

Two Routes to Interoceptive Prioritisation: Dissociating Likelihood Precision and Preference Reweighting in Active Inference

St John Grimble^{1,2*}, Nicolas Kuske^{1,2}, Evert A. Boonstra², Charel van Hoof⁶,
Rowan Hodson³, Benjamin Rosman⁴, Ryan Smith³, Mark Solms^{2**}, and
Jonathan P. Shock^{1,2,5**}

¹ Dept. of Mathematics & Applied Mathematics, Univ. of Cape Town, South Africa

² Neuroscience Institute, Univ. of Cape Town, South Africa

³ Laureate Institute for Brain Research, Tulsa, OK, USA

⁴ MIND Institute and CSAM, University of the Witwatersrand, South Africa

⁵ INRS, Montreal, Canada

⁶ Delft University of Technology, Dept. of Cognitive Robotics

Abstract. An embodied agent must prioritise among competing bodily needs. In active inference, it can respond to an urgent need by increasing the precision of information about that need or by increasing its preference for outcomes that relieve it. We compare these two interventions, separately and together, in a foraging gridworld with three lethal needs, and ask how each changes survival, perception and the body model the agent learns. Precision sharpens interoceptive posteriors and improves survival. Preference reweighting keeps bodily states further from their lethal boundaries and changes the states from which the agent learns, so its learned model covers fewer need levels, with lower quality across all levels but slightly higher quality on the levels it visits. An experiment on 72 further layouts does not establish the predicted preference benefit on difficult layouts, but both interventions improve average survival there and their interaction is positive. Survival alone therefore gives an incomplete account of how a model prioritises needs and learns about its body.

Keywords: active inference · interoception · likelihood precision · prior preferences · expected free energy · Dirichlet learning · homeostasis

1 Introduction

A multi-need embodied agent cannot service every drive at once. Hunger, thirst and the need to breathe can each demand a different action, and survival depends on which takes priority. The agent must also infer how urgent these needs are from its interoceptive signals. Active inference allows us to model this relationship between inference over bodily states and homeostatic control [22,2,23,1].

* Corresponding author: uct@stjohngrimble.com

** Co-senior authors.

But responding to an urgent need can involve different mechanisms. We ask what difference the choice of mechanism makes to behaviour and learning.

We compare two approaches to prioritising a bodily need. An agent that believes itself thirsty could increase the precision of its thirst signal, making information about thirst more reliable at the expense of other needs. Alternatively, it could increase its preference for water, so that policies leading to water become more attractive. The first approach is *likelihood-precision routing*, related to accounts of attention as precision-weighted sensory evidence [7,16,3] and to context-dependent likelihood precision in discrete-state visual attention [14]. The second is *preference reweighting*, drawing on motivational and homeostatic accounts of active inference [18,19,27].

These are not two implementations of a single mechanism. Both alter policy selection through expected free energy, but they need not change learning in the same way. Our precision intervention changes both the observations emitted by the body and the likelihood used for inference and planning. Preference reweighting changes which outcomes the agent seeks. Although preferences do not enter the likelihood update directly, the actions they favour change the agent’s experience and hence what it learns. Comparing the learned body model therefore tells us something that a survival comparison alone cannot.

The choice also matters for biological interpretation. Precision points to gain control on interoceptive afferents [21,22], while preference points to changes in regulated outcomes [25,18]. That the same behaviour can be produced by distinct underlying mechanisms is the familiar problem of degeneracy in biological systems [5,20]. A model that reproduces an animal’s behaviour does not fix which mechanism the animal uses.

We test each intervention separately and together in the same foraging agent. To our knowledge, these two ways of responding to interoceptive need have not previously been varied together in the same task while measuring their effects on the learned body model. We relate survival to posterior uncertainty, bodily regulation and the range of states represented in that model. We then test the most uncertain of these findings, the dependence of the preference effect on layout difficulty, on an independently selected set of layouts. The question is not only whether the agent survives, but how its way of prioritising needs changes what it perceives, does and learns.

2 Related Work

Smith et al. [24] dissociate interoceptive sensory precision from a bodily-state prior in a cardiac-perception task. We adopt the same strategy of varying distinct model parameters, here comparing likelihood precision with preferences. The distinction matters because a state prior expresses a belief about the body, whereas a preference specifies observations the agent *wants*.

Precision-as-attention accounts describe attention through the precision of sensory evidence or prediction errors [7,16,3]. In a discrete-state visual-search model, context-dependent likelihood precision has been used to model selec-

tive attention [14]. The motivational and homeostatic literature encodes desirable bodily or contextual outcomes through prior preferences [18,19,27]. Related homeostatic reinforcement learning defines reward as reduced physiological deficit [12].

Active inference also admits other ways to change behaviour. Hesp et al. [10] derive valence from the dynamics of policy precision, rather than the likelihood precision studied here. Other accounts couple preferences and attentional precision [19,17]. Our comparison makes the two interventions separately controllable within the discrete active-inference framework synthesised by Da Costa et al. [4].

Selective precision allocation is studied in our prior model [9], while Torresan et al. [26] compare forms of preference without varying likelihood precision. Neacsu et al. [15] show that weakening prior preferences can alter likelihood learning. We build on this indirect relationship by comparing how precision and preference, each responding to interoceptive need, change the experience from which the agent learns.

