents a particular computational error, namely treating a preferred interpretation as a newly observed favorable event. It also reveals the dependence of appreciation on the reference. Changing the assumed hazard or cost changes the assigned value even when the experienced observations remain identical. That dependence is a feature to expose, not an inconvenience to hide.

Human counterfactual reflection may have effects absent from this implementation. Imagining alternatives could change memory, attention, affect, beliefs about future hazards, or social judgments. Gratitude toward a benefactor may additionally involve intentions and perceived sacrifice (Yu et al. 2018). It may evoke indebtedness (He et al. 2022). Consequently, the exact belief null here should be understood as a clean software separation and a candidate empirical contrast, not a prediction that real appreciation is psychologically isolated from inference.

Interpreting observed absence also requires a detection model. An interval without an alarm can favor safety when alarms reliably detect harm, convey little when the detector is unreliable, or be misleading when reporting is selectively suppressed. Our missingness sentinel assumes no update; real missingness may instead be informative. Seaman and colleagues distinguish realized and everywhere missing-at-random conditions and explain that ignorability depends on the inference being made (Seaman et al. 2013, sec. 2 and 5). Their sufficient Bayesian conditions also concern the relation between priors for data and missingness parameters. A missing symbol alone establishes none of these conditions, and missing at random is not synonymous with zero information. The finite audit clarifies our choice but does not solve inference under arbitrary observation failure. A model of gratitude for an absence should therefore state what could have been detected, what was recorded, and how counterfactual value was assigned.

S6. What these models do and do not identify

The source reconstruction covers selected numerical cells under a documented adapter. It does not reproduce every interface action, excluded plotting helper, software environment of the original session, or later version of the notebooks. Detailed paper analysis is limited to preprint version one; the journal abstract establishes publication identity and broad scope only. The scholarship catalog records whether each supporting source was inspected in full, at abstract level, or only through bibliographic metadata. Bibliographic presence alone is not evidence of an effect.

The corrected learning model is intentionally small. It uses discrete valence categories, a fixed world grid, a stationary process within each stratum, and a restricted observation model. It does not represent interpersonal intentions, memory decay independent of the chosen update schedule, self-report demand, embodied behavior, or an agent’s ability to change its environment. Replay and precision are imposed computational operations. The optimism readout is partly determined by its construction. Learned gaze transfer now has an explicit embedding, but its sign depends on that mapping and its predictive consequences can be adverse. The active scene model assumes known candidate channels and a hand-authored region partition; it neither infers painting semantics from pixels nor fits human eye movements. A richer model is not automatically better; its additional parameters would require identifiability and predictive tests.

The intersection additionally assumes known query channels, independent uniform scene/encoding priors, an authored cyclic map and an available reference anchor. It learns within a supplied menu; it does not learn how to construct a new representation or a metacognitive architecture. Query costs are declared utilities, not measured effort. The matched-query sham ignores the demonstrator’s adaptive-schedule likelihood. Its numerical identities therefore concern that updater, not complete absence of available information. Near-deterministic stochastic scores can remain poorly characterized by finite-sample bootstrap approximations even after numerical centering is corrected.

The practice model has additional, different limits. Its candidate repertoire, structural restriction, neutral channel schedules, opportunity value, access probability and cost are authored design choices. Restricted and open priors are compared by exact refitting within an episodic window, not by a fitted model of human metacognition. The reversal context gives the omitted predictor a known advantage and cannot establish a general advantage of relaxation. Evidence gain is scaled into utility by a declared coefficient. There is no learned gratitude state, no measurement of first-person realization and no validated translation from the sector labels to experience. The full appreciation–practice–realization proposal is therefore a conceptual research program with one executable enabling mechanism, not a completed reduction.

S7. Human research: manipulations, measurements and recovery checks

Distinguishing absence, missingness and valuation

A first study would separate positive-event recall, dose-matched random-event recall, appreciation of an observed uneventful interval, and counterfactual reflection about avoided harm. It would measure gratitude, relief, indebtedness, perceived safety, and well-being separately while matching expectancy and task engagement as closely as feasible. The absence and missingness manipulation would include explicit information about detector reliability. If participants treat missing reports as favorable evidence despite knowing the detector failed, the resulting behavior would challenge the simplest missingness convention rather than validate it.

Mechanism testing would require repeated behavioral observations, parameter recovery, alternative models, and held-out prediction. A precision-based explanation should outperform simpler changes in response bias or outcome coding and predict transfer where transfer is claimed. An attention study could compare pre/post scene behavior while controlling preferences and exposure; failure of transfer would directly constrain the proposed bridge. An optimism shift accompanied by worse calibration or adverse outcomes would count against a general benefit claim. Statistical planning would use empirical variability and clinically or theoretically meaningful outcomes, not the number of synthetic agents reported here.

Testing appreciation, access and appropriate revision

A prospective inquiry study could randomize appreciation of enabling conditions, pleasant-event reflection and neutral factual reflection, crossed with actual availability of an alternative explanation. Stable and changed contexts would test useful retention and warranted revision. Recognition probes would measure whether an available alternative was noticed; later independent trials would evaluate predictions. These are proposed studies, not experiments or preregistrations completed here.

For example, let μ_{zx} be expected later predictive loss under orientation z and assigned access x . A candidate primary estimand compares enabling-condition appreciation E with pleasant-event reflection P :

$$\theta_{\text{emp}} = (\mu_{E1} - \mu_{P1}) - (\mu_{E0} - \mu_{P0}). \quad (29)$$

A negative interaction would favor an access-dependent predictive consequence, provided the randomized design, missing-data plan and intended population justify that interpretation. It would not establish gratitude as the mediator or realization as the outcome. This estimand would need prospective selection, empirically grounded sample-size planning, a meaningful effect scale and separate process measures before data collection.

Rival explanation	Discriminating manipulation or measure	Observation that constrains the proposal
Pleasant affect or expectancy	Pleasant-event comparator and measured expectations	Only favorable-language ratings change
Instrumental information value	Useful and noninstrumental information at comparable uncertainty	Choices follow expected task benefit alone
Anticipated good news	Welcome and unwelcome information with usefulness controlled	Inquiry increases only for desired answers
Patience rather than inquiry value	Independently varied delay, effort and monetary costs	Waiting changes without generalizing to other inquiry costs
Availability without awareness	Alternative-recognition and comprehension probes	Alternatives are exposed but neither noticed nor used
Compliance or indiscriminate doubt	Stable accurate and contradicted representations	Accurate beliefs weaken regardless of evidence
Temporary task switching	Later independent prediction and renewed access checks	Immediate switching does not sustain appropriate revisability

An appreciation increase without access or appropriate revision would interrupt the proposed chain. Indiscriminate weakening of accurate beliefs would fail its retention requirement. Separate social-attribution, curiosity, patience and demand measures would help compare mechanisms, but adding questionnaires alone would not solve identification. Philosophical interpretation and first-person reports require their own methods and could disagree with predictive scores.

Before interpreting fitted human parameters, parameter recovery must cover the relevant parameter range, competing models must be tested through model-recovery confusion matrices, and simulated behavior must reproduce patterns not used for fitting (Wilson and Collins 2019). The initial policy’s value–cost confounding illustrates why exact software replay cannot substitute for identification. The exhaustive count oracle characterizes this finite menu; a winning menu member may still be an inadequate explanation. Social attribution and matched positive affect also belong in the inquiry comparison, given the boundaries of costly interpersonal gratitude (Bartlett and DeSteno 2006). Access and control need separate process measures rather than being inferred from the final choice (Sandved-Smith et al. 2021).

Finally, adverse and unstable environments should be central boundary conditions in later work rather than extrapolations from benign worlds. Any human study would need independent ethical review, informed consent, appropriate support, and a design that does not encourage false reassurance or avoidance of necessary care. These requirements define what new evidence would be needed. They are not recommendations to deploy the simulated operations as an intervention.

Coupled gaze and boundary inquiry in a human study

Experiential reports would require a specified collection and analysis method. The inspected visual-task study used training and participant-derived experiential categories before grouping trials for behavioral and EEG analyses (Lutz et al. 2002). We take this as a methodological example: reports can add constraints to a study, while their interpretation remains distinct from first-order performance, gratitude and philosophical sufficiency.

A direct empirical extension would combine a scene-search task with an explicit action-contingency task. Participants could choose scene observations or costly questions about a learned reference mapping; a separate label-only condition would lack any distinguishing observation law. Randomized action assignments, event-pair identities and access to both components would be retained. First-order scene/reference predictions, trialwise confidence, encoding judgments, dwell times and spatial sampling would be measured separately. Confidence bias, sensitivity and efficiency would require their own estimands and models (Fleming and Lau 2014).

Equal decision counts would not establish equal effort: query duration, motor demands and the number of observable components require measurement and appropriate comparators. A replayed-query comparator would need to account for information conveyed by the demonstrator’s adaptive schedule or explicitly restrict the available record. A reference-frame transfer assessment would need to instantiate its anchor experimentally, rather than reinterpret an ordinary future trial after the fact. Proper recovery studies would test omitted encodings, misspecified channels and curiosity-only alternatives before assigning a fitted parameter a metacognitive meaning.

Crossing that task with enabling-condition appreciation, pleasant reflection and neutral reflection could test the proposed orientation toward examinability. The focal evidence would be appropriate allocation and later prediction, including useful retention, rather than a generalized rise in doubt or positive reports. An informative action–response relation would identify a task encoding, not the source preprint’s physical or contemplative thesis. Human sample-size planning would require empirical variability, participant-level dependence and prespecified effect targets; the larger synthetic cohorts cannot supply those quantities.

S8. Source custody, operating diagnostics and complete interval tables

Digest-bound source reconstruction

The primary source is the archived version-one preprint and two separately preserved notebooks: the gratitude/belief-updating notebook linked by the user and the modified Yarbus notebook identified in the preprint’s code-availability section. Whole-file digests bind execution to those artifacts. The reconstruction executes selected numerical definitions and cells with isolated

seeded sampling. It excludes package installation, plotting, downloads, interactive exports, and stored notebook outputs. It is a bounded headless reconstruction, not a complete replay of the original Colab interface.

The gratitude adapter permits reduced dimensions only through a manifest of integer-literal substitutions in the pinned source. The original dimensions remain available. Baseline arrays are copied before subsequent mutation so that the reconstruction can expose aliasing rather than silently correcting it. The Yabus adapter extracts the original agent class and preserves its action-first transition convention. Numerical findings are reported separately from static observations about excluded plotting helpers. This approach makes the source executable while retaining the behavior that the corrected study later changes.

An independent unequal-column-mass example returned posterior well probability 0.9000 from raw counts versus 0.5000 after column normalization. The seeded source adapter measured a maximum aliased snapshot mutation of 75.6530. The original plot binning excluded 1 endpoint-grid agent(s). In the attention reconstruction (seed 42, 100 steps), 33 sampled observations had zero likelihood under the separately declared process. The independent preference-plus-ambiguity oracle matched the source energy to maximum absolute error 0.

The source audit distinguishes executable facts from judgments about how a model should be interpreted. In the gratitude notebook, accumulated likelihood entries are used as weights without column normalization. This makes accumulated mass relevant in addition to relative observation probabilities. Baseline and intervention snapshots are saved at different points relative to daily learning, and a view of a saved baseline can be mutated during intervention. The phases also change environmental observation and transition processes, so their before/after difference does not isolate the rehearsal operation. Updating-task priors are redrawn rather than paired. The implementation’s clamp includes selected positive or neutral observations, whereas the preprint’s description emphasizes positive-event rehearsal.

