The Posterior-Prior Epistemic Gap: When World Model Imagination Fails to Carry Task-Relevant Signals

Jaerock Kwon[1] Elahe Delavari[1] Donghyun Kim[2] Haewoon Nam[2]
[1]Department of Electrical and Computer Engineering, University of Michigan–Dearborn
[2]Department of Electrical Engineering, Hanyang University, South Korea
{jrkwon, elahed}@umich.edu {kissw, hnam}@hanyang.ac.kr

Abstract

Model-based planners that roll out imagined futures assume the transition model faithfully propagates task-relevant signals—a property rarely measured directly. In a single-seed case study of an RSSM world model with active-inference planning in CARLA, we identify a **posterior-prior gap**—posterior (observation-filtered) states retain obstacle information that imagined rollout states progressively lose: a supervised obstacle probe achieves AUC 0.85 on posterior states, but the same probe applied to imagined prior states shows monotonic decay (AUC 0.85→0.71 over 15 steps; logit separation drops 57%). Because the probe is identical, the decay isolates the transition model—not the representation—as the information bottleneck. Obstacle avoidance drops to 0% without privileged runtime coordinates, and the planner's own ensemble-based epistemic term does not flag the loss—ensemble disagreement yields AUROC 0.52 as an obstacle classifier (chance level), making the failure an unflagged imagination failure—an unknown unknown at planning time.[1] As a secondary contribution, we show that **task-conditioned weighting of EFE components** over the same world model yields lane keeping ($81.2 \pm 1.8\%$ route completion) and obstacle avoidance (28/28 avoided, *using privileged coordinates*) in CARLA, with runtime weight modulation enabling post-evasion centering.

1. Introduction

Model-based planning with learned world models (Hafner et al., 2023) relies on rolling out imagined futures and scoring candidate action sequences. Planning quality therefore depends on whether the transition model faithfully propagates task-relevant signals through imagination—a property that is usually assumed but rarely measured directly.

Active inference (Friston, 2010; Parr et al., 2022) provides one concrete instantiation: an agent selects actions that minimize Expected Free Energy (EFE), naturally balancing goal-seeking (instrumental value) with information-seeking (epistemic value). We use it here as the planning framework, but the diagnostic we propose—probing the same auxiliary head on posterior versus imagined prior states—applies to any world-model-based planner that distinguishes filtered and predicted latent states.

We study an RSSM world model (Hafner et al., 2023) with iCEM planning (Pinneri et al., 2021) in the CARLA driving simulator (Dosovitskiy et al., 2017) and ask: *does the posterior's competence at encoding task-relevant signals transfer to the imagined prior that the planner actually uses?* In this single-seed case study, our answer is negative. The filtered posterior encodes obstacle presence at AUC 0.85, but probing the same auxiliary head on imagined prior states reveals 57% logit separation decay over 15 steps (AUC 0.85→0.71), and avoidance remains at 0%. Crucially, the planner's own ensemble-based epistemic term G_e does not flag this loss: ensemble disagreement as an obstacle classifier yields AUROC 0.52 (chance level), making the failure an unflagged imagination failure—an unknown unknown at planning time. Our contributions:

1. A **directly measured posterior-prior gap** in RSSM-based planning: probing imagined states shows 57% logit-separation decay over 15 imagination steps (AUC 0.85→0.71), yielding 0% avoidance without privileged information. The planner's ensemble-based epistemic term does not flag this loss (AUROC 0.52 as obstacle classifier—chance level).
2. A **diagnostic probe** that applies the same auxiliary obstacle head to both posterior and imagined prior states, isolating the transition model—not the representation—as the primary bottleneck (with the

[1]Project page: https://bimilab.github.io/paper-posterior-prior-epistemic-gap/

caveat that the probe is trained on posterior states, so distribution shift may contribute; see Section 6). This design can be applied to other world-model-based planners.

3. **Task-conditioned weighting of EFE components** over a shared RSSM world model, producing lane keeping ($81.2 \pm 1.8\%$ route completion) and obstacle avoidance (28/28 with *privileged* runtime coordinates) from a single objective via runtime weight modulation—demonstrating behavioral flexibility, not learned obstacle reasoning.