3 Methods

3.1 A foraging task with three lethal needs in tension

We study interoceptive prioritisation in a minimal model of embodied homeostatic regulation. A *body* has internal variables to maintain, an *environment* holds resources to replenish them, and an active-inference *agent* senses the body through interoception and acts to keep it alive. The body couples to the environment through the agent, so survival is a closed perception–action loop.

The agent forages in *AffectWorld*, a 6×6 gridworld with hunger, thirst and oxygen. Hunger and thirst fall by one level per step. Eating restores hunger, while drinking restores one third of the thirst range. Water submerges the agent and lowers oxygen by two levels per step, whereas on land oxygen recovers by one

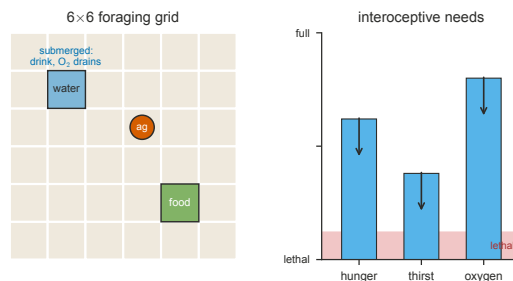


Fig. 1: **The foraging task** (illustrative, with layouts varying): one food and one water tile on a 6×6 grid. The needs panel (right) tracks three interoceptive levels and their lethal floors. Hunger and thirst decline over time. Drinking restores thirst but lowers oxygen, which recovers only on land. The best priority therefore depends on the current body state.

level per step. Food and water tiles remain fixed and reusable, so the resource map does not deplete or change during these experiments. The agent dies when any need crosses its lethal boundary. Drinking can therefore relieve thirst while endangering oxygen, which places all three needs in tension (Fig. 1).⁷

3.2 An active-inference forager that learns its body model online

We model the task as a discrete partially observable Markov decision process (POMDP) and use active inference to infer states and select actions [8,4]. The POMDP’s state factors are grid position, a context factor held fixed in these experiments, and the true level of each of the three needs. Position is observed exactly. The current tile’s resource cue is observed with 0.95 accuracy, and each need produces a noisy ordinal interoceptive observation. The agent therefore infers only the need levels. Actions combine a move with an optional interaction. Movement, body dynamics and the fixed resource map are given to the agent. Only the interoceptive observation likelihood is learned online.

Writing hidden states as s , observations as o and policies as π , the generative model factorises over a horizon as

$$P(o_{1:T}, s_{1:T}, \pi) = P(\pi) P(s_1) \prod_{t=1}^T P(o_t | s_t) \prod_{t=2}^T P(s_t | s_{t-1}, \pi). \quad (1)$$

Here the categorical likelihood is $P(o_t | s_t) = \text{Cat}(A)$, transitions are $P(s_t | s_{t-1}, \pi) = \text{Cat}(B)$, C contains raw log-preference scores over observations, and D is the initial-state prior. The corresponding preference distribution is $P_C = \sigma(C)$. The interoceptive likelihood factorises over the $M = 3$ channels, $A = \{A^{(m)}\}$. The agent infers a posterior $q(s)$ by fixed-point iteration and scores each policy by its expected free energy,

$$G(\pi) = \underbrace{D_{\text{KL}}[Q(o | \pi) \| P_C(o)]}_{\text{risk}} + \underbrace{\mathbb{E}_{Q(s|\pi)}[H[P(o | s)]]}_{\text{ambiguity}}, \quad (2)$$

where $Q(o | \pi) = \sum_s A Q(s | \pi)$ and actions follow $P(\pi) = \sigma(-\gamma G(\pi))$. Up to the normalisation constant in P_C , $-G(\pi)$ equals $\mathbb{E}_{Q(o|\pi)}[C(o)] + I_Q(s; o | \pi)$, the expected utility plus the state information gain. Thus C enters expected utility, while A enters predicted outcomes and the state information gain $I_Q(s; o | \pi) = H[Q(o | \pi)] - \mathbb{E}_{Q(s|\pi)} H[P(o | s)]$. All four conditions keep this state information gain and omit parameter information gain over the learned Dirichlet counts, so no agent chooses actions in order to improve its body model. What differs between conditions is the experience that their policies provide for learning.

Planning enumerates policies to a horizon of three steps. The agent accumulates Dirichlet counts $\mathbf{a}$ over A and updates them after each observation [4],

$$\mathbf{a}^{(m)} += \eta (\mathbf{o}^{(m)} \otimes q(s^{(m)})), \quad (3)$$

⁷ The model studied here has three interoceptive channels ($M = 3$). It descends from our prior model [9], which has four channels. The fourth, a sleep/fatigue channel, and its tiles are removed from the model rather than disabled.

where $\mathbf{o}^{(m)}$ is the one-hot observation on channel m , $q(s^{(m)})$ is its state posterior, η is the learning rate and $\otimes$ is the outer product. Each observation therefore updates the model of that interoceptive channel, weighted by the inferred body state. Code, configurations and the supplementary material, which gives the constants, procedures and additional analyses, are available at <https://github.com/sgrimby/interoceptive-prioritisation-iwai2026-snapshot>.