These findings motivate the corrected design; they do not demonstrate that every possible count-weight interpretation is invalid. Nor should a difference in tensor convention be called a defect without checking its use. The Yabus source indexes its transition tensor action-first, and that convention is internally valid. Its observation sampling uses the agent likelihood despite a separately declared environmental likelihood. Its implemented policy score contains a conditional-entropy term, which must not be relabeled as an independently calculated information-gain quantity. The source’s plotted paths remain illustrative discrete visits with plotting transformations, not eye-tracking measurements.

Exogenous attention: implementation and interpretation

The attention extension uses a 13-location scene with positive, negative, and neutral categories. The agent consumes declared likelihood, transition, preference, and initial-state tensors, with the canonical transition axes ordered next state, previous state, action. Native GNN execution qualifies a matched agent/process model. A separate counterpart invokes the active-inference agent while holding the environmental observation likelihood fixed across exogenous precision profiles. Preferences and initial beliefs are perturbed as discriminating controls. Numerical receipts retain observations, actions, beliefs, and policy probabilities.

Category summaries aggregate mean policy probability over the designated locations. They describe model policy mass, not measured gaze, subjective preference, or a learned effect of gratitude practice. The selective-attention literature motivates this precision manipulation (Mirza et al. 2019). In an adjacent affect-congruent attention experiment, modeled attention and gaze were not identical measures, and gaze findings depended on probe type and normalization (Bennett et al. 2023). That study did not manipulate gratitude. This exogenous comparator has no learning bridge; the separate primary extension in the transfer model consumes actual frozen learned channels. The distinction between native matched-model qualification and a fixed-process comparison is essential to interpreting the result.

The positive precision model assigned mean policy probability 0.9870 to positive, 0.0049 to neutral, and 0.0081 to negative locations over 8 steps.

The negative precision model assigned mean policy probability 0.6218 to positive, 0.3731 to neutral, and 0.0051 to negative locations over 8 steps.

The attention comparison asks whether changing declared precision profiles changes policy probabilities under the specified agent and process. It does not estimate a gratitude-to-gaze effect. Native execution and the fixed-process counterpart answer different qualification questions, so numerical validity of the former cannot substitute for the latter. Preference and initial-prior controls help establish that the implementation consumes its declared inputs rather than merely returning a plausible-looking path.

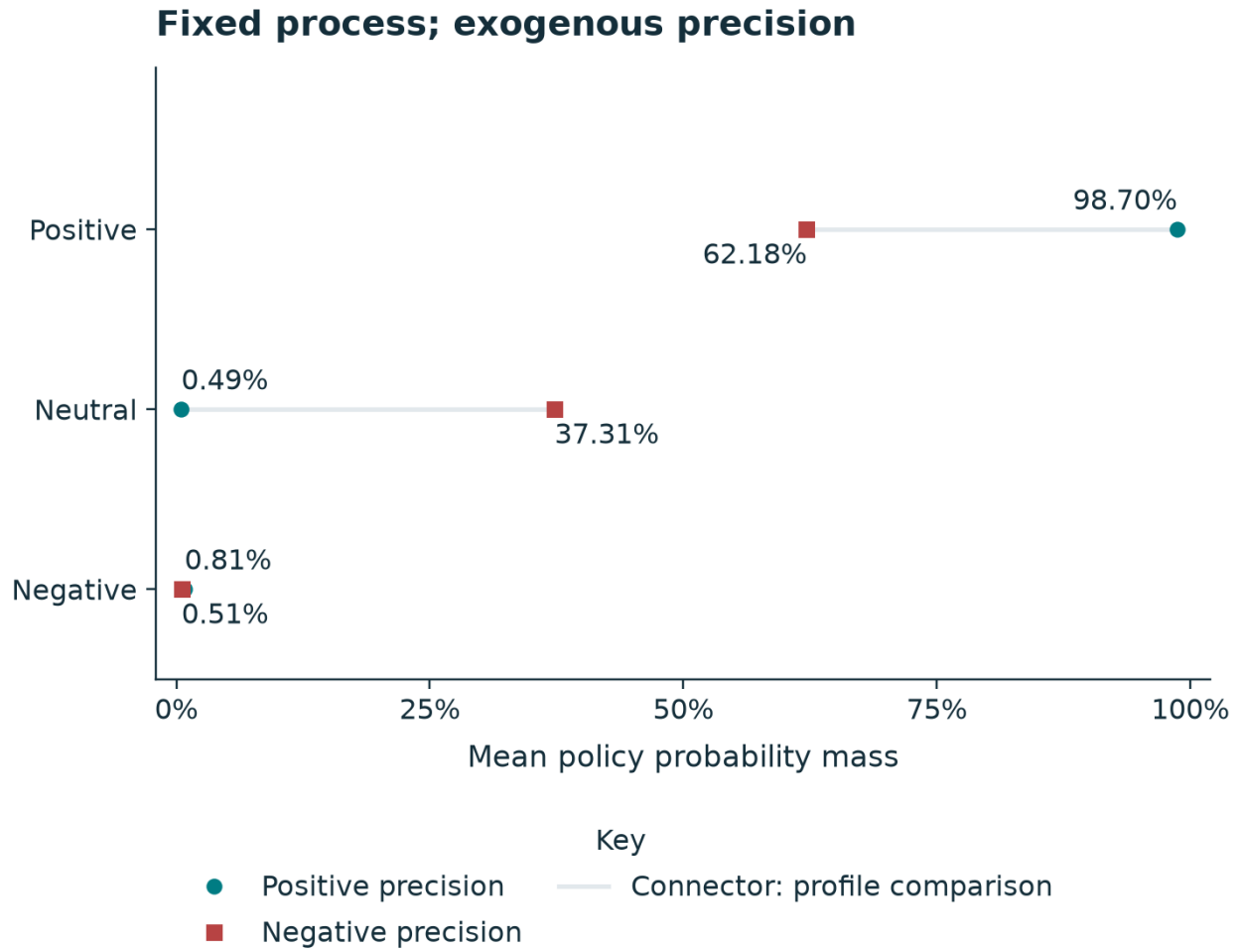


Figure 38: Precision profiles alter policy mass in a fixed-process comparator. Positive and negative attention profiles are exogenously assigned rather than learned from gratitude rehearsal. Circles and squares identify profiles. Values aggregate policy probability over the declared positive, negative and neutral sites; all masses, including small negative-site values, are labeled. Gray connectors compare profiles and are not uncertainty intervals. Native matched-model qualification is separate. These policy probabilities are neither measured gaze nor evidence of learned transfer.

Statistical operating checks and model sensitivity

Each fixture used 4096 independent experiments, with 64 paired units per stratum, 1023 bootstrap draws and 1023 sign-flip draws. Methods share each generated sample.

Fixture	Percentile cover	Analytic t cover	Bootstrap t cover	t available
Gaussian null	3864/4096	3898/4096	3887/4096	4096/4096
Gaussian shift (+1 SD)	3860/4096	3893/4096	3881/4096	4096/4096
Heavy-tail null (t, df=3)	3804/4096	3885/4096	3786/4096	4096/4096
Skewed null (exponential)	3835/4096	3864/4096	3896/4096	4096/4096
5 Gaussian variance strata	3869/4096	3895/4096	3891/4096	4096/4096
Discrete null (Bernoulli 0.1)	3837/4096	3895/4096	1495/4096	1576/4096

Coverage counts use the total denominator: unavailable bootstrap t intervals count as noncoverage. Availability is shown separately, without dropping degenerate samples. Exact binomial intervals and Monte Carlo standard errors accompany rates in the data.

Mean-zero fixture	Rejections	Exact binomial 95% interval	Sign symmetry
Gaussian null	190/4096	[0.040, 0.053]	Satisfied
Heavy-tail null (t, df=3)	224/4096	[0.048, 0.062]	Satisfied
Skewed null (exponential)	231/4096	[0.050, 0.064]	Violated
5 Gaussian variance strata	191/4096	[0.040, 0.054]	Satisfied
Discrete null (Bernoulli 0.1)	232/4096	[0.050, 0.064]	Violated

For the rare Bernoulli fixture, exact bootstrap t availability under the declared failure rule is 38.82%; observed availability is 38.48%. Unconditional coverage is 36.50%, whereas coverage conditional on an available interval is 94.86%. The exact calculation benchmarks availability alone; it does not establish a coverage guarantee.

The +1 SD Gaussian control rejected in 4096/4096 experiments (95% binomial interval [0.999, 1.000]). Skewed mean-zero fixtures deliberately violate the sign-flip assumption; any rejection distortion remains a reported boundary. No seed, benchmark or method is selected to achieve a desired coverage rate.

Across 25 separately admitted grid cells, the positive-minus-random optimism change contrast ranged from -0.00329 to -0.00128; the corresponding Brier contrast ranged from -0.00276 to -0.00106. Each cell used 128 independent agents per fixed stratum, sharing seeds across cells.

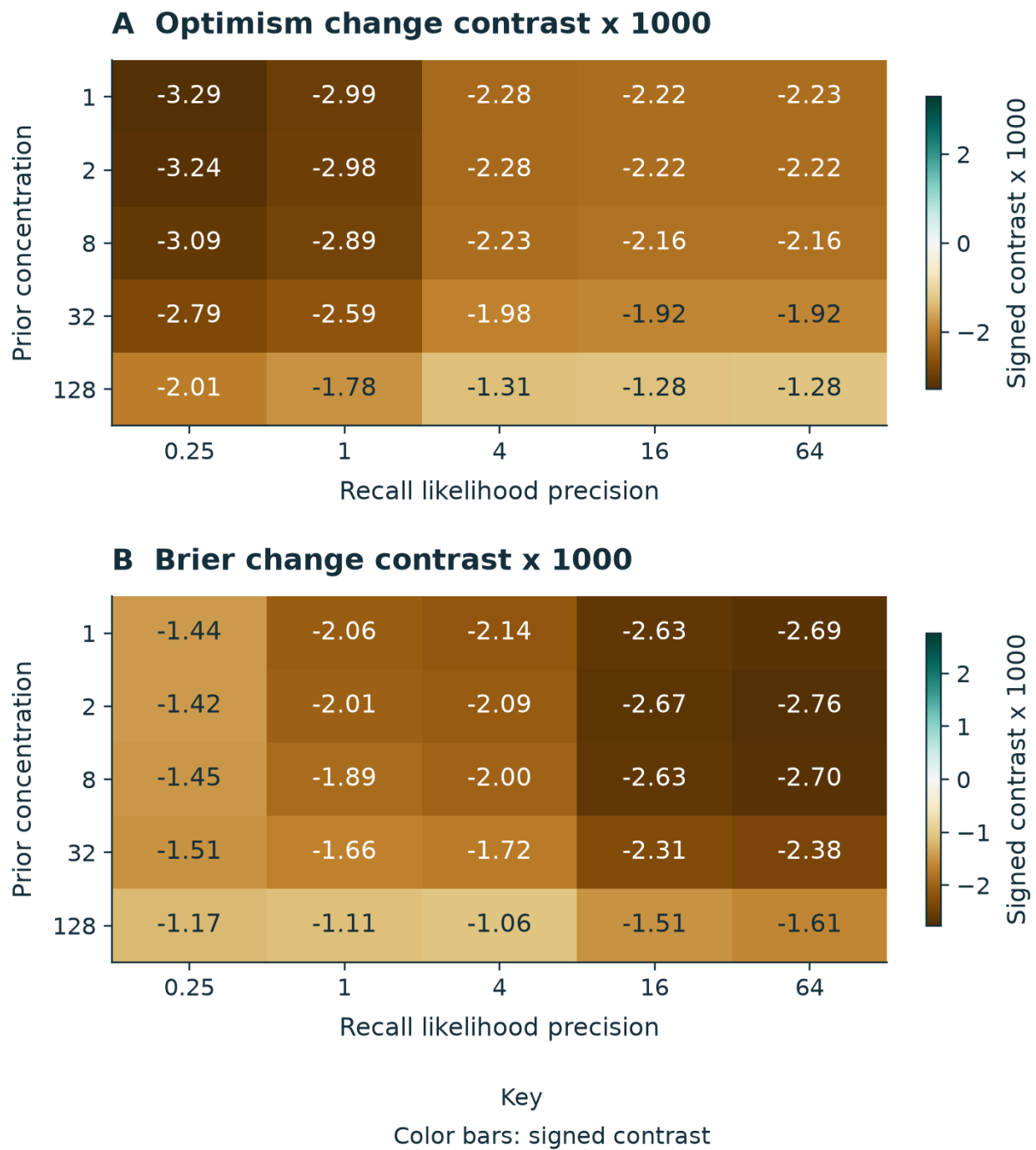
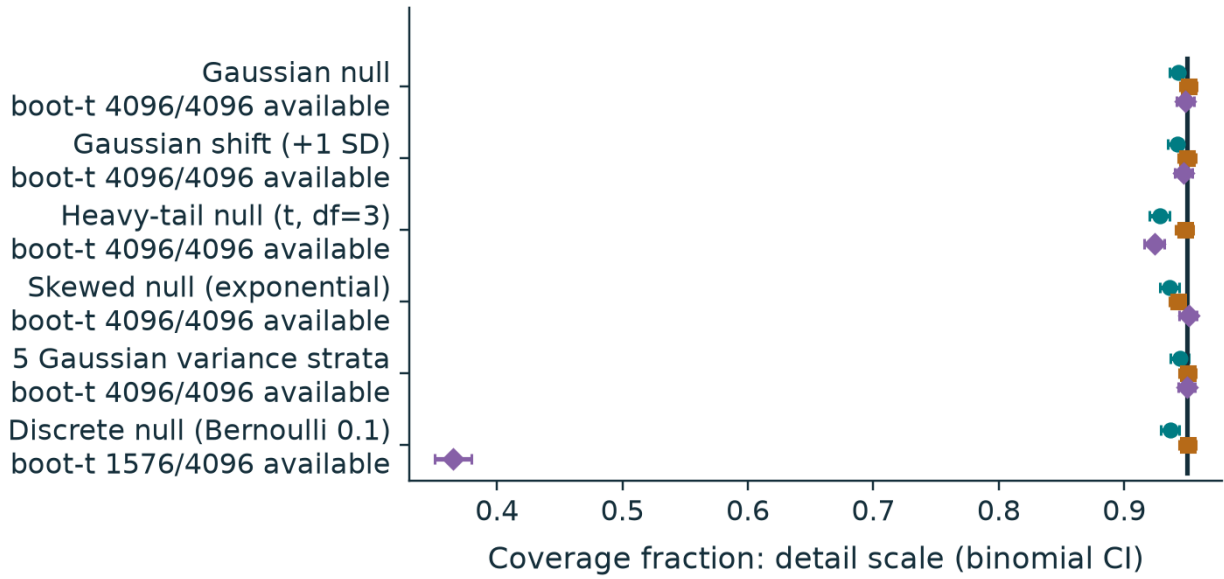