2. Related Work

World models for control. DreamerV3 (Hafner et al., 2023) uses RSSM imagination for RL-based policy learning, achieving strong results across diverse domains. MILE (Hu et al., 2022) learns driving world models for imitation. Transformer-based world models (Chen et al., 2022; Micheli et al., 2023) replace the recurrent transition with attention, potentially altering how information propagates through multi-step imagination; whether the posterior-prior gap we identify persists under attention-based transitions is an open question. These systems evaluate their world models on reconstruction and reward prediction, but not on whether safety-critical object signals survive multi-step imagination rollouts—the angle this paper takes.

Deep active inference. Catal et al. (2020) and Fountas et al. (2020) apply deep active inference with VAE world models to visual navigation and Atari games. Millidge (2020) connects active inference with variational policy gradients, and Lanillos et al. (2021) surveys active inference in robotics. Delavari et al. (2025) demonstrate active inference for lateral vehicle control via perceptual motor learning. These works focus on single behaviors; we additionally show that runtime weight modulation of EFE components can produce multiple behaviors from a shared world model.

Epistemic uncertainty in planning. Deep ensembles (Lakshminarayanan et al., 2017) provide calibrated uncertainty via disagreement among independently trained networks and are widely used in model-based RL (Chua et al., 2018). Evidential deep learning (Sensoy et al., 2018; Amini et al., 2020) places distributional priors over network outputs to separate aleatoric from epistemic uncertainty in a single forward pass. Sekar et al. (2020) plan to explore via world model uncertainty. These methods quantify epistemic uncertainty at *observation time*, but assume that the resulting signals transfer faithfully into model-based planning rollouts. Our results suggest that this assumption may not hold for RSSM-based active inference: posterior epistemic competence does not appear to guarantee prior (imagination) competence, even when the representation encodes the task-relevant signal with high fidelity.

Precision in active inference. Precision-weighting—modulating the relative influence of prediction errors—is central to biological active inference (Parr et al., 2022). Computational implementations typically use fixed weights; we demonstrate runtime modulation of EFE component weights to produce qualitatively different behaviors from a shared world model.

3. Background

3.1. Active Inference and EFE

Under active inference, an agent selects policies π that minimize Expected Free Energy:

$$G(\pi) = \sum_t \gamma^t \Big[\underbrace{\mathbb{E}_q[-\log p_{\text{pref}}(z_t)]}_{\text{instrumental (risk)}} + \underbrace{H[p(o_t|s_t)]}_{\text{ambiguity}} \Big], \quad (1)$$

where the instrumental term drives the agent toward preferred states p_{pref} and the ambiguity term penalizes states where observations are unpredictable from the latent state. The full epistemic value—mutual information $I(o;s) = H[q(o)] - H[p(o|s)]$—further includes an expected-novelty term, but neither decomposition is computed directly in our implementation; instead, Eq. 4 uses ensemble disagreement as an empirical proxy for epistemic uncertainty. We define p_{pref} in latent space rather than observation space because the RSSM planning rollout operates entirely in latent space, making latent-space preferences directly compatible with the planner without an additional decoding step. Precision-weighting modulates the relative influence of these terms based on their estimated reliability.

A note on "epistemic." We use this term in two distinct senses throughout the paper: (i) the *epistemic value* G_e in our EFE (Eq. 4), which uses ensemble disagreement as an empirical proxy for information gain, and (ii) the *posterior-prior epistemic gap*, which quantifies how much task-relevant signal the transition model loses when rolling out imagined states without observations. The first is a planning-time uncertainty signal; the second is a diagnostic of the world model's imagination fidelity.