3.3 Likelihood precision and preference reweighting

Both interventions respond to the agent’s posterior estimates of need. Precision routing gives more reliable information about the need with the largest expected deficit. Preference reweighting instead uses the estimated hunger and thirst deficits to set preferences for food and water (Fig. 2).

Likelihood precision (κ -routing). A scalar precision κ on the diagonal of the interoceptive likelihood $A^{(m)}$ controls how reliably a true need level is reported. The same allocation changes the body’s emitted observations and the likelihood used for inference and planning. At higher κ the observation model places more mass on the correct ordinal level, and at lower κ it spreads that mass over the alternatives. Thus, when thirst is selected, the agent receives more reliable evidence about thirst and less reliable evidence about hunger and oxygen. Here κ is the diagonal probability mass of the likelihood, not an inverse-variance precision. The total precision across channels is fixed at $K = M \cdot \bar{\kappa}$, so prioritising one channel necessarily de-prioritises the others. Because κ scales A , it affects state inference, policy evaluation and the Dirichlet update in Eq. (3). This is the precision-routing mechanism of our prior model [9].

Preference reweighting (dynamic- C). This intervention rewrites the raw resource preference scores C from the posterior deficits and does not directly alter A . Greater expected thirst makes water more desirable, and greater expected hunger does the same for food. If either deficit crosses the critical threshold, the alternative resource and empty tiles receive a large penalty. Because C appears only in the pragmatic part of the policy score, this intervention changes action selection without entering the likelihood update in Eq. (3).

Oxygen exposes an asymmetry between the interventions. It recovers passively on land and has no navigable resource cue, so preference reweighting does not target it directly. It can protect oxygen only by favouring policies that leave or avoid water. Precision routing can instead assign precision to the oxygen channel itself. The survival comparison therefore concerns the two interventions as implemented in a mixture of actionable and passive needs. It does not isolate their performance under symmetric access to every channel.

We compare four conditions using the same agent architecture. The uniform baseline keeps κ uniform and C static. The κ -routing condition changes precision alone, dynamic- C changes preferences alone, and the combined condition applies both interventions.

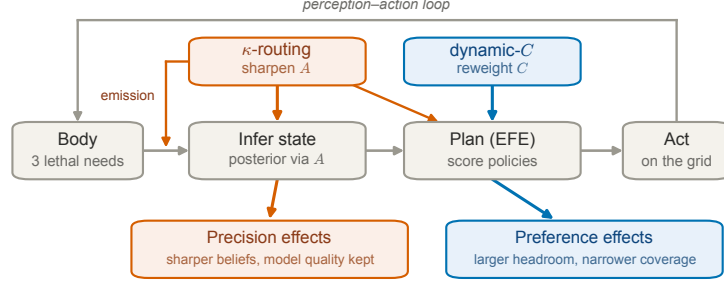


Fig. 2: **The perception–action loop and the two implemented mechanisms.** Precision (κ) sharpens the likelihood A , which feeds inference and planning and governs the body’s emission. Preference reweighting changes C , which enters only the planner’s pragmatic term. The lower boxes show principal internal effects on the original layouts, not effects exclusive to either intervention.

3.4 What changes in behaviour and learning?

Survival tells us whether prioritisation helps the agent stay alive. To understand how it does so, we also examine the agent’s uncertainty, bodily states and learned likelihood. In particular, preference can alter learning by changing the states the agent visits, even though it does not appear explicitly in Eq. (3). We distinguish the quality of the learned model across all need levels from its quality on visited levels, and measure the fraction of levels visited. This tests whether a lower overall score accompanies poorer learning on visited levels or narrower experience.

The *survival fraction* is the proportion of the 120-step episode completed. Comparisons pair the same seed and layout across conditions.

On the perceptual side, *posterior entropy* $\bar{H} = \langle -\sum_s q(s^{(m)}) \ln q(s^{(m)}) \rangle$ measures uncertainty about the need levels, averaged over channels and surviving steps in the second half of the episodes. Lower entropy means a sharper posterior. *Learned-model quality* is the probability that the learned likelihood assigns to the observation favoured by the true model, averaged over latent levels, $p_\star^{(m)} = \langle \hat{A}^{(m)}[\arg \max_o A_{o,s}^{(m)}, s] \rangle_s$. We report this both across all levels and across levels the agent visits.

On the behavioural side, *visited-state entropy* $H[p_{\text{visit}}]$ measures the breadth of body states occupied. The *minimum headroom* $b = \langle \min_m s_t^{(m)} / L_m \rangle_t$ is the mean distance of the most endangered need from its lethal boundary, expressed as a fraction of that channel’s range. It is calculated from true body levels, so sharper perception cannot change the measurement itself.