Figure 39: A finite parameter grid tests dependence on declared choices. Each prior-concentration and recall-precision cell is independently admitted from a modified GNN declaration. Colors show positive-minus-random contrasts in endpoint changes; the panels retain their own units and zero-centered scales. Cells share seeds and are conditional fixed-grid comparisons, not independent replications or confidence regions. The grid makes model dependence visible. It does not estimate uncertainty over unexamined parameters or establish human robustness.

The operating-characteristic output separates successful interval coverage, availability and null rejection. Its exact binomial intervals describe uncertainty across independent synthetic experiments. Skew and discrete-boundary stressors show where a mean-zero target does not supply sign symmetry or a usable studentized variance. The controls characterize each tested procedure and distribution; they do not turn software qualification into empirical validation.

Rare Bernoulli bootstrap t
Exact availability: 38.8%; observed: 38.5%
Coverage given available: 94.9%

A Interval coverage



B Mean-zero stress tests

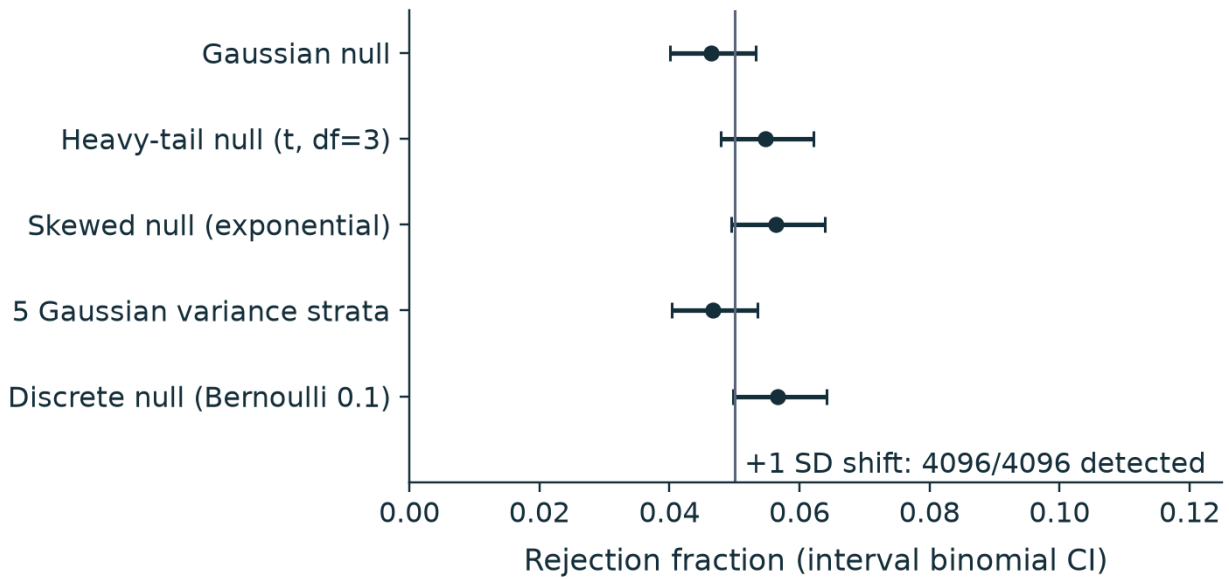


Figure 40: Distribution shape limits the interpretation of nominal confidence. Known-answer fixtures test coverage, interval availability and mean-zero rejection using the same generated samples for every method. The coverage panel spans 6 fixtures and uses a data-derived detail axis; the rejection panel separates null distributions from the positive-shift control. Bootstrap t availability is displayed for every fixture. The Bernoulli annotation distinguishes exact availability, observed availability and coverage conditional on an available interval. Whiskers are exact binomial 95% intervals over independent experiments; unavailable intervals count as noncoverage. Reference lines are 0.95 coverage and 0.05 rejection. Centered exponential and Bernoulli differences violate sign symmetry despite zero means. The exact Bernoulli calculation benchmarks availability under the strict zero-variance failure rule, not coverage. The benchmarks are not retuned to make a method pass.

Complete primary interval comparisons

The following tables retain all retained pointwise 95% procedures on the same primary contrasts. Display scaling matches the primary figures. The analytic interval is a robustness approximation outside the single-Gaussian control; unavailable bootstrap t intervals remain unavailable rather than borrowing another method.

Optimism-task change; all values x1000.

Contrast	Percentile 95%	Bootstrap t 95%	Analytic t 95%
Positive - None	[-2.278, -2.263]	[-2.278, -2.262]	[-2.278, -2.263]
Random - None	[-0.031, -0.004]	[-0.031, -0.004]	[-0.031, -0.004]
Absence - None	[0.000, 0.000]	Unavailable	[0.000, 0.000]
Positive - Random	[-2.267, -2.239]	[-2.267, -2.239]	[-2.267, -2.239]

Brier-loss change; all values x1000.

Contrast	Percentile 95%	Bootstrap t 95%	Analytic t 95%
Positive - None	[0.135, 0.207]	[0.135, 0.207]	[0.135, 0.208]
Random - None	[2.124, 2.228]	[2.125, 2.229]	[2.124, 2.229]
Absence - None	[0.000, 0.000]	Unavailable	[0.000, 0.000]
Positive - Random	[-2.064, -1.945]	[-2.066, -1.946]	[-2.065, -1.946]

Complete practice interval comparisons

All 8 exploratory policy-minus-rigid contrasts retain percentile, studentized bootstrap and analytic t intervals at the paired-replicate level. Units are nats per held-out observation. Unavailable studentized intervals remain labeled; their degenerate resample counts are retained in the numerical receipt.

Context / policy	Percentile 95%	Bootstrap t 95%	Analytic t 95%
Stationary / Inquiry	[+0.0000, +0.0000]	Unavailable (degenerate sample)	[+0.0000, +0.0000]
Stationary / No preference	[+0.0000, +0.0000]	Unavailable (degenerate sample)	[+0.0000, +0.0000]
Stationary / Sham	[+0.0000, +0.0000]	Unavailable (degenerate sample)	[+0.0000, +0.0000]
Stationary / No access	[+0.0000, +0.0000]	Unavailable (degenerate sample)	[+0.0000, +0.0000]
Reversal / Inquiry	[-0.1064, -0.1030]	[-0.1064, -0.1029]	[-0.1064, -0.1030]
Reversal / No preference	[-0.0543, -0.0467]	[-0.0543, -0.0466]	[-0.0543, -0.0466]
Reversal / Sham	[+0.0000, +0.0000]	Unavailable (degenerate sample)	[+0.0000, +0.0000]
Reversal / No access	[+0.0000, +0.0000]	Unavailable (degenerate sample)	[+0.0000, +0.0000]

Complete gaze interval comparisons

These tables include all 58 declared contrasts, including direct comparisons between open sampling policies. Estimates retain their native units. Complete episodes and agents, rather than locations or observations, determine interval precision. Studentization failures are explicit; the underlying exact null is retained.

Transfer: likelihood.

Contrast; endpoint	Mean	Percentile 95%	Bootstrap t 95%	Analytic t 95%
Positive - None; Positive probability change	+0.038492	[+0.038314, +0.038673]	[+0.038312, +0.038670]	[+0.038309, +0.038674]
Positive - None; Scene-risk change	+0.002362	[+0.002166, +0.002548]	[+0.002175, +0.002557]	[+0.002175, +0.002549]
Random - None; Positive probability change	+0.026996	[+0.026838, +0.027152]	[+0.026840, +0.027153]	[+0.026841, +0.027151]
Random - None; Scene-risk change	-0.007860	[-0.008103, -0.007621]	[-0.008099, -0.007619]	[-0.008112, -0.007608]
Absence - None; Positive probability change	+0.000000	[+0.000000, +0.000000]	Unavailable	[+0.000000, +0.000000]
Absence - None; Scene-risk change	+0.000000	[+0.000000, +0.000000]	Unavailable	[+0.000000, +0.000000]
Positive - Random; Positive probability change	+0.011496	[+0.011280, +0.011711]	[+0.011282, +0.011712]	[+0.011281, +0.011710]
Positive - Random; Scene-risk change	+0.010222	[+0.009907, +0.010537]	[+0.009908, +0.010538]	[+0.009908, +0.010536]

Transfer: posterior readout.