3.2. RSSM World Model

The RSSM maintains deterministic state $h_t \in \mathbb{R}^{256}$ (GRU) and stochastic state $z_t \in \mathbb{R}^{64}$ (Gaussian):

$$\text{Posterior:} \quad z_t \sim q(z_t|h_t, o_t) \qquad \text{(filtering)} \quad (2)$$

$$\text{Prior:} \quad \hat{z}_t \sim p(z_t|h_t) \qquad \text{(imagination)} \quad (3)$$

The concatenated feature $[h_t; z_t] \in \mathbb{R}^{320}$ serves as the latent state for decoding and planning. During planning, only the prior is available—this distinction becomes critical in Section 6.

4. Method

4.1. Architecture

The world model comprises a ConvEncoder (4-layer CNN processing 64×64 RGB + 4D proprioceptive state $[\text{speed}, \text{steer}, h_{\text{err}}, c_{\text{err}}] \rightarrow$ 256D embedding), RSSM (deter=256, stoch=64), ObsDecoder (image reconstruction), StateDecoder (4D state prediction), and 5-head EnsembleTransitionHeads (epistemic uncertainty). Training minimizes Variational Free Energy (VFE) with dual KL divergence (Hafner et al., 2023), image and state reconstruction losses, on 96K frames of expert driving from CARLA Towns 04 and 06.

4.2. Multi-Component EFE

Our EFE scoring evaluates imagined trajectories from iCEM (500 samples, 50 elites, 5 iterations, horizon 12–15) with four components:

$$G(\tau) = \sum_{t=1}^{H} \gamma^t \Big[\beta_i G_i(t) - \beta_e G_e(t) + \beta_s G_s(t) + \beta_o G_o(t)\Big] \tag{4}$$

Instrumental value G_i: Cross-entropy $\mathbb{E}_q[-\log p_{\text{pref}}(z_t)]$ between imagined latents and a GMM preference fitted on expert driving latents (K=5, 64D). Z-scored across the CEM population per timestep.

Epistemic value G_e: Ensemble disagreement among 5 transition heads. Because G is minimized, the negative sign on $\beta_e G_e$ in Eq. 4 maximizes epistemic value, promoting information-seeking trajectories.

State-space value G_s: Decoded heading error and crosstrack error from the state decoder, penalizing deviation from lane center: $G_s(t) = \min(\hat{h}_t^2 + \hat{c}_t^2, 4.0)$. Not z-scored—state decoder predictions during imagination are noisy, and z-scoring amplifies this noise.

Obstacle penalty G_o: Runtime lane-change incentive using detected obstacle proximity and decoded crosstrack shift. Penalizes imagined trajectories that maintain the current lateral position when an obstacle is ahead.

4.3. Runtime Weight Modulation

The β weights are modulated at runtime based on prediction errors and context. These hand-designed schedules approximate precision-weighting (Parr et al., 2022) without formal optimality guarantees. Three rules apply: (1) *state error modulation*—when $e = h_{\text{err}}^2 + c_{\text{err}}^2 > 0.3$, β_s increases (up to 4×) and β_i decreases proportionally; (2) *obstacle proximity modulation*—when proximity $\rho > 0.1$, β_i is suppressed ($\beta_i \leftarrow \beta_i \cdot \max(0.2, 1{-}0.8\rho)$) to let G_o dominate; (3) *post-evasion centering*—after obstacle clearance, β_s is temporarily set to 0.5, pulling the ego back toward lane center.

Table 1. Task-conditioned EFE component weights. Same world model checkpoint; only β weights and horizon differ.

Parameter	Task A	Task B
$\beta_{\text{instrumental}}$	1.0	1.0
$\beta_{\text{epistemic}}$	0.1	0.1
β_{state}	0.5	0.0
β_{obstacle}	0.0	40.0
Horizon H	12	15

Table 2. Obstacle avoidance results using *privileged* runtime coordinates (mean ± std over episodes). SR: success rate; LC: lane changes per episode; MLD: mean lateral deviation (m).

Route	Eps	SR	Avoid	MLD	LC
R0 (373m, 2 obs)	5	100%	10/10	$0.67_{\pm.05}$	6.0
R1 (518m, 3 obs)	3	100%	9/9	$0.66_{\pm.03}$	6.0
R2 (518m, 3 alt)	3	100%	9/9	$0.67_{\pm.02}$	6.3
Total	11	**100%**	**28/28**	$0.66_{\pm.04}$	6.1

Table 1 shows the key difference: Task A uses β_s=0.5 for continuous centering while Task B sets β_s=0.0 to allow lateral deviation during lane changes, with β_o=40.0 for obstacle avoidance.