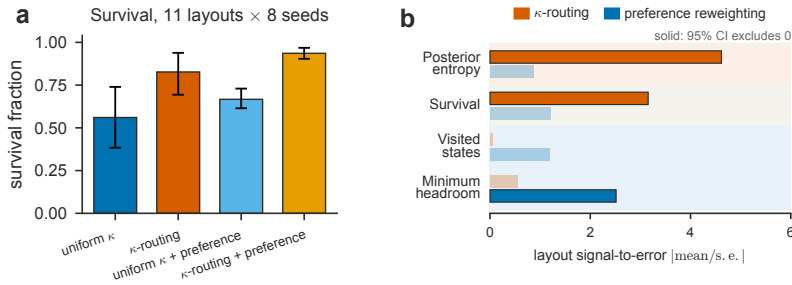


Fig. 3: **Survival and the internal dissociation.** (a) Condition means and layout-bootstrap 95% intervals across eleven layouts and eight seeds. (b) Layout-level signal-to-error (|mean/s.e.|). Bar length is descriptive. Solid bars have 95% intervals excluding zero (Tables 1, 3).

3.5 Experiments and analysis

The original experiment compares the four conditions on eleven distinct layouts with eight paired seeds. Each run comprises 20 learning episodes of 120 steps, and the learned-model comparison uses the same layouts and seeds over a longer run. These layouts were selected in an earlier task to distinguish controllers that respond differently to urgency, so we measured their difficulty in the present task from baseline survival rather than inheriting labels from that selection.

Intervals resample layouts while preserving the paired comparisons [6]. For the three original survival contrasts, we use exact two-sided sign-flip tests with Holm correction [11]. The internal-measurement intervals are not adjusted for multiple comparisons. With only eleven independent layouts, the original intervals are wide.

We also test whether the preference benefit depends on difficulty, as suggested by a post-hoc split of the original layouts. This experiment uses 72 additional layouts selected without inspecting intervention outcomes. Separate baseline runs assign them to three difficulty groups of 24, and evaluation uses eight new paired seeds. The primary test asks whether preference improves survival on the difficult layouts, averaging its effect with and without precision routing. The continuous relationship with difficulty and the effects across all 72 layouts are secondary analyses. The design and analysis were fixed before evaluation outcomes were inspected.

4 Results

4.1 How does each intervention change survival?

With the uniform baseline, the agent survives 0.56 of the episode, or about 67 of 120 steps. κ -routing raises survival to 0.83, and the combined condition reaches 0.94. Preference reweighting alone reaches 0.67 (Fig. 3a and Table 1).

Table 1: Survival contrasts on the eleven original layouts with eight paired seeds. Intervals are paired layout bootstraps. The p values use exact layout-level sign flips with Holm adjustment across the three contrasts.

Contrast	Δ survival	95% CI	Holm p
κ -routing vs uniform baseline	+0.27	[0.12, 0.44]	0.0176
dynamic- C vs uniform baseline	+0.11	[-0.06, 0.27]	0.248
κ -routing+dynamic- C vs uniform baseline	+0.38	[0.18, 0.56]	0.0176

The direction of the precision result agrees with our prior model [9]. Precision and the combined intervention improve survival after correction for multiple comparisons. The preference-alone estimate is less certain. Its interval includes both a decrease and an increase in survival, and an equivalence test is also inconclusive.

In a post-hoc split of the original layouts, preference alone improved survival on the five where the baseline struggled and reduced it on the other six. This suggested that its benefit might depend on difficulty. We tested that possibility on the additional layouts, this time averaging the preference effect across both precision conditions rather than testing preference alone.

Does the difficulty pattern generalise? The predicted preference benefit on the layouts calibrated as difficult was not established (Table 2). The primary estimate is close to zero, with an interval spanning both signs.

The continuous analysis also does not establish the predicted decrease in benefit with increasing baseline survival. The difficulty ranking is limited, however, by 51 tied baseline-calibration values.

Across all 72 layouts, the planned secondary analyses show positive average survival effects for both interventions and a positive interaction (Table 2). These results support complementarity on the survival-fraction scale.

On these additional layouts, the internal measurements show the same dominant pattern as on the original layouts. κ -routing lowers posterior entropy, and dynamic- C raises headroom and lowers conflict, although each also changes measurements associated with the other intervention.

4.2 Precision sharpens perception, preference changes the learned model

On the original layouts, precision sharpens the interoceptive posterior (Table 3, Fig. 3b). The preference contrast is smaller and its interval includes zero, so these data do not resolve a change in posterior entropy.

The learned-model comparison reveals why it matters which states an agent visits. Preference reweighting lowers model quality when averaged over all need levels, yet slightly improves quality on the levels actually visited. It also visits fewer levels. Those it rarely visits remain close to the vague prior and lower the

Table 2: Survival effects in the additional experiment. The primary test uses the 24 layouts calibrated as difficult. Secondary effects average over all 72 layouts. Intervals resample layouts. Secondary p values are Holm-adjusted jointly with the continuous difficulty test.

Comparison	Effect	95% CI	p
Preference, difficult layouts (primary)	-0.008	[-0.025, +0.016]	0.5793
Precision, all layouts (secondary)	+0.111	[+0.076, +0.147]	< 0.0001
Preference, all layouts (secondary)	+0.080	[+0.046, +0.118]	< 0.0001
Interaction, all layouts (secondary)	+0.114	[+0.049, +0.182]	0.002

Table 3: Internal contrasts against baseline, shown as mean [layout-bootstrap 95% CI]. Entropies use 10^{-3} nats (posterior) and nats (visited state). Headroom is a fraction.