Contrast; endpoint	Mean	Percentile 95%	Bootstrap t 95%	Analytic t 95%
Positive - None; Positive probability change	-0.004677	[-0.004693, -0.004661]	[-0.004693, -0.004661]	[-0.004694, -0.004661]
Positive - None; Scene-risk change	+0.000096	[+0.000092, +0.000100]	[+0.000091, +0.000100]	[+0.000092, +0.000100]
Random - None; Positive probability change	-0.000051	[-0.000079, -0.000023]	[-0.000078, -0.000022]	[-0.000078, -0.000023]
Random - None; Scene-risk change	+0.000026	[+0.000024, +0.000027]	[+0.000024, +0.000027]	[+0.000024, +0.000027]
Absence - None; Positive probability change	+0.000000	[+0.000000, +0.000000]	Unavailable	[+0.000000, +0.000000]
Absence - None; Scene-risk change	+0.000000	[+0.000000, +0.000000]	Unavailable	[+0.000000, +0.000000]
Positive - Random; Positive probability change	-0.004627	[-0.004657, -0.004599]	[-0.004654, -0.004597]	[-0.004655, -0.004598]
Positive - Random; Scene-risk change	+0.000070	[+0.000066, +0.000074]	[+0.000066, +0.000074]	[+0.000066, +0.000074]

Active: stationary.

Contrast; endpoint	Mean	Percentile 95%	Bootstrap t 95%	Analytic t 95%
Open utility - Restricted utility; Scene risk	+0.009202	[+0.008492, +0.009968]	[+0.008517, +0.010023]	[+0.008459, +0.009945]
Open utility - Restricted utility; Positive occupancy	-0.034261	[-0.036431, -0.032132]	[-0.036556, -0.032209]	[-0.036362, -0.032160]
Open utility - Restricted utility; Movement penalty	+0.000964	[+0.000847, +0.001082]	[+0.000850, +0.001086]	[+0.000846, +0.001083]
Open inquiry - Restricted utility; Scene risk	+0.009639	[+0.008862, +0.010479]	[+0.008892, +0.010535]	[+0.008832, +0.010447]

Contrast; endpoint	Mean	Percentile 95%	Bootstrap t 95%	Analytic t 95%
Open inquiry - Restricted utility; Positive occupancy	-0.039551	[-0.041865, -0.037277]	[-0.041970, -0.037382]	[-0.041877, -0.037225]
Open inquiry - Restricted utility; Movement penalty	+0.000553	[+0.000437, +0.000669]	[+0.000440, +0.000671]	[+0.000441, +0.000666]
Restricted inquiry - Restricted utility; Scene risk	+0.000003	[-0.000034, +0.000044]	[-0.000032, +0.000047]	[-0.000036, +0.000043]
Restricted inquiry - Restricted utility; Positive occupancy	+0.001424	[+0.001166, +0.001709]	[+0.001164, +0.001716]	[+0.001148, +0.001701]
Restricted inquiry - Restricted utility; Movement penalty	-0.000179	[-0.000238, -0.000124]	[-0.000238, -0.000123]	[-0.000237, -0.000121]
Sham - Restricted utility; Scene risk	+0.000000	[+0.000000, +0.000000]	Unavailable	[+0.000000, +0.000000]
Sham - Restricted utility; Positive occupancy	+0.000000	[+0.000000, +0.000000]	Unavailable	[+0.000000, +0.000000]
Sham - Restricted utility; Movement penalty	+0.000000	[+0.000000, +0.000000]	Unavailable	[+0.000000, +0.000000]
Uniform open - Restricted utility; Scene risk	+0.008750	[+0.007942, +0.009592]	[+0.007950, +0.009602]	[+0.007923, +0.009578]
Uniform open - Restricted utility; Positive occupancy	-0.606906	[-0.610460, -0.603317]	[-0.610493, -0.603334]	[-0.610542, -0.603271]
Uniform open - Restricted utility; Movement penalty	+0.069156	[+0.068742, +0.069572]	[+0.068737, +0.069573]	[+0.068742, +0.069569]
Open inquiry - Open utility; Scene risk	+0.000437	[-0.000043, +0.000918]	[-0.000017, +0.000957]	[-0.000037, +0.000911]
Open inquiry - Open utility; Positive occupancy	-0.005290	[-0.006700, -0.003893]	[-0.006806, -0.003973]	[-0.006693, -0.003887]
Open inquiry - Open utility; Movement penalty	-0.000411	[-0.000514, -0.000308]	[-0.000517, -0.000308]	[-0.000515, -0.000307]
Open inquiry - Uniform open; Scene risk	+0.000889	[-0.000135, +0.001964]	[-0.000137, +0.001956]	[-0.000155, +0.001933]
Open inquiry - Uniform open; Positive occupancy	+0.567356	[+0.563151, +0.571642]	[+0.562984, +0.571531]	[+0.563093, +0.571618]
Open inquiry - Uniform open; Movement penalty	-0.068603	[-0.069032, -0.068183]	[-0.069017, -0.068175]	[-0.069029, -0.068176]

Active: reversal.

Contrast; endpoint	Mean	Percentile 95%	Bootstrap t 95%	Analytic t 95%
Open utility - Restricted utility; Scene risk	-0.115899	[-0.116915, -0.114866]	[-0.116828, -0.114730]	[-0.116933, -0.114865]
Open utility - Restricted utility; Positive occupancy	-0.328586	[-0.331394, -0.325675]	[-0.331356, -0.325587]	[-0.331494, -0.325678]
Open utility - Restricted utility; Movement penalty	-0.000588	[-0.000799, -0.000368]	[-0.000809, -0.000377]	[-0.000804, -0.000373]

Contrast; endpoint	Mean	Percentile 95%	Bootstrap t 95%	Analytic t 95%
Open inquiry - Restricted utility; Scene risk	-0.115884	[-0.116868, -0.114839]	[-0.116824, -0.114765]	[-0.116929, -0.114839]
Open inquiry - Restricted utility; Positive occupancy	-0.335232	[-0.338230, -0.332334]	[-0.338052, -0.332119]	[-0.338137, -0.332327]
Open inquiry - Restricted utility; Movement penalty	-0.001136	[-0.001338, -0.000930]	[-0.001343, -0.000935]	[-0.001340, -0.000933]
Restricted inquiry - Restricted utility; Scene risk	-0.000033	[-0.000071, +0.000005]	[-0.000072, +0.000004]	[-0.000072, +0.000006]
Restricted inquiry - Restricted utility; Positive occupancy	+0.001289	[+0.001044, +0.001560]	[+0.001041, +0.001566]	[+0.001025, +0.001552]
Restricted inquiry - Restricted utility; Movement penalty	-0.000150	[-0.000209, -0.000093]	[-0.000212, -0.000095]	[-0.000208, -0.000092]
Sham - Restricted utility; Scene risk	+0.000000	[+0.000000, +0.000000]	Unavailable	[+0.000000, +0.000000]
Sham - Restricted utility; Positive occupancy	+0.000000	[+0.000000, +0.000000]	Unavailable	[+0.000000, +0.000000]
Sham - Restricted utility; Movement penalty	+0.000000	[+0.000000, +0.000000]	Unavailable	[+0.000000, +0.000000]
Uniform open - Restricted utility; Scene risk	-0.114215	[-0.115239, -0.113110]	[-0.115212, -0.113041]	[-0.115285, -0.113145]
Uniform open - Restricted utility; Positive occupancy	-0.593872	[-0.597421, -0.590350]	[-0.597410, -0.590268]	[-0.597454, -0.590290]
Uniform open - Restricted utility; Movement penalty	+0.067621	[+0.067178, +0.068087]	[+0.067152, +0.068071]	[+0.067168, +0.068073]
Open inquiry - Open utility; Scene risk	+0.000015	[-0.000565, +0.000595]	[-0.000553, +0.000615]	[-0.000570, +0.000601]
Open inquiry - Open utility; Positive occupancy	-0.006646	[-0.008296, -0.005018]	[-0.008400, -0.005119]	[-0.008286, -0.005006]
Open inquiry - Open utility; Movement penalty	-0.000548	[-0.000656, -0.000442]	[-0.000658, -0.000443]	[-0.000656, -0.000441]
Open inquiry - Uniform open; Scene risk	-0.001669	[-0.003012, -0.000253]	[-0.003080, -0.000332]	[-0.003064, -0.000274]
Open inquiry - Uniform open; Positive occupancy	+0.258640	[+0.254304, +0.263075]	[+0.254268, +0.263073]	[+0.254221, +0.263059]
Open inquiry - Uniform open; Movement penalty	-0.068757	[-0.069169, -0.068339]	[-0.069177, -0.068343]	[-0.069182, -0.068331]

Exact constant contrasts have zero-width percentile and analytic intervals; the bootstrap t method reports degenerate `_sample` rather than manufacturing an interval. Probability endpoints are in probability units; loss uses natural-log nats per outcome; movement is the GNN-weighted unit-square distance per sampled step.

Complete joint gaze bounds

These bounds belong to the single exploratory family in [the joint method](#). Each is expressed in the same native units as its pointwise estimate. The family status and conditional critical-quantile bracket appear in [the joint results](#). Exact identities are distinct from unavailable sample constants; no stochastic bound is borrowed from another method.

Transfer: likelihood.

Contrast; endpoint	Joint 95% bounds	Status
Positive - None; Positive probability change	[+0.038193, +0.038790]	estimated
Positive - None; Scene-risk change	[+0.002055, +0.002668]	estimated
Random - None; Positive probability change	[+0.026742, +0.027250]	estimated
Random - None; Scene-risk change	[-0.008273, -0.007448]	estimated
Absence - None; Positive probability change	[+0.000000, +0.000000]	mechanism exact
Absence - None; Scene-risk change	[+0.000000, +0.000000]	mechanism exact
Positive - Random; Positive probability change	[+0.011145, +0.011847]	estimated
Positive - Random; Scene-risk change	[+0.009708, +0.010736]	estimated

Transfer: posterior readout.

Contrast; endpoint	Joint 95% bounds	Status
Positive - None; Positive probability change	[-0.004704, -0.004650]	estimated
Positive - None; Scene-risk change	[+0.000089, +0.000102]	estimated
Random - None; Positive probability change	[-0.000095, -0.000006]	estimated
Random - None; Scene-risk change	[+0.000023, +0.000028]	estimated
Absence - None; Positive probability change	[+0.000000, +0.000000]	mechanism exact
Absence - None; Scene-risk change	[+0.000000, +0.000000]	mechanism exact
Positive - Random; Positive probability change	[-0.004674, -0.004580]	estimated
Positive - Random; Scene-risk change	[+0.000064, +0.000076]	estimated

Active: stationary.

Contrast; endpoint	Joint 95% bounds	Status
Open utility - Restricted utility; Scene risk	[+0.007988, +0.010416]	estimated
Open utility - Restricted utility; Positive occupancy	[-0.037694, -0.030828]	estimated
Open utility - Restricted utility; Movement penalty	[+0.000770, +0.001158]	estimated
Open inquiry - Restricted utility; Scene risk	[+0.008320, +0.010959]	estimated
Open inquiry - Restricted utility; Positive occupancy	[-0.043351, -0.035751]	estimated
Open inquiry - Restricted utility; Movement penalty	[+0.000369, +0.000737]	estimated
Restricted inquiry - Restricted utility; Scene risk	[-0.000062, +0.000068]	estimated
Restricted inquiry - Restricted utility; Positive occupancy	[+0.000972, +0.001876]	estimated
Restricted inquiry - Restricted utility; Movement penalty	[-0.000273, -0.000084]	estimated
Sham - Restricted utility; Scene risk	[+0.000000, +0.000000]	mechanism exact
Sham - Restricted utility; Positive occupancy	[+0.000000, +0.000000]	mechanism exact
Sham - Restricted utility; Movement penalty	[+0.000000, +0.000000]	mechanism exact
Uniform open - Restricted utility; Scene risk	[+0.007398, +0.010103]	estimated
Uniform open - Restricted utility; Positive occupancy	[-0.612847, -0.600966]	estimated
Uniform open - Restricted utility; Movement penalty	[+0.068480, +0.069831]	estimated

Contrast; endpoint	Joint 95% bounds	Status
Open inquiry - Open utility; Scene risk	[-0.000338, +0.001212]	estimated
Open inquiry - Open utility; Positive occupancy	[-0.007582, -0.002997]	estimated
Open inquiry - Open utility; Movement penalty	[-0.000582, -0.000240]	estimated
Open inquiry - Uniform open; Scene risk	[-0.000816, +0.002595]	estimated
Open inquiry - Uniform open; Positive occupancy	[+0.560391, +0.574320]	estimated
Open inquiry - Uniform open; Movement penalty	[-0.069299, -0.067906]	estimated

Active: reversal.