5. Behavioral Results (Privileged System Context)

All experiments use CARLA 0.9.16 with synchronous mode at 20 FPS. Mean lateral deviation (MLD) measures lane-center accuracy; success rate (SR) counts collision-free episodes.

5.1. Task A: Lane Keeping

On Town04 moderate curves (378m, 0°–85° heading change, 3 episodes), the agent achieves $81.2 \pm 1.8\%$ average route completion with 0.63 ± 0.13 m MLD. The state-space penalty (β_s=0.5) provides a Gaussian preference $\tilde{p}(s) = \mathcal{N}(0, I)$ over heading and crosstrack errors. The agent terminates at a sharp curve beyond the training distribution (>85°).

5.2. Task B: Obstacle Avoidance

On Town06 highway routes (373–518m, 2–3 static obstacles per route, using privileged obstacle coordinates), the agent avoids 28/28 obstacles across 11 episodes on three route configurations (Table 2). We note the limited scale: 11 episodes on 3 fixed routes with static obstacles.

Each obstacle is avoided via an active lane-change maneuver (∼6 lane changes per episode). Post-evasion centering

Table 3. Ablation on R1 (518m, 3 obstacles, 3 episodes each). Same 64×64 model throughout. "w/o privileged" removes all six privileged mechanisms simultaneously (not individually isolated); Compl.: route completion.

Variant	SR	Avoid	MLD	Compl.
Full system	100%	9/9	0.66	99.2%
w/o privileged	0%	0/9	0.28	43.8%
+ Centering	100%	9/9	0.52	99.1%

"w/o privileged": all 6 obstacle mechanisms removed (obstacle penalty G_o, proximity computation, evasion direction, reflexive steer, directional CEM penalty, motor primitives). "+ Centering": full system with post-evasion β_s=0.5.

($\beta_s \to 0.5$ after obstacle clearance) returns the ego toward the route lane between obstacles. Both tasks use the same trained world model checkpoint; the behavioral difference arises entirely from β configuration (Table 1).

5.3. Ablation

Table 3 shows Route 1 ablation results.

Removing all privileged mechanisms yields 0% avoidance despite reasonable driving (0.28 m MLD), confirming their necessity. Post-evasion centering improves MLD (0.66→0.52) without degrading avoidance.

6. Toward Vision-Based Obstacle Detection

The obstacle avoidance results above rely on runtime obstacle coordinates. We investigate whether this information can emerge from the latent space, removing all six privileged mechanisms (proximity computation, evasion direction, reflexive steer override, motor primitives, directional CEM penalty, obstacle EFE penalty).

6.1. Preference Gap Analysis

We fit a GMM on obstacle-free expert latents and measure the gap $\Delta = \log p(z_{\text{clean}}) - \log p(z_{\text{obs}})$. A positive gap would enable EFE-driven avoidance.

6.2. Contrastive Preference

We replace the single GMM with a log-ratio model: $\log p_{\text{pref}}(z) = \log p_{\text{clean}}(z) - \alpha \log p_{\text{avoid}}(z)$, where α controls how strongly obstacle-present latents are downweighted. This achieves correct posterior separation (37 nats raw at α=1.0, 1.1 effective nats at α=0.2), but avoidance remains 0%.

6.3. Auxiliary Prediction Head

To determine whether the bottleneck is the representation or the transition model, we add a supervised obstacle prediction head (MLP: 320→128→1, BCE loss, β_{aux}=1.0) during training only. On the training distribution, this achieves near-perfect posterior separation (0.000 vs 1.000 on in-distribution task labels) and amplifies the contrastive gap to 276 nats (133.1 effective). Cross-dataset probing (clean vs. obstacle-present data, Table 5) yields AUC 0.85—strong but imperfect, reflecting distribution shift between training and probing data. Yet avoidance remains at **0%**.