Metric (vs. baseline)	κ -routing	dynamic- C	κ -routing+dynamic- C
Posterior entropy (10^{-3})	-1.37 [-1.95, -0.84]	-0.26 [-0.83, 0.28]	-1.18 [-1.77, -0.58]
Visited-state entropy	+0.00 [-0.09, 0.07]	-0.05 [-0.14, 0.02]	-0.15 [-0.24, -0.08]
Minimum headroom	+0.01 [-0.02, 0.04]	+0.05 [0.01, 0.09]	+0.11 [0.07, 0.15]

all-level average. In other words, the lower score accompanies narrower experience rather than poorer performance on the states encountered. Precision shows no comparable drop in all-level quality in the same learning runs.

4.3 Preference keeps needs further from their lethal boundaries

Preference increases minimum headroom on the original layouts by +0.05, and the combined condition by +0.11 (Table 3). The precision contrast is +0.01, with an interval that includes zero. Both interventions also reduce the time spent with two needs simultaneously near their lethal boundaries, from about 2% of baseline steps to less than 0.1%, a comparison we report without an interval. They reach this low-conflict state in different ways. Precision sharpens the evidence about need, while preference keeps the most endangered need further from its boundary.

4.4 Robustness on the original layouts

Replacing posterior deficits with true body levels leaves preference survival unchanged on every original layout. Sweeping the critical penalty over two orders of magnitude and removing its threshold also preserves the easy-layout cost, as do held-out seeds. The original pattern is therefore not explained by posterior error or one penalty setting, although it did not generalise to the additional layouts.

A belief-only κ variant keeps emitted observation precision uniform while routing precision during inference and planning. On six of the original layouts it

improves survival by +0.33. The full-intervention contrast of +0.27 in Table 1 uses the original eleven layouts, so these estimates are not a paired decomposition of the emission effect. The subset result suggests that sharper emission is not required for a benefit, while the primary intervention remains composite.

5 Discussion

The choice of prioritisation mechanism changes more than survival. Precision alters the evidence the agent receives and how it uses that evidence in inference and planning, while preference changes which outcomes the agent seeks. Both also change where the agent goes, and so what it later observes, believes and learns. On the additional layouts preference raised posterior entropy and precision slightly lowered headroom, so neither intervention is confined to perception or to action.

Why does preference change the learned likelihood when it never enters the learning rule? The answer is the agent’s experience. An agent that commits to relieving one need visits a narrower range of bodily states, and the levels it rarely visits stay close to the prior. It learns those states no less accurately, only less often. This is an interoceptive instance of the indirect coupling that Neacsu et al. [15] describe, and it is why an average over the whole likelihood says little about what the agent has learned from the states it actually encountered.

On the original layouts, preference helped where the baseline struggled and cost survival where the baseline already coped. Reweighting appears to overcommit the agent to one need, which is useful when the agent is failing and harmful when it is not. The additional layouts did not reproduce this pattern. Preference did not help on the layouts calibrated as difficult, no continuous relation with baseline survival was detected, and the unadjusted tertile contrasts run in the opposite direction, with the largest benefit on the easiest layouts. We therefore keep the overcommitment account for the original layouts only. What the additional layouts do show is that both interventions improve survival on average, and that the two together improve it by more than the sum of their separate effects. That is evidence of complementarity in survival, though not of how the two mechanisms combine.

This matters for how active inference is used to model biological systems. Fitting behaviour does not tell us whether prioritisation works through interoceptive gain or through preferred outcomes, because both can produce the same survival. The internal measurements suggest what would. Uncertainty or gain, distance from physiological limits and the range of states experienced can all be measured alongside behaviour. We compare two implemented mechanisms rather than testing formal parameter identifiability, and the same ambiguity applies to homeostatic reinforcement learners [12].

The original experiment uses eleven selected layouts and its difficulty split is post hoc. The additional experiment separates calibration from evaluation but is still one task. Preference also has no direct target for oxygen, which recovers passively. A symmetric comparison would need actionable oxygen or preferences

over body outcomes, and the joint hunger–thirst gate of the preference rule was never triggered in these runs.

Precision could also act on transition beliefs or on policy selection [13]. Varying emitted and believed noise separately would show which part of the present precision intervention carries its effect. More channels, richer environments and longer planning horizons would show how far the results extend beyond this gridworld and its three-step planner.

6 Conclusion

Likelihood precision and preference reweighting provide distinct mechanisms for interoceptive prioritisation, with different dominant effects on posterior uncertainty, bodily regulation and state coverage. The predicted preference benefit on independently calibrated difficult layouts was not established. Across all 72 layouts, the secondary analyses show positive average survival effects for both interventions and a positive interaction. Their internal observables reveal effects on later inference and learning that survival alone does not resolve. Comparing these observables is therefore important when using behavioural success to evaluate a model of prioritisation.

Declarations

Acknowledgements Funded by the Oppenheimer Memorial Trust (UCT Neuroscience Institute, 475201 NSI1006) and Conscium, Ltd (UCT Psychology, PSY526 428430). EAB: NWO grant 019.223SG.002. RS: Laureate Institute for Brain Research.