Contrast; endpoint	Joint 95% bounds	Status
Open utility - Restricted utility; Scene risk	[-0.117588, -0.114210]	estimated
Open utility - Restricted utility; Positive occupancy	[-0.333338, -0.323834]	estimated
Open utility - Restricted utility; Movement penalty	[-0.000940, -0.000236]	estimated
Open inquiry - Restricted utility; Scene risk	[-0.117591, -0.114177]	estimated
Open inquiry - Restricted utility; Positive occupancy	[-0.339979, -0.330486]	estimated
Open inquiry - Restricted utility; Movement penalty	[-0.001469, -0.000804]	estimated
Restricted inquiry - Restricted utility; Scene risk	[-0.000096, +0.000030]	estimated
Restricted inquiry - Restricted utility; Positive occupancy	[+0.000858, +0.001719]	estimated
Restricted inquiry - Restricted utility; Movement penalty	[-0.000245, -0.000055]	estimated
Sham - Restricted utility; Scene risk	[+0.000000, +0.000000]	mechanism exact
Sham - Restricted utility; Positive occupancy	[+0.000000, +0.000000]	mechanism exact
Sham - Restricted utility; Movement penalty	[+0.000000, +0.000000]	mechanism exact
Uniform open - Restricted utility; Scene risk	[-0.115964, -0.112467]	estimated
Uniform open - Restricted utility; Positive occupancy	[-0.599724, -0.588020]	estimated
Uniform open - Restricted utility; Movement penalty	[+0.066882, +0.068359]	estimated
Open inquiry - Open utility; Scene risk	[-0.000941, +0.000972]	estimated
Open inquiry - Open utility; Positive occupancy	[-0.009326, -0.003966]	estimated
Open inquiry - Open utility; Movement penalty	[-0.000724, -0.000372]	estimated
Open inquiry - Uniform open; Scene risk	[-0.003948, +0.000611]	estimated
Open inquiry - Uniform open; Positive occupancy	[+0.251420, +0.265860]	estimated
Open inquiry - Uniform open; Movement penalty	[-0.069452, -0.068061]	estimated

Complete scene/boundary intersection bounds

The following is a separate 84-member exploratory family. Z scene strata are fixed and equally weighted; paired independent replicates are the resampling units. Anchored loss refers to the stipulated H=0 reference intervention. Exact nonzero offsets retain their adverse or favorable signs. Unavailable constants, if any, do not receive artificial zero-width intervals.

Regime / H	Contrast	Endpoint	Mean	Joint interval	Status
diagnostic / H0	Integrated - Visual only	Anchor loss	-0.101220	[-0.107449, -0.094991]	estimated
diagnostic / H0	Integrated - Visual only	Boundary Brier	-0.612571	[-0.638418, -0.586724]	estimated
diagnostic / H0	Boundary only - Visual only	Anchor loss	0.000000	[0.000000, 0.000000]	mechanism exact
diagnostic / H0	Boundary only - Visual only	Boundary Brier	-0.666667	[-0.666667, -0.666667]	estimated
diagnostic / H0	Uniform - Visual only	Anchor loss	-0.096426	[-0.101357, -0.091495]	estimated
diagnostic / H0	Uniform - Visual only	Boundary Brier	-0.591520	[-0.614552, -0.568488]	estimated
diagnostic / H0	Restricted - Visual only	Anchor loss	-0.105880	[-0.110288, -0.101473]	estimated
diagnostic / H0	Restricted - Visual only	Boundary Brier	-0.616534	[-0.641923, -0.591145]	estimated
diagnostic / H0	Yoked sham - Visual only	Anchor loss	0.000000	[0.000000, 0.000000]	mechanism exact
diagnostic / H0	Yoked sham - Visual only	Boundary Brier	0.000000	[0.000000, 0.000000]	mechanism exact
diagnostic / H0	Integrated - Uniform	Anchor loss	-0.004794	[-0.012426, 0.002838]	estimated
diagnostic / H0	Integrated - Uniform	Boundary Brier	-0.021051	[-0.053491, 0.011389]	estimated
diagnostic / H0	Integrated - Boundary only	Anchor loss	-0.101220	[-0.107449, -0.094991]	estimated
diagnostic / H0	Integrated - Boundary only	Boundary Brier	0.054096	[0.028249, 0.079943]	estimated
diagnostic / H1	Integrated - Visual only	Anchor loss	-0.054426	[-0.064589, -0.044263]	estimated
diagnostic / H1	Integrated - Visual only	Boundary Brier	-0.611300	[-0.637446, -0.585153]	estimated
diagnostic / H1	Boundary only - Visual only	Anchor loss	0.000000	[0.000000, 0.000000]	mechanism exact
diagnostic / H1	Boundary only - Visual only	Boundary Brier	-0.666667	[-0.666667, -0.666667]	estimated
diagnostic / H1	Uniform - Visual only	Anchor loss	-0.046550	[-0.055812, -0.037287]	estimated
diagnostic / H1	Uniform - Visual only	Boundary Brier	-0.584887	[-0.608699, -0.561076]	estimated
diagnostic / H1	Restricted - Visual only	Anchor loss	-0.057989	[-0.068054, -0.047924]	estimated
diagnostic / H1	Restricted - Visual only	Boundary Brier	-0.625979	[-0.648755, -0.603203]	estimated
diagnostic / H1	Yoked sham - Visual only	Anchor loss	0.000000	[0.000000, 0.000000]	mechanism exact
diagnostic / H1	Yoked sham - Visual only	Boundary Brier	0.000000	[0.000000, 0.000000]	mechanism exact
diagnostic / H1	Integrated - Uniform	Anchor loss	-0.007876	[-0.021187, 0.005435]	estimated
diagnostic / H1	Integrated - Uniform	Boundary Brier	-0.026412	[-0.060705, 0.007881]	estimated
diagnostic / H1	Integrated - Boundary only	Anchor loss	-0.054426	[-0.064589, -0.044263]	estimated
diagnostic / H1	Integrated - Boundary only	Boundary Brier	0.055367	[0.029220, 0.081514]	estimated
diagnostic / H2	Integrated - Visual only	Anchor loss	-0.079314	[-0.086584, -0.072044]	estimated
diagnostic / H2	Integrated - Visual only	Boundary Brier	-0.618498	[-0.642679, -0.594317]	estimated
diagnostic / H2	Boundary only - Visual only	Anchor loss	0.000000	[0.000000, 0.000000]	mechanism exact
diagnostic / H2	Boundary only - Visual only	Boundary Brier	-0.666667	[-0.666667, -0.666667]	estimated

Regime / H	Contrast	Endpoint	Mean	Joint interval	Status
diagnostic / H2	Uniform - Visual only	Anchor loss	-0.065590	[-0.072833, -0.058347]	estimated
diagnostic / H2	Uniform - Visual only	Boundary Brier	-0.600923	[-0.621706, -0.580141]	estimated
diagnostic / H2	Restricted - Visual only	Anchor loss	0.115680	[0.101109, 0.130252]	estimated
diagnostic / H2	Restricted - Visual only	Boundary Brier	1.127626	[1.110108, 1.145144]	estimated
diagnostic / H2	Yoked sham - Visual only	Anchor loss	0.000000	[0.000000, 0.000000]	mechanism exact
diagnostic / H2	Yoked sham - Visual only	Boundary Brier	0.000000	[0.000000, 0.000000]	mechanism exact
diagnostic / H2	Integrated - Uniform	Anchor loss	-0.013723	[-0.023993, -0.003454]	estimated
diagnostic / H2	Integrated - Uniform	Boundary Brier	-0.017575	[-0.049050, 0.013901]	estimated
diagnostic / H2	Integrated - Boundary only	Anchor loss	-0.079314	[-0.086584, -0.072044]	estimated
diagnostic / H2	Integrated - Boundary only	Boundary Brier	0.048169	[0.023988, 0.072349]	estimated
label only / H0	Integrated - Visual only	Anchor loss	-0.000195	[-0.001288, 0.000898]	estimated
label only / H0	Integrated - Visual only	Boundary Brier	0.000000	[0.000000, 0.000000]	mechanism exact
label only / H0	Boundary only - Visual only	Anchor loss	0.113064	[0.111446, 0.114681]	estimated
label only / H0	Boundary only - Visual only	Boundary Brier	0.000000	[0.000000, 0.000000]	mechanism exact
label only / H0	Uniform - Visual only	Anchor loss	0.005338	[0.001974, 0.008701]	estimated
label only / H0	Uniform - Visual only	Boundary Brier	0.000000	[0.000000, 0.000000]	mechanism exact
label only / H0	Restricted - Visual only	Anchor loss	-0.000195	[-0.001288, 0.000898]	estimated
label only / H0	Restricted - Visual only	Boundary Brier	-0.166667	[-0.166667, -0.166667]	mechanism exact
label only / H0	Yoked sham - Visual only	Anchor loss	-0.000195	[-0.001288, 0.000898]	estimated
label only / H0	Yoked sham - Visual only	Boundary Brier	0.000000	[0.000000, 0.000000]	mechanism exact
label only / H0	Integrated - Uniform	Anchor loss	-0.005533	[-0.008889, -0.002177]	estimated
label only / H0	Integrated - Uniform	Boundary Brier	0.000000	[0.000000, 0.000000]	mechanism exact
label only / H0	Integrated - Boundary only	Anchor loss	-0.113259	[-0.114819, -0.111698]	estimated
label only / H0	Integrated - Boundary only	Boundary Brier	0.000000	[0.000000, 0.000000]	mechanism exact
label only / H1	Integrated - Visual only	Anchor loss	0.000142	[-0.001480, 0.001763]	estimated
label only / H1	Integrated - Visual only	Boundary Brier	0.000000	[0.000000, 0.000000]	mechanism exact
label only / H1	Boundary only - Visual only	Anchor loss	0.112967	[0.111599, 0.114334]	estimated
label only / H1	Boundary only - Visual only	Boundary Brier	0.000000	[0.000000, 0.000000]	mechanism exact
label only / H1	Uniform - Visual only	Anchor loss	0.004775	[0.001831, 0.007719]	estimated
label only / H1	Uniform - Visual only	Boundary Brier	0.000000	[0.000000, 0.000000]	mechanism exact
label only / H1	Restricted - Visual only	Anchor loss	0.000142	[-0.001480, 0.001763]	estimated
label only / H1	Restricted - Visual only	Boundary Brier	-0.166667	[-0.166667, -0.166667]	mechanism exact