Table 4. Preference gap Δ across modalities and methods. Top rows vary input modality (separate world models); bottom rows vary the preference objective (same 64×64 model). α: contrastive down-weighting factor.

Method	Δ	MLD	Avoid
Single GMM (different modalities)			
64×64 RGB	−24.6	0.31	0%
128×128 RGB	−2.3	0.17	0%
64×64 RGBD	−68.1	—	0%
Contrastive preference (α=0.2)			
Without aux head	+1.1	0.17	0%
With aux head	+133.1	0.61	0%

Δ: effective nats (positive = correct). Aux head: binary obstacle prediction during training only.

6.4. The Posterior-Prior Epistemic Gap

Protocol. We collect 100 obstacle-free and 100 obstacle-present sequences from the evaluation routes, encode each with the RSSM posterior (using ground-truth observations), and record the auxiliary-head logit at the obstacle encounter frame. For imagined priors, we seed the RSSM at the same posterior state and unroll via `img_step` for k=1 to 15 steps using the agent's own planned actions, applying the same frozen auxiliary head at each depth (Table 5).

Caveat. The auxiliary head is trained on posterior states; evaluating it on prior states introduces a distribution-shift confound. Two observations mitigate this concern: (1) a fixed posterior–prior distribution offset would produce a *constant* AUC drop at k=1 with no further depth dependence, whereas we observe monotonic decay from k=1 through k=15 (Figure 1); (2) the sharpest drop occurs at k=1 (separation 4.89→3.59), consistent with the transition model's single-step information loss compounding over depth. Nevertheless, a probe retrained on imagined prior states is the clean next test; our setup lacks ground-truth obstacle labels aligned with imagination, which we leave to future work.

The posterior encodes obstacle presence with AUC 0.85, but imagination progressively erodes this signal: AUC drops to 0.71 and logit separation decays 57% over 15 steps

Table 5. Obstacle probe on posterior vs. imagined prior states. AUC and logit separation (mean obstacle − mean clean) decay with imagination depth.

State	AUC	Separation
Posterior	0.851	4.89
Prior k=1	0.788	3.59
Prior k=5	0.768	3.01
Prior k=10	0.742	2.54
Prior k=15	0.707	2.11

200 sequences (balanced), aux-head model. Separation decays 57% over 15 imagination steps.

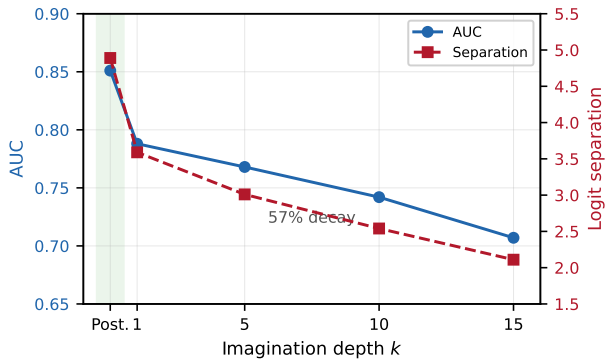


Figure 1. Posterior-prior epistemic gap across imagination depth. AUC (left axis) and logit separation (right axis) both decay monotonically as the RSSM rolls out imagined states without observations. AUC drops from 0.85 (posterior) to 0.71 (k=15), with logit separation decaying 57% over 15 steps. The sharpest drop occurs at k=1.

(Figure 1). This is consistent with the RSSM transition model $f(h_t, z_t, a_t) \to h_{t+1}$ failing to propagate obstacle information through action-conditioned rollouts (subject to the distribution-shift caveat above). The sharpest drop occurs at k=1 (separation 4.89→3.59), indicating that even a single imagination step substantially degrades obstacle encoding.

Why G_e does not help. The ensemble-based epistemic term G_e measures disagreement among five transition heads. During the same imagined rollouts, we measure whether G_e can distinguish obstacle-present from obstacle-free trajectories: the AUROC of ensemble disagreement as a binary obstacle classifier is 0.52 at k=1 and 0.51 at k=15 (chance level), with Cohen's $d < 0.1$ at all depths. The ensemble heads agree on *how* obstacles are forgotten—they share the same blind spot. Consequently, G_e provides no planning-time signal that obstacle information has been lost, making the failure an unknown unknown.