Use of AI tools AI-based tools assisted with writing, code development and research planning. The corresponding author verified all results, claims and code.

Data and code availability The project repository at <https://github.com/sgrimby/interoceptive-prioritisation-iwai2026-snapshot> contains the scoped data, code and exact environment needed to reproduce the reported experiments and analyses, together with the supplementary material.

Disclosure of Interests The authors have no competing interests to declare.

References

1. Allen, M., Tsakiris, M.: The body as first prior: interoceptive predictive processing and the primacy of self-models. In: Tsakiris, M., De Preester, H. (eds.) *The Interoceptive Mind: From Homeostasis to Awareness*, pp. 27–45. Oxford University Press (2018). <https://doi.org/10.1093/oso/9780198811930.003.0002>
2. Barrett, L.F., Simmons, W.K.: Interoceptive predictions in the brain. *Nature Reviews Neuroscience* **16**(7), 419–429 (2015). <https://doi.org/10.1038/nrn3950>
3. Bastos, A.M., Usrey, W.M., Adams, R.A., Mangun, G.R., Fries, P., Friston, K.J.: Canonical microcircuits for predictive coding. *Neuron* **76**(4), 695–711 (2012). <https://doi.org/10.1016/j.neuron.2012.10.038>
4. Da Costa, L., Parr, T., Sajid, N., Veselic, S., Neacsu, V., Friston, K.J.: Active inference on discrete state-spaces: A synthesis. *Journal of Mathematical Psychology* **99**, 102447 (2020). <https://doi.org/10.1016/j.jmp.2020.102447>
5. Edelman, G.M., Gally, J.A.: Degeneracy and complexity in biological systems. *Proceedings of the National Academy of Sciences* **98**(24), 13763–13768 (2001). <https://doi.org/10.1073/pnas.231499798>
6. Efron, B., Tibshirani, R.J.: *An Introduction to the Bootstrap*. Chapman & Hall, New York (1993)
7. Feldman, H., Friston, K.J.: Attention, uncertainty, and free-energy. *Frontiers in Human Neuroscience* **4**, 215 (2010). <https://doi.org/10.3389/fnhum.2010.00215>
8. Friston, K., FitzGerald, T., Rigoli, F., Schwartenbeck, P., Pezzulo, G.: Active inference: a process theory. *Neural Computation* **29**(1), 1–49 (2017). https://doi.org/10.1162/NECO_a_00912
9. Grimbly, S.J., Kuske, N., Boonstra, E.A., Bassett, B.A., van Hoof, C., Hodson, R., Rosman, B., Smith, R., Solms, M., Shock, J.P.: Interoceptive attention as dynamic homeostatic prioritization in a foraging agent (2026), arXiv:2608.04232. Accepted at SAB 2026 (From Animals to Animats 18)
10. Hesp, C., Smith, R., Parr, T., Allen, M., Friston, K.J., Ramstead, M.J.D.: Deeply felt affect: The emergence of valence in deep active inference. *Neural Computation* **33**(2), 398–446 (2021). https://doi.org/10.1162/neco_a_01341
11. Holm, S.: A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics* **6**(2), 65–70 (1979)
12. Keramati, M., Gutkin, B.: Homeostatic reinforcement learning for integrating reward collection and physiological stability. *eLife* **3**, e04811 (2014). <https://doi.org/10.7554/eLife.04811>
13. Limanowski, J., Adams, R.A., Kilner, J., Parr, T.: The many roles of precision in action. *Entropy* **26**(9), 790 (2024). <https://doi.org/10.3390/e26090790>
14. Mirza, M.B., Adams, R.A., Friston, K.J., Parr, T.: Introducing a Bayesian model of selective attention based on active inference. *Scientific Reports* **9**, 13915 (2019). <https://doi.org/10.1038/s41598-019-50138-8>
15. Neacsu, V., Mirza, M.B., Adams, R.A., Friston, K.J.: Structure learning enhances concept formation in synthetic active inference agents. *PLOS ONE* **17**(11), e0277199 (2022). <https://doi.org/10.1371/journal.pone.0277199>
16. Parr, T., Friston, K.J.: Uncertainty, epistemics and active inference. *Journal of the Royal Society Interface* **14**(136), 20170376 (2017). <https://doi.org/10.1098/rsif.2017.0376>
17. Parvizi-Wayne, D.: How preferences enslave attention: calling into question the endogenous/exogenous dichotomy from an active inference perspective. *Phenomenology and the Cognitive Sciences* (2024). <https://doi.org/10.1007/s11097-024-10028-5>