Regime / H	Contrast	Endpoint	Mean	Joint interval	Status
label only / H1	Yoked sham - Visual only	Anchor loss	0.000142	[-0.001480, 0.001763]	estimated
label only / H1	Yoked sham - Visual only	Boundary Brier	0.000000	[0.000000, 0.000000]	mechanism exact
label only / H1	Integrated - Uniform	Anchor loss	-0.004634	[-0.007733, -0.001534]	estimated
label only / H1	Integrated - Uniform	Boundary Brier	0.000000	[0.000000, 0.000000]	mechanism exact
label only / H1	Integrated - Boundary only	Anchor loss	-0.112825	[-0.114504, -0.111147]	estimated
label only / H1	Integrated - Boundary only	Boundary Brier	0.000000	[0.000000, 0.000000]	mechanism exact
label only / H2	Integrated - Visual only	Anchor loss	-0.000270	[-0.001567, 0.001027]	estimated
label only / H2	Integrated - Visual only	Boundary Brier	0.000000	[0.000000, 0.000000]	mechanism exact
label only / H2	Boundary only - Visual only	Anchor loss	0.112203	[0.110722, 0.113684]	estimated
label only / H2	Boundary only - Visual only	Boundary Brier	0.000000	[0.000000, 0.000000]	mechanism exact
label only / H2	Uniform - Visual only	Anchor loss	0.003161	[0.000451, 0.005871]	estimated
label only / H2	Uniform - Visual only	Boundary Brier	0.000000	[0.000000, 0.000000]	mechanism exact
label only / H2	Restricted - Visual only	Anchor loss	-0.000270	[-0.001567, 0.001027]	estimated
label only / H2	Restricted - Visual only	Boundary Brier	0.833333	[0.833333, 0.833333]	mechanism exact
label only / H2	Yoked sham - Visual only	Anchor loss	-0.000270	[-0.001567, 0.001027]	estimated
label only / H2	Yoked sham - Visual only	Boundary Brier	0.000000	[0.000000, 0.000000]	mechanism exact
label only / H2	Integrated - Uniform	Anchor loss	-0.003431	[-0.006028, -0.000834]	estimated
label only / H2	Integrated - Uniform	Boundary Brier	0.000000	[0.000000, 0.000000]	mechanism exact
label only / H2	Integrated - Boundary only	Anchor loss	-0.112473	[-0.113747, -0.111198]	estimated
label only / H2	Integrated - Boundary only	Boundary Brier	0.000000	[0.000000, 0.000000]	mechanism exact

S9. Model settings injected from admitted declarations

Model / setting	Admitted value
study seed	20261015
study replicates	1024
study luckiness	0.1, 0.3, 0.5, 0.7, 0.9
study baseline days	7
study intervention days	7
study steps per day	12
study recall count	3
study recall precision	4
study learning rate	2
study forgetting	1
study prior concentration	8
study bootstrap draws	8191
study permutation draws	8191
study assessment trials	70
study task prior min	0.05
study task prior max	0.95
study sensitivity replicates	128
study sensitivity prior concentrations	1, 2, 8, 32, 128

Model / setting	Admitted value
study sensitivity recall precisions	0.25, 1, 4, 16, 64
study calibration replicates	4096
study calibration sample size	64
study calibration bootstrap draws	1023
study calibration permutation draws	1023
study evaluation trials	1024
study reference hazard	0.3
study harm cost	1
practice seed	1729
practice replicates	512
practice calibration samples	24
practice assessment samples	512
practice bootstrap draws	4095
practice action precision	6
practice practice cost	0.2
practice examinability preference	0.4
practice previous gain weight	1
practice metacognitive access	1
practice prior precision	2
gaze seed	91073
gaze source seed	42
gaze source steps	1000
gaze source runs	5
gaze transfer steps	128
gaze transfer ambiguity weight	0.5
gaze transfer precision	1.5
gaze trajectory steps	20
gaze example strata	0.3, 0.7
gaze example window	1
gaze replicates	512
gaze steps	48
gaze bootstrap draws	4095
gaze action precision	4
gaze information weight	4
gaze saccade cost	0.2
gaze structural model	emptiness_practice
gaze initial location	10
gaze regions	0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 2, 2, 2
boundary seed	620715
boundary replicates	512
boundary steps	24
boundary prefixes	16, 32, 64, 128, 256, 512
boundary probe reliabilities	0.65, 0.8, 0.95
boundary probe costs	0.02, 0.05, 0.1
boundary endpoints	anchored_regret, boundary_brier
boundary gaze model	gaze
boundary anchor mapping	0
boundary joint seed	620716
joint seed	91074
joint draws	16383
joint confidence	0.95
joint monte carlo confidence	0.99
joint calibration experiments	4096
joint calibration units	64
joint calibration draws	1023
replicates	1024
practice reversal schedule	0.85 for 4 windows; 0.15 for 4 windows; 0.85 for 4 windows
practice stationary probability	0.85
boundary middle reliability	0.8

S10. Textual editions, translation choices and comparison limits

The [Figure 2](#) adds further edition-specific distinctions. Guru Nanak’s *Japji* is read in Sant Singh Khalsa’s undated translation, with the conclusion and immediate context separately identified ([Guru Nanak, n.d.](#), ang 1, pauris 1–2). The *Analects* reciprocity

exchange is located under different numbering in the two inspected editions: Legge’s edition and Eno’s teaching translation (Confucius 2002, XV.23; 2015, 15.24). Kulp’s *Avot* wording is lot, with its psalm and labor context retained (Kulp, n.d.-b, 4:1). The Qur’an addition separates translation from tafsir, as before (Khatab, n.d.-b, 31:12), and the *Gītā* addition distinguishes sacrificial participants in the verses from modern commentary’s broader analogies (Swami Mukundananda, n.d.-b, 3.10–13). Our MMK summary is an attributed paraphrase of the inspected Sanskrit e-text, whose print-source field is blank; the digital conversion date does not date Nāgārjuna or establish a critical edition (Nagarjuna 2020, 24.8–10 and 24.18). The source companion attaches these qualifications to the authored comparison rows. The selective atlas gives no common scale and makes no claim about traditions or witnesses omitted from it.

The expanded **main-text comparison** uses named primary-text editions. Thich Nhat Hanh’s *Heart Sutra* rendering is explicitly a contemporary revision; its explanation is attributed to the translator rather than silently treated as ancient scripture (Thich Nhat Hanh 2014). The *Bhagavadgītā* quotation follows Mukundananda’s Translation field, distinguished from his commentary (Swami Mukundananda, n.d.-a). Christian thanksgiving follows the NRSVUE text, whose date identifies translation copyright (*New Revised Standard Version Updated Edition* 2021). The Qur’an passage follows the credited Khatab translation rather than the site’s tafsir (Khatab, n.d.-a). Mishnah wording follows Joshua Kulp’s named version; the Talmudic blessing and discussion follow the separately identified William Davidson annotated English edition (Kulp, n.d.-a; *Berakhot*, n.d.). The ellipsis within the short blessing excerpt belongs to that displayed edition. Rolt’s Pseudo-Dionysius translation is Christian theological writing rather than a biblical passage; interleaved translator notes remain distinct from the translated argument (Pseudo-Dionysius the Areopagite, n.d.). Where an edition’s publication year was not established, the bibliography leaves it undated. These targeted readings do not constitute a systematic or philological survey, establish historical influence, or support a shared psychological mechanism.

The main text quotes Thānissaro Bhikkhu’s catalogued translations of the Kaccayanagotta discourse and the Cula-suññata discourse (Thanissaro Bhikkhu 1997b, 1997a). The Access to Insight edition identifies translation copyright separately from its later site revision. Current dhammatalks editions revise introductions, notes and some wording; those later translator notes are interpretations and are not silently substituted for the cited discourse. A separately inspected Bhikkhu Sujato comparison preserves the absent/present distinction and the conditioned character of refined concentration. It is a translation comparison, not an assertion of doctrinal consensus. The source dossier records exact pages, segments, acquisition failures and quotation checks. The added readings distinguish a translated discourse from its translator’s introduction, and ancient composition from an inspected translation or digital edition’s publication. Siderits’ account explains the reliance on transmitted texts and subsequent interpretation (Siderits 2026, sec. 1). These witnesses do not supply unmediated access to a single original doctrine. The Epictetus citation follows the visually inspected scanned title and pages; the conflicting TEI date remains disclosed in the source dossier. These texts constrain the philosophical interpretation of nothingness; they do not validate a gratitude intervention or identify Bayesian model revision with the realization they describe.

References

- Algoe, Sara B. 2012. “Find, Remind, and Bind: The Functions of Gratitude in Everyday Relationships.” *Social and Personality Psychology Compass* 6 (6): 455–69. <https://doi.org/10.1111/j.1751-9004.2012.00439.x>.
- Algoe, Sara B., Jonathan Haidt, and Shelly L. Gable. 2008. “Beyond Reciprocity: Gratitude and Relationships in Everyday Life.” *Emotion* 8 (3): 425–29. <https://doi.org/10.1037/1528-3542.8.3.425>.
- Bartlett, Monica Y., and David DeSteno. 2006. “Gratitude and Prosocial Behavior: Helping When It Costs You.” *Psychological Science* 17 (4): 319–25. <https://doi.org/10.1111/j.1467-9280.2006.01705.x>.
- Bennett, Daniel, Angela Radulescu, Sam Zorowitz, Valkyrie Falso, and Yael Niv. 2023. “Affect-congruent attention modulates generalized reward expectations.” *PLOS Computational Biology* 19 (12): e1011707. <https://doi.org/10.1371/journal.pcbi.1011707>.
- Berakhot*. n.d. William Davidson Edition – English, Sefaria. Accessed October 20, 2026. <https://www.sefaria.org/Berakhot.60b.5?lang=bi&ven=William%20Davidson%20Edition%20-%20English>.
- boot package documentation. n.d. *Nonparametric Bootstrap Confidence Intervals (boot.ci)*. R Documentation. <https://www.stat.math.ethz.ch/R-manual/R-devel/library/boot/html/boot.ci.html>.
- BRIER, GLENN W. 1950. “VERIFICATION OF FORECASTS EXPRESSED IN TERMS OF PROBABILITY.” *Monthly Weather Review* 78 (1): 1–3. [https://doi.org/10.1175/1520-0493\(1950\)078%3C0001:vofeit%3E2.0.co;2](https://doi.org/10.1175/1520-0493(1950)078%3C0001:vofeit%3E2.0.co;2).
- Castelhano, Monica S., Michael L. Mack, and John M. Henderson. 2009. “Viewing Task Influences Eye Movement Control During Active Scene Perception.” *Journal of Vision* 9 (3): 6, 1–15. <https://doi.org/10.1167/9.3.6>.
- Chan, Alan. 2025. “Laozi.” In *The Stanford Encyclopedia of Philosophy*, Winter 2025, edited by Edward N. Zalta and Uri Nodelman. Metaphysics Research Lab, Stanford University. <https://plato.stanford.edu/archives/win2025/entries/laozi/>.
- Choi, Hyewon, Youngjae Cha, Michael E. McCullough, Nicholas A. Coles, and Shigehiro Oishi. 2025. “A meta-analysis of the effectiveness of gratitude interventions on well-being across cultures.” *Proceedings of the National Academy of Sciences* 122 (28). <https://doi.org/10.1073/pnas.2425193122>.

