EGWM Part 7: Experiment 7 – Adaptive Inner Voice

Mel Howard

December 2025

1 Introduction

The EGWM framework treats value channels ("feelings") such as confusion, competence, and elegance as low-dimensional signals that guide structural actions on a bank of specialist models (SPAWN, UPDATE, MERGE, FORGET). In earlier parts, these feelings were combined with a fixed governor to produce structural decisions.

However, a robust agent should not only act on its feelings; it should also learn *how much to trust each feeling* in different environments. In a stable world, an aggressive push for elegance and compression is desirable. In a rapidly changing world, the same preference for elegance can be catastrophic, preventing the system from allocating enough heads to track new regimes.

This part introduces an "adaptive inner voice" mechanism: a learnable value-weight vector that scales the feelings before they enter the governor. We then show, in a simple continual-learning benchmark, that meta-learning these value weights leads to different, environment-specific trade-offs between competence and elegance.

2 Experiment 7: Adaptive Inner Voice

2.1 Setup

We consider a simplified EGWM-style agent that maintains a bank of logistic heads for binary classification in a two-dimensional input space. Each "world" is defined by a linear separator in $\mathbb{R}^2$; labels are generated as

$$y = \begin{cases} 1, & \text{if } W^{(w)} \cdot x + b^{(w)} > 0, \\ 0, & \text{otherwise,} \end{cases}$$

with $x \in [-2, 2]^2$.

An *episode* consists of 40 phases. In each phase we sample a batch of points from the active world, train a scratch logistic regressor on that batch, and evaluate both the scratch