18. Pezzulo, G., Rigoli, F., Friston, K.J.: Active inference, homeostatic regulation and adaptive behavioural control. *Progress in Neurobiology* **134**, 17–35 (2015). <https://doi.org/10.1016/j.pneurobio.2015.09.001>
19. Pezzulo, G., Rigoli, F., Friston, K.J.: Hierarchical active inference: A theory of motivated control. *Trends in Cognitive Sciences* **22**(4), 294–306 (2018). <https://doi.org/10.1016/j.tics.2018.01.009>
20. Prinz, A.A., Bucher, D., Marder, E.: Similar network activity from disparate circuit parameters. *Nature Neuroscience* **7**(12), 1345–1352 (2004). <https://doi.org/10.1038/nn1352>
21. Quattrocki, E., Friston, K.J.: Autism, oxytocin and interoception. *Neuroscience and Biobehavioral Reviews* **47**, 410–430 (2014). <https://doi.org/10.1016/j.neubiorev.2014.09.012>
22. Seth, A.K.: Interoceptive inference, emotion, and the embodied self. *Trends in Cognitive Sciences* **17**(11), 565–573 (2013). <https://doi.org/10.1016/j.tics.2013.09.007>
23. Seth, A.K., Friston, K.J.: Active interoceptive inference and the emotional brain. *Philosophical Transactions of the Royal Society B: Biological Sciences* **371**(1708), 20160007 (2016). <https://doi.org/10.1098/rstb.2016.0007>
24. Smith, R., Kuplicki, R., Feinstein, J., Forthman, K.L., Stewart, J.L., Paulus, M.P., Tulsa 1000 investigators, Khalsa, S.S.: A Bayesian computational model reveals a failure to adapt interoceptive precision estimates across depression, anxiety, eating, and substance use disorders. *PLoS Computational Biology* **16**(12), e1008484 (2020). <https://doi.org/10.1371/journal.pcbi.1008484>
25. Sterling, P.: Allostasis: a model of predictive regulation. *Physiology & Behavior* **106**(1), 5–15 (2012). <https://doi.org/10.1016/j.physbeh.2011.06.004>
26. Torresan, F., Kanai, R., Baltieri, M.: Prior preferences in active inference agents: soft, hard, and goal shaping (2025), arXiv:2512.03293
27. Tschantz, A., Barca, L., Maisto, D., Buckley, C.L., Seth, A.K., Pezzulo, G.: Simulating homeostatic, allostatic and goal-directed forms of interoceptive control using active inference. *Biological Psychology* **169**, 108266 (2022). <https://doi.org/10.1016/j.biopsycho.2022.108266>

Two Routes to Interoceptive Prioritisation: Dissociating Likelihood Precision and Preference Reweighting in Active Inference Supplementary Material

St John Grimby^{1,2*}, Nicolas Kuske^{1,2}, Evert A. Boonstra², Charel van Hoof⁶,
Rowan Hodson³, Benjamin Rosman⁴, Ryan Smith³, Mark Solms^{2**}, and
Jonathan P. Shock^{1,2,5**}

¹ Dept. of Mathematics & Applied Mathematics, Univ. of Cape Town, South Africa

² Neuroscience Institute, Univ. of Cape Town, South Africa

³ Laureate Institute for Brain Research, Tulsa, OK, USA

⁴ MIND Institute and CSAM, University of the Witwatersrand, South Africa

⁵ INRS, Montreal, Canada

⁶ Delft University of Technology, Dept. of Cognitive Robotics

A Model and constants

This section states the parts of the model that the main paper summarises.

Interoceptive likelihood. Each channel m has a latent level $s \in \{0, \dots, L_m\}$, with $L_m = 10$ for hunger and thirst and $L_m = 6$ for oxygen. The agent observes $n = 6$ ordinal need levels $\beta_m(s) = \text{round}((n-1)(L_m-s)/L_m)$, and the likelihood places mass κ_m on the reported level,

$$A_{o,s}^{(m)} = \begin{cases} \kappa_m & o = \beta_m(s), \\ (1 - \kappa_m)/(n-1) & o \neq \beta_m(s). \end{cases} \quad (\text{S1})$$

Under the budget $K = M\bar{\kappa} = 1.95$ the attended channel receives $\kappa = 0.9$ and the other two 0.525 each. The uniform allocation gives 0.65 to every channel.

Precision selector. The attended channel is the one the agent believes closest to its lethal boundary, with ties broken by the lowest index,

$$m^* = \arg \max_m \mathbb{E}_{Q(s(m))} [(L_m - s)/L_m]. \quad (\text{S2})$$

Preference reweighting. At each step dynamic- C sets the preferences for food and water from the believed deficits $t_m = \max(0, \bar{L} - \mathbb{E}_{Q(s(m))}[s])$, with a fixed

* Corresponding author: uct@stjohngrimby.com

** Co-senior authors.

Table S1: Constants, identical across the four conditions.

Quantity	Symbol	Value
Interoceptive channels	M	3 (hunger, thirst, oxygen)
Ordinal levels per channel	n	6
Base precision	$\bar{\kappa}$	0.65
Attended precision cap	κ_{att}	0.90
Precision budget	$K = M\bar{\kappa}$	1.95
Background floor	κ_{min}	0.05
Deficit reference	$\bar{L}$	15
Critical threshold	τ	14
Preference critical penalty	P_{crit}	-500
Policy precision	γ	1 (pymdp default)
Dirichlet learning rate	η	1.0
Dirichlet concentration	α_0	0.1
Planning horizon	H	3 (full enumeration)
Grid	—	6×6 , one food and one water tile
Original layouts	—	12 IDs (11 layouts), 8 seeds; 20×120
Learned-model run	—	12 IDs (11 layouts), 8 seeds; 30×150
Additional layouts	—	72 layouts; calibration 4 seeds, evaluation 8 seeds; 20×120
Resamples	—	10^4 bootstraps (original), 2×10^4 boot- straps and 10^5 permutations (additional)

reference $\bar{L} = 15$ for both hunger and thirst. Because $\bar{L}$ exceeds their ceiling of 10, each deficit lies between 5 and 15:

$$C^{\text{res}}[\text{food}] = \begin{cases} P_{\text{crit}} & t_{\text{thirst}} > \tau, \\ t_{\text{hunger}} & \text{otherwise,} \end{cases} \quad C^{\text{res}}[\text{water}] = \begin{cases} P_{\text{crit}} & t_{\text{hunger}} > \tau, \\ t_{\text{thirst}} & \text{otherwise.} \end{cases} \quad (\text{S3})$$