- CLOPPER, C. J., and E. S. PEARSON. 1934. "THE USE OF CONFIDENCE OR FIDUCIAL LIMITS ILLUSTRATED IN THE CASE OF THE BINOMIAL." *Biometrika* 26 (4): 404–13. <https://doi.org/10.1093/biomet/26.4.404>.
- Confucius. 2002. *The Analects of Confucius (from the Chinese Classics)*. Translated by James Legge. Project Gutenberg, ebook 3330. <https://www.gutenberg.org/cache/epub/3330/pg3330-images.html>.
- Confucius. 2015. *The Analects of Confucius: An Online Teaching Translation*. Translated by Robert Eno. IUScholarWorks, Indiana University. <https://hdl.handle.net/2022/23420>.
- Coricelli, Giorgio, Hugo D Critchley, Mateus Joffily, John P O'Doherty, Angela Sirigu, and Raymond J Dolan. 2005. "Regret and its avoidance: a neuroimaging study of choice behavior." *Nature Neuroscience* 8 (9): 1255–62. <https://doi.org/10.1038/nm1514>.
- Cregg, David R., and Jennifer S. Cheavens. 2021. "Gratitude Interventions: Effective Self-help? A Meta-analysis of the Impact on Symptoms of Depression and Anxiety." *Journal of Happiness Studies* 22 (1): 413–45. <https://doi.org/10.1007/s10902-020-00236-6>.
- Cronbach, Lee J., and Paul E. Meehl. 1955. "Construct Validity in Psychological Tests." *Psychological Bulletin* 52 (4): 281–302. <https://doi.org/10.1037/h0040957>.
- Cunha, Lúzie Fofonka, Lucia Campos Pellanda, and Caroline Tozzi Reppold. 2019. "Positive Psychology and Gratitude Interventions: A Randomized Clinical Trial." *Frontiers in Psychology* 10. <https://doi.org/10.3389/fpsyg.2019.00584>.
- Dahl, Cortland J., Antoine Lutz, and Richard J. Davidson. 2015. "Reconstructing and Deconstructing the Self: Cognitive Mechanisms in Meditation Practice." *Trends in Cognitive Sciences* 19 (9): 515–23. <https://doi.org/10.1016/j.tics.2015.07.001>.
- Davis, Don E., Elise Choe, Joel Meyers, et al. 2016. "Thankful for the little things: A meta-analysis of gratitude interventions." *Journal of Counseling Psychology* 63 (1): 20–31. <https://doi.org/10.1037/cou0000107>.
- DeSteno, David, Ye Li, Leah Dickens, and Jennifer S. Lerner. 2014. "Gratitude: A Tool for Reducing Economic Impatience." *Psychological Science* 25 (6): 1262–67. <https://doi.org/10.1177/0956797614529979>.
- Efron, B. 1979. "Bootstrap Methods: Another Look at the Jackknife." *The Annals of Statistics* 7 (1). <https://doi.org/10.1214/aos/1176344552>.
- Epictetus. 1887. *The Discourses of Epictetus; with the Encheiridion and Fragments*. Translated by George Long. George Bell; Sons. <https://archive.org/details/dscrsesepictetus00epiciala>.
- Epstude, Kai, and Neal J. Roese. 2008. "The Functional Theory of Counterfactual Thinking." *Personality and Social Psychology Review* 12 (2): 168–92. <https://doi.org/10.1177/1088868308316091>.
- Fisher, Elizabeth L., Lars Sandved Smith, and Jakob Hohwy. 2025. *Computational Mechanisms of Gratitude Practice*. PsyArXiv preprint, version 1. https://osf.io/preprints/psyarxiv/xm25y_v1.
- Fisher, Elizabeth L., Lars Sandved Smith, and Jakob Hohwy. 2026. "Computational Mechanisms of Gratitude Practice." *Neural Computation*, 1–23. <https://doi.org/10.1162/neco.a.1591>.
- Fleming, Stephen M., and Hakwan C. Lau. 2014. "How to Measure Metacognition." *Frontiers in Human Neuroscience* 8: 443. <https://doi.org/10.3389/fnhum.2014.00443>.
- Friedman, Daniel Ari. 2026. *Realizing Emptiness*. Research software, version 3.0.0. https://github.com/docxology/realizing_emptiness/tree/bd91fa1c230551180ed741d1951840e3628da35d.
- Friston, Karl, Thomas FitzGerald, Francesco Rigoli, Philipp Schwartenbeck, and Giovanni Pezzulo. 2017. "Active Inference: A Process Theory." *Neural Computation* 29 (1): 1–49. https://doi.org/10.1162/neco_a_00912.
- Garfield, Jay L. 1995. *The Fundamental Wisdom of the Middle Way: Nāgārjuna's Mūlamadhyamakakārikā*. Oxford University Press. <https://doi.org/10.1093/oso/9780195103175.001.0001>.
- Garfield, Jay L. 2001. *Empty Words: Buddhist Philosophy and Cross-Cultural Interpretation*. Oxford University Press. <https://doi.org/10.1093/oso/9780195145519.001.0001>.
- Garfield, Jay L., and Sonam Thakchöe. 2011. "Identifying the Object of Negation and the Status of Conventional Truth: Why the dGag Bya Matters so Much to Tibetan Mādhyamikas." In *Moonshadows: Conventional Truth in Buddhist Philosophy*, edited by The Cowherds. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199751426.003.0005>.
- Garrett, Neil, and Tali Sharot. 2017. "Optimistic update bias holds firm: Three tests of robustness following Shah et al." *Consciousness and Cognition* 50: 12–22. <https://doi.org/10.1016/j.concog.2016.10.013>.
- Garrett, Neil, Tali Sharot, Paul Faulkner, Christoph W. Korn, Jonathan P. Roiser, and Raymond J. Dolan. 2014. "Losing the rose tinted glasses: neural substrates of unbiased belief updating in depression." *Frontiers in Human Neuroscience* 8. <https://doi.org/10.3389/fnhum.2014.00639>.

- Gelman, Andrew, and Cosma Rohilla Shalizi. 2013. "Philosophy and the Practice of Bayesian Statistics." *British Journal of Mathematical and Statistical Psychology* 66 (1): 8–38. <https://doi.org/10.1111/j.2044-8317.2011.02037.x>.
- Gneiting, Tilmann, and Adrian E Raftery. 2007. "Strictly Proper Scoring Rules, Prediction, and Estimation." *Journal of the American Statistical Association* 102 (477): 359–78. <https://doi.org/10.1198/016214506000001437>.
- Goh, Rui Zhe, Ian B. Phillips, and Chaz Firestone. 2023. "The perception of silence." *Proceedings of the National Academy of Sciences* 120 (29). <https://doi.org/10.1073/pnas.2301463120>.
- Guru Nanak. n.d. *Japji Sahib, in Sri Guru Granth Sahib*. Translated by Sant Singh Khalsa. SriGranth.org. Accessed October 20, 2026. <https://www.sriganth.org/servlet/gurbani.gurbani?Action=Page&Param=1>.
- Hall, Peter. 1988. "Theoretical Comparison of Bootstrap Confidence Intervals." *The Annals of Statistics* 16 (3): 927–53. <https://doi.org/10.1214/aos/1176350933>.
- Hanley, James A., and Abby Lippman-Hand. 1983. "If Nothing Goes Wrong, Is Everything All Right? Interpreting Zero Numerators." *JAMA* 249 (13): 1743–45. <https://doi.org/10.1001/jama.1983.03330370053031>.
- Hartanto, Andree, Manmeet Kaur, K. T. A. Sandeeshwara Kasturiratna, Frosch Y. X. Quek, and Nadyanna M. Majeed. 2023. "A critical examination of the effectiveness of gratitude intervention on well-Being Outcomes: A within-person experimental daily diary approach." *The Journal of Positive Psychology* 18 (6): 942–57. <https://doi.org/10.1080/17439760.2022.2154704>.
- He, Wuming, Junjie Qiu, Yingying Chen, and Yufang Zhong. 2022. "Gratitude Intervention Evokes Indebtedness: Moderated by Perceived Social Distance." *Frontiers in Psychology* 13. <https://doi.org/10.3389/fpsyg.2022.824326>.
- Heckendorf, Hanna, Dirk Lehr, David Daniel Ebert, and Henning Freund. 2019. "Efficacy of an internet and app-based gratitude intervention in reducing repetitive negative thinking and mechanisms of change in the intervention's effect on anxiety and depression: Results from a randomized controlled trial." *Behaviour Research and Therapy* 119: 103415. <https://doi.org/10.1016/j.brat.2019.103415>.
- Hesp, Casper, Ryan Smith, Thomas Parr, Micah Allen, Karl J. Friston, and Maxwell J. D. Ramstead. 2021. "Deeply Felt Affect: The Emergence of Valence in Deep Active Inference." *Neural Computation* 33 (2): 398–446. https://doi.org/10.1162/neco_a_01341.
- Hobbs, Catherine, Petra Vozarova, Aarushi Sabharwal, Punit Shah, and Katherine Button. 2022. "Is depression associated with reduced optimistic belief updating?" *Royal Society Open Science* 9 (2). <https://doi.org/10.1098/rsos.190814>.
- Holm, Sture. 1979. "A Simple Sequentially Rejective Multiple Test Procedure." *Scandinavian Journal of Statistics* 6 (2): 65–70. <https://www.jstor.org/stable/4615733>.
- Hothorn, Torsten, Frank Bretz, and Peter Westfall. 2008. "Simultaneous Inference in General Parametric Models." *Biometrical Journal* 50 (3): 346–63. <https://doi.org/10.1002/bimj.200810425>.
- Kerry, Nicholas, Ria Chhabra, and Jeremy Clifton. 2023. "Being Thankful for What You Have: A Systematic Review of Evidence for the Effect of Gratitude on Life Satisfaction." *Psychology Research and Behavior Management* Volume 16: 4799–816. <https://doi.org/10.2147/prbm.s372432>.
- Khattab, Mustafa, trans. n.d.-a. *The Clear Quran*. Quran.com. Accessed October 20, 2026. <https://quran.com/en/ibrahim/7>.
- Khattab, Mustafa, trans. n.d.-b. *The Clear Quran*. Quran.com. Accessed October 20, 2026. <https://quran.com/31/12>.
- Kobayashi, Kenji, and Ming Hsu. 2019. "Common Neural Code for Reward and Information Value." *Proceedings of the National Academy of Sciences* 116 (26): 13061–66. <https://doi.org/10.1073/pnas.1820145116>.
- Koehler, Elizabeth, Elizabeth Brown, and Sebastien J.-P. A. Haneuse. 2009. "On the Assessment of Monte Carlo Error in Simulation-Based Statistical Analyses." *The American Statistician* 63 (2): 155–62. <https://doi.org/10.1198/tast.2009.0030>.
- Koo, Minkyung, Sara B. Algoe, Timothy D. Wilson, and Daniel T. Gilbert. 2008. "It's a wonderful life: Mentally subtracting positive events improves people's affective states, contrary to their affective forecasts." *Journal of Personality and Social Psychology* 95 (5): 1217–24. <https://doi.org/10.1037/a0013316>.
- Kulp, Joshua, trans. n.d.-a. *Mishnah Berakhot*. Mishnah Yomit by Dr. Joshua Kulp, Sefaria. Accessed October 20, 2026. https://www.sefaria.org/Mishnah_Berakhot.9.5?lang=bi&ven=Mishnah%20Yomit%20by%20Dr.%20Joshua%20Kulp.
- Kulp, Joshua, trans. n.d.-b. *Pirkei Avot*. Mishnah Yomit by Dr. Joshua Kulp, Sefaria. Accessed October 20, 2026. https://www.sefaria.org/Pirkei_Avot.4.1?lang=bi&ven=Mishnah%20Yomit%20by%20Dr.%20Joshua%20Kulp.
- Lakens, Daniël. 2017. "Equivalence Tests: A Practical Primer for t Tests, Correlations, and Meta-Analyses." *Social Psychological and Personality Science* 8 (4): 355–62. <https://doi.org/10.1177/1948550617697177>.
- Laozi. 2008. *The Tao Teh King, or the Tao and Its Characteristics*. Ebook 216, HTML release 2008-07-12. Translated by James Legge. Project Gutenberg. <https://www.gutenberg.org/files/216/216-h/216-h.htm>.