Combined with the 0% avoidance in Table 4, this shows that the residual prior signal (AUC 0.71) is insufficient to drive EFE-based obstacle avoidance through CEM planning.

7. Discussion

Task-conditioned weighting over a shared world model. Different β configurations over the same world model produce lane keeping, obstacle avoidance (with privileged coordinates), and post-evasion centering—consistent with precision-weighting in active inference (Parr et al., 2022), though our schedules are hand-designed.

Posterior ≠ prior epistemic competence. Direct probing (Table 5, Figure 1) reveals a measurable dissociation in this single-seed study: the posterior encodes obstacles at AUC 0.85, but imagination erodes this to AUC 0.71 over 15 steps (57% logit-separation loss). This suggests that planning quality may be bounded by the transition model's ability to propagate task-relevant information through imagination, not solely by the representation's expressiveness. We note the distribution-shift caveat (Section 6): because the probe is trained on posterior states, some of the measured decay may reflect probe–distribution mismatch rather than pure signal loss. Whether this limitation extends to other RSSM-based systems (e.g., DreamerV3 (Hafner et al., 2023)) is an open question.

An unflagged imagination failure. In this setup, the gap constitutes an operationally "unknown unknown": ensemble disagreement (G_e) evaluated as a binary obstacle classifier yields AUROC 0.52 at k=1 and 0.51 at k=15 (chance level; Cohen's $d < 0.1$), confirming that the planner's own uncertainty signal does not distinguish safe from dangerous imagined futures. From the planner's perspective, high-collision and safe trajectories are indistinguishable—a failure mode that cannot be detected without external observation. This aligns with the EIML workshop theme: in this case study, the agent does not know what it does not know during planning.

Limitations. This is a single-seed case study: one trained world model, 11 episodes across 3 fixed highway routes with static obstacles, and no training-seed variation. Results should be interpreted as a diagnostic case study, not as broadly generalizable claims. The auxiliary-head probe is trained on posterior states and evaluated on prior states, introducing a distribution-shift confound (Section 6); the monotonic depth-dependent decay is suggestive but not conclusive without a matched prior-state probe. The "w/o privileged" ablation (Table 3) removes all six privileged mechanisms simultaneously, so it confirms necessity of the bundle but does not isolate individual components. Obstacle avoidance relies on runtime obstacle coordinates—a form of privileged information—rather than learned perception. The component weighting schedules are hand-designed, not learned. We do not compare against non-AIF

planners (e.g., MPC with hand-designed cost) or imitation baselines. Sharp curves beyond $\sim 85^\circ$ exceed the training distribution, and no junctions or dynamic traffic are tested.

Future directions. Closing the posterior-prior gap requires transition models that learn object-conditioned dynamics. Attention-based transitions that selectively propagate object features through imagination, or object-centric world models (SLATE (Singh et al., 2022), SAVi (Kipf et al., 2022)) that decompose scenes into slots, may enable the transition model to predict how obstacles move through the visual field as a function of steering actions.

8. Conclusion

In a single-seed case study, we measured a posterior-prior gap in RSSM-based planning: a supervised obstacle probe drops from AUC 0.85 (posterior) to 0.71 (15-step imagination), with 57% logit-separation decay. Avoidance remains at 0% without privileged coordinates, and the planner's ensemble-based epistemic term does not flag the loss—making the failure an unknown unknown at planning time. While the probe's distribution-shift caveat prevents a fully conclusive attribution, the monotonic depth-dependent decay points to the transition model as the primary bottleneck. Closing this gap—through object-conditioned transitions, attention-based imagination, or imagination-aware uncertainty estimation—is a key challenge for world-model-based agents requiring object-level reasoning.

As a secondary contribution, task-conditioned weighting of EFE components over the same RSSM produced lane keeping ($81.2 \pm 1.8\%$ route completion) and obstacle avoidance (28/28 with privileged coordinates) from a single objective.