A deficit exceeds $\tau = 14$ only when the expected level is below 1. The other outcome slots receive P_{crit} when either deficit is critical and an empty bias of -1 otherwise, and poison keeps -4 . Oxygen has no navigable resource and is not reweighted. C does not enter the likelihood or its Dirichlet update.

Original layouts. The food and water positions were selected in an earlier task without lethal oxygen, to separate an urgency-following controller from one that ignores urgency. The IDs `layout0001` and `layout0068` share coordinates and are averaged before inference, giving 88 paired layout–seed records over eleven layouts. In a post-hoc median split, preference alone improves survival by $+0.36$ on the five layouts with lower baseline survival and reduces it by -0.11 on the other six. Its overall survival contrast has a 90% interval of $[-0.03, +0.24]$, which does not establish equivalence within ± 0.06 .

B Additional layouts

The 72 layouts are the first eligible placements from a fixed seed stream that are distinct under rotation and reflection of the grid and from the original layouts. They were chosen without reference to any agent outcome. Baseline-only runs on seeds 200–203 ranked them into 24 hard, 24 middle and 24 easy layouts. Fifty-one layouts share the modal calibration survival of 0.083, and travel burden, the largest start-to-resource distance and a fixed geometry signature break those ties. The four conditions were then run on seeds 300–307.

For layout g , let $\bar{Y}_g(k, c)$ be survival averaged over the eight seeds, where k and c indicate precision routing and preference reweighting. The marginal preference effect is

$$d_g^C = \frac{1}{2} \{ [\bar{Y}_g(0, 1) - \bar{Y}_g(0, 0)] + [\bar{Y}_g(1, 1) - \bar{Y}_g(1, 0)] \}, \quad (\text{S4})$$

and the primary test is its mean over the 24 hard layouts. By tertile, the marginal preference effect is -0.008 on hard, $+0.059$ $[+0.003, +0.122]$ on middle and $+0.190$ $[+0.128, +0.257]$ on easy layouts (unadjusted; Fig. S1). Its slope on calibration survival is $+0.025$ $[-0.102, +0.266]$ (permutation $p = 0.7595$). Using seed medians gives a primary estimate of -0.011 , and leaving out each hard layout in turn gives means from -0.018 to -0.006 .

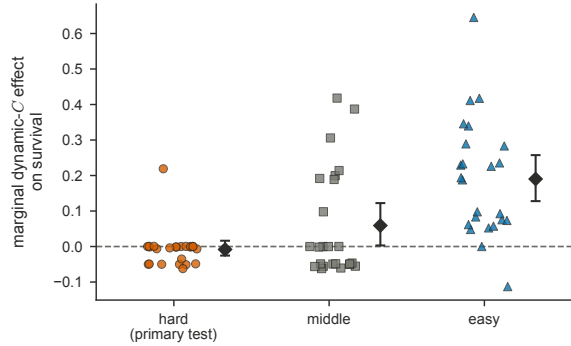


Fig. S1: **Marginal dynamic- C effect on survival for each of the 72 additional layouts.** Points are layouts, averaged over eight seeds and grouped by calibrated tier (24 each). Diamonds are tier means with 95% layout-bootstrap intervals. The hard tier is the primary test.

Among the internal measures, marginal κ -routing lowers posterior entropy, broadens coverage and slightly lowers headroom. Marginal dynamic- C raises headroom, lowers conflict, narrows coverage and raises posterior entropy. Each of these effects has a 95% interval that excludes zero, while the κ -routing effect on conflict does not.

C State coverage and the learned-model measure

Preference reweighting does not enter the Dirichlet update, but it changes which levels the agent visits and so which columns of the likelihood are updated. With preference reweighting the agent visits fewer levels, and it learns the levels it does visit at least as well (Table S2). These values come from a separate learning run of 30 episodes of 150 steps, with the uniform baseline and dynamic- C only, in which hunger and thirst ranged up to 15 rather than 10. The main text’s comparison with precision uses an earlier three-layout run of the same runner with all four conditions.

Table S2: Preference effects relative to the uniform baseline over 96 paired layout–seed records on the 12 layout IDs, two of which share coordinates. Quality is averaged over all levels of the three channels. Intervals resample layout IDs.

Measurement	Difference	95% CI
Model quality, all levels	−0.030	[−0.040, −0.021]
Model quality, visited levels	+0.008	[+0.002, +0.015]
Fraction of levels visited	−0.091	[−0.123, −0.068]