- Laukkonen, Ruben E., and Heleen A. Slagter. 2021. "From Many to (n)one: Meditation and the Plasticity of the Predictive Mind." *Neuroscience & Biobehavioral Reviews* 128: 199–217. <https://doi.org/10.1016/j.neubiorev.2021.06.021>.
- Lieshout, Lieke L. F. van, Annelinde R. E. Vandenbroucke, Nils C. J. Müller, Roshan Cools, and Floris P. de Lange. 2018. "Induction and Relief of Curiosity Elicit Parietal and Frontal Activity." *Journal of Neuroscience* 38 (10): 2579–88. <https://doi.org/10.1523/JNEUROSCI.2816-17.2018>.
- Lutz, Antoine, Amishi P. Jha, John D. Dunne, and Clifford D. Saron. 2015. "Investigating the Phenomenological Matrix of Mindfulness-Related Practices from a Neurocognitive Perspective." *American Psychologist* 70 (7): 632–58. <https://doi.org/10.1037/a0039585>.
- Lutz, Antoine, Jean-Philippe Lachaux, Jacques Martinerie, and Francisco J. Varela. 2002. "Guiding the Study of Brain Dynamics by Using First-Person Data: Synchrony Patterns Correlate with Ongoing Conscious States During a Simple Visual Task." *Proceedings of the National Academy of Sciences* 99 (3): 1586–91. <https://doi.org/10.1073/pnas.032658199>.
- McCullough, Michael E., Shelley D. Kilpatrick, Robert A. Emmons, and David B. Larson. 2001. "Is gratitude a moral affect?" *Psychological Bulletin* 127 (2): 249–66. <https://doi.org/10.1037/0033-2909.127.2.249>.
- Mirza, M. Berk, Rick A. Adams, Karl Friston, and Thomas Parr. 2019. "Introducing a Bayesian model of selective attention based on active inference." *Scientific Reports* 9 (1). <https://doi.org/10.1038/s41598-019-50138-8>.
- Moore, Graham F., Suzanne Audrey, Mary Barker, et al. 2015. "Process evaluation of complex interventions: Medical Research Council guidance." *BMJ* 350: h1258. <https://doi.org/10.1136/bmj.h1258>.
- Morris, Tim P., Ian R. White, and Michael J. Crowther. 2019. "Using simulation studies to evaluate statistical methods." *Statistics in Medicine* 38 (11): 2074–102. <https://doi.org/10.1002/sim.8086>.
- Munch, Edvard. 1893. *The Scream*, NG.m.00939. Nasjonalmuseet. <https://www.nasjonalmuseet.no/en/collection/object/NG.M.00939>.
- Nagarjuna. 2020. *Mulamadhyamakakarika*. GRETEL, SUB Gottingen. https://gretel.sub.uni-goettingen.de/gretel/corpustei/transformations/html/sa_nAgArjuna-mUlamadhyamakakArikA.htm.
- Nelson, Thomas O., and Louis Narens. 1990. "Metamemory: A Theoretical Framework and New Findings." In *Psychology of Learning and Motivation*, vol. 26. [https://doi.org/10.1016/S0079-7421\(08\)60053-5](https://doi.org/10.1016/S0079-7421(08)60053-5).
- New Revised Standard Version Updated Edition*. 2021. Authorized text displayed by Bible Gateway. <https://www.biblegateway.com/passage/?search=1%20Thessalonians%205%3A18&version=NRSVUE>.
- Newman, David B., Amie M. Gordon, and Wendy Berry Mendes. 2021. "Comparing daily physiological and psychological benefits of gratitude and optimism using a digital platform." *Emotion* 21 (7): 1357–65. <https://doi.org/10.1037/emo0001025>.
- Palminteri, Stefano, Valentin Wyart, and Etienne Koechlin. 2017. "The Importance of Falsification in Computational Cognitive Modeling." *Trends in Cognitive Sciences* 21 (6): 425–33. <https://doi.org/10.1016/j.tics.2017.03.011>.
- Phipson, Belinda, and Gordon K Smyth. 2010. "Permutation P-values Should Never Be Zero: Calculating Exact P-values When Permutations Are Randomly Drawn." *Statistical Applications in Genetics and Molecular Biology* 9 (1). <https://doi.org/10.2202/1544-6115.1585>.
- Pseudo-Dionysius the Areopagite. n.d. *Dionysius the Areopagite: On the Divine Names and the Mystical Theology*. Translated by C. E. Rolt. Christian Classics Ethereal Library. Accessed October 20, 2026. <https://www.ccel.org/ccel/rold/dionysius.v.html>.
- Romano, Joseph P., and Michael Wolf. 2005. "Exact and Approximate Stepdown Methods for Multiple Hypothesis Testing." *Journal of the American Statistical Association* 100 (469): 94–108. <https://doi.org/10.1198/016214504000000539>.
- RUBIN, DONALD B. 1976. "Inference and missing data." *Biometrika* 63 (3): 581–92. <https://doi.org/10.1093/biomet/63.3.581>.
- Sajid, Noor, Philip J. Ball, Thomas Parr, and Karl J. Friston. 2021. "Active Inference: Demystified and Compared." *Neural Computation* 33 (3): 674–712. https://doi.org/10.1162/neco_a_01357.
- Saltelli, Andrea, and Paola Annoni. 2010. "How to avoid a perfunctory sensitivity analysis." *Environmental Modelling & Software* 25 (12): 1508–17. <https://doi.org/10.1016/j.envsoft.2010.04.012>.
- Sandved-Smith, Lars, Chris Fields, Thomas Doctor, Ruben Laukkonen, and Jakob Hohwy. 2026. *There is no self-evidence: A physics of emptiness realisation*. PsyArXiv preprint, version 1. https://osf.io/preprints/psyarxiv/m78z2_v1.
- Sandved-Smith, Lars, Casper Hesp, Jérémie Mattout, Karl Friston, Antoine Lutz, and Maxwell J. D. Ramstead. 2021. "Towards a Computational Phenomenology of Mental Action: Modelling Meta-Awareness and Attentional Control with Deep Parametric Active Inference." *Neuroscience of Consciousness* 2021 (1): niab018. <https://doi.org/10.1093/nc/niab018>.
- Satterthwaite, F. E. 1946. "An Approximate Distribution of Estimates of Variance Components." *Biometrics Bulletin* 2 (6): 110–14. <https://doi.org/10.2307/3002019>.

- Seaman, Shaun, John Galati, Dan Jackson, and John Carlin. 2013. “What Is Meant by ‘Missing at Random’?” *Statistical Science* 28 (2): 257–68. <https://doi.org/10.1214/13-STS415>.
- Seymour, Ben, John P O’Doherty, Martin Koltzenburg, et al. 2005. “Opponent appetitive-aversive neural processes underlie predictive learning of pain relief.” *Nature Neuroscience* 8 (9): 1234–40. <https://doi.org/10.1038/nn1527>.
- Shah, Punit, Adam J. L. Harris, Geoffrey Bird, Caroline Catmur, and Ulrike Hahn. 2016. “A pessimistic view of optimistic belief updating.” *Cognitive Psychology* 90: 71–127. <https://doi.org/10.1016/j.cogpsych.2016.05.004>.
- Siderits, Mark. 2026. “Buddha.” In *The Stanford Encyclopedia of Philosophy*, Summer 2026, edited by Edward N. Zalta and Uri Nodelman. Metaphysics Research Lab, Stanford University. <https://plato.stanford.edu/archives/sum2026/entries/buddha/>.
- Smith, Ryan, Karl J. Friston, and Christopher J. Whyte. 2022. “A step-by-step tutorial on active inference and its application to empirical data.” *Journal of Mathematical Psychology* 107: 102632. <https://doi.org/10.1016/j.jmp.2021.102632>.
- Swami Mukundananda. n.d.-a. *Bhagavad Gita: The Song of God*. Translated by Swami Mukundananda. Online translation and commentary. Accessed October 20, 2026. <https://www.holy-bhagavad-gita.org/chapter/2/verse/47/>.
- Swami Mukundananda. n.d.-b. *Bhagavad Gita: The Song of God*. Translated by Swami Mukundananda. Online translation and commentary. Accessed October 20, 2026. <https://www.holy-bhagavad-gita.org/chapter/3/verse/12/>.
- Tatler, B. W., M. M. Hayhoe, M. F. Land, and D. H. Ballard. 2011. “Eye Guidance in Natural Vision: Reinterpreting Saliency.” *Journal of Vision* 11 (5): 5, 1–23. <https://doi.org/10.1167/11.5.5>.
- Thanissaro Bhikkhu. 1997a. *Cula-suññata Sutta: The Lesser Discourse on Emptiness (MN 121)*. Translation from Pali, Access to Insight. <https://www.accesstoinsight.org/tipitaka/mn/mn.121.than.html>.
- Thanissaro Bhikkhu. 1997b. *Kaccayanagotta Sutta: To Kaccayana Gotta (on Right View), SN 12.15*. Access to Insight. <https://www.accesstoinsight.org/tipitaka/sn/sn12/sn12.015.than.html>.
- Thanissaro Bhikkhu. 1997c. *Suñña Sutta: Empty (SN 35.85)*. Access to Insight, BCBS edition. <https://www.accesstoinsight.org/tipitaka/sn/sn35/sn35.085.than.html>.
- Thanissaro Bhikkhu. 2002. *Kataññu Suttas: Gratitude (AN 2.31–32)*. Access to Insight, BCBS edition. <https://www.accesstoinsight.org/tipitaka/an/an02/an02.031.than.html>.
- Thanissaro Bhikkhu. 2004. *Alagaddupama Sutta: The Water-Snake Simile (MN 22)*. Access to Insight, BCBS edition. <https://www.accesstoinsight.org/tipitaka/mn/mn.022.than.html>.
- Thich Nhat Hanh. 2014. *The Insight That Brings Us to the Other Shore*. Translated by Thich Nhat Hanh. Plum Village. <https://plumvillage.org/about/thich-nhat-hanh/letters/thich-nhat-hanh-new-heart-sutra-translation>.
- Van Dam, Nicholas T., Marieke K. van Vugt, David R. Vago, et al. 2018. “Mind the Hype: A Critical Evaluation and Prescriptive Agenda for Research on Mindfulness and Meditation.” *Perspectives on Psychological Science* 13 (1): 36–61. <https://doi.org/10.1177/1745691617709589>.
- Võ, Melissa L.-H., Avigael M. Aizenman, and Jeremy M. Wolfe. 2016. “You Think You Know Where You Looked? You Better Look Again.” *Journal of Experimental Psychology: Human Perception and Performance* 42 (10): 1477–81. <https://doi.org/10.1037/xhp0000264>.
- WELCH, B. L. 1947. “THE GENERALIZATION OF ‘STUDENT’S’ PROBLEM WHEN SEVERAL DIFFERENT POPULATION VARIANCES ARE INVOLVED.” *Biometrika* 34 (1-2): 28–35. <https://doi.org/10.1093/biomet/34.1-2.28>.
- White, Ian R, Tra My Pham, Matteo Quartagno, and Tim P Morris. 2024. “How to check a simulation study.” *International Journal of Epidemiology* 53 (1). <https://doi.org/10.1093/ije/dyad134>.
- Wilson, Robert C., and Anne G. E. Collins. 2019. “Ten Simple Rules for the Computational Modeling of Behavioral Data.” *eLife* 8: e49547. <https://doi.org/10.7554/eLife.49547>.
- Wood, Alex M., Jeffrey J. Froh, and Adam W. A. Geraghty. 2010. “Gratitude and well-being: A review and theoretical integration.” *Clinical Psychology Review* 30 (7): 890–905. <https://doi.org/10.1016/j.cpr.2010.03.005>.
- Yarbus, Alfred L. 1967. “Eye Movements During Perception of Complex Objects.” In *Eye Movements and Vision*. Springer. https://doi.org/10.1007/978-1-4899-5379-7_8.
- Yarkoni, Tal. 2022. “The Generalizability Crisis.” *Behavioral and Brain Sciences* 45: e1. <https://doi.org/10.1017/S0140525X20001685>.
- Yu, Hongbo, Xiaoxue Gao, Yuanyuan Zhou, and Xiaolin Zhou. 2018. “Decomposing Gratitude: Representation and Integration of Cognitive Antecedents of Gratitude in the Brain.” *The Journal of Neuroscience* 38 (21): 4886–98. <https://doi.org/10.1523/jneurosci.2944-17.2018>.