References

Alexander Amini, Wilko Schwarting, Amin Soleimany, and Daniela Rus. Deep evidential regression. In *Advances in Neural Information Processing Systems*, volume 33, pages 14927–14937, 2020.

Ozan Catal, Samuel Wauthier, Cedric De Boom, Tim Verbelen, and Bart Dhoedt. Learning perception and planning with deep active inference. *arXiv preprint arXiv:2001.11841*, 2020.

Chang Chen, Yi-Fu Wu, Jaesik Yoon, and Sungjin Ahn. TransDreamer: Reinforcement learning with transformer world models. *arXiv preprint arXiv:2202.09481*, 2022.

Kurtland Chua, Roberto Calandra, Rowan McAllister, and Sergey Levine. Deep reinforcement learning in a handful of trials using probabilistic dynamics models. In *Advances in Neural Information Processing Systems*, volume 31, 2018.

Elahe Delavari, John Moore, Junho Hong, and Jaerock Kwon. Perceptual motor learning with active inference framework for robust lateral control. *arXiv preprint arXiv:2503.01676*, 2025.

Alexey Dosovitskiy, German Ros, Felipe Codevilla, Antonio Lopez, and Vladlen Koltun. CARLA: An open urban driving simulator. In *Conference on Robot Learning*, pages 1–16, 2017.

Zafeirios Fountas, Noor Sajid, Pedro AM Mediano, and Karl Friston. Deep active inference agents using Monte-Carlo methods. In *Advances in Neural Information Processing Systems*, volume 33, pages 11662–11675, 2020.

Karl Friston. The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, 11(2):127–138, 2010.

Danijar Hafner, Jurgis Pasukonis, Jimmy Ba, and Timothy Lillicrap. Mastering diverse domains through world models. *arXiv preprint arXiv:2301.04104*, 2023.

Anthony Hu et al. Model-based imitation learning for urban driving. In *Advances in Neural Information Processing Systems*, volume 35, 2022.

Thomas Kipf, Gamaleldin F Elsayed, Aravindh Mahendran, Austin Stone, Sara Sabour, Georg Heigold, Rico Jonschkowski, Alexey Dosovitskiy, and Klaus Greff. Conditional object-centric learning from video. In *International Conference on Learning Representations*, 2022.

Balaji Lakshminarayanan, Alexander Pritzel, and Charles Blundell. Simple and scalable predictive uncertainty estimation using deep ensembles. In *Advances in Neural Information Processing Systems*, volume 30, 2017.

Pablo Lanillos et al. Active inference in robotics and artificial agents: Survey and challenges. *arXiv preprint arXiv:2112.01871*, 2021.

Vincent Micheli, Eloi Alonso, and François Fleuret. Transformers are sample-efficient world models. In *International Conference on Learning Representations*, 2023.

Beren Millidge. Deep active inference as variational policy gradients. *Journal of Mathematical Psychology*, 96: 102348, 2020.

Thomas Parr, Giovanni Pezzulo, and Karl J Friston. *Active Inference: The Free Energy Principle in Mind, Brain, and Behavior*. MIT Press, 2022.

Cristina Pinneri, Shambhuraj Sawant, Sebastian Blaes, Jan Achterhold, Joerg Stueckler, Michal Rolinek, and Georg Martius. Sample-efficient cross-entropy method for real-time planning. In *Conference on Robot Learning*, pages 1049–1065, 2021.

Ramanan Sekar, Oleh Rybkin, Kostas Daniilidis, Pieter Abbeel, Danijar Hafner, and Deepak Pathak. Planning to explore via self-supervised world models. In *International Conference on Machine Learning*, pages 8583–8592, 2020.

Murat Sensoy, Lance Kaplan, and Melih Kandemir. Evidential deep learning to quantify classification uncertainty. In *Advances in Neural Information Processing Systems*, volume 31, 2018.

Gautam Singh, Fei Deng, and Sungjin Ahn. Illiterate DALL-E learns to compose. In *International Conference on Learning Representations*, 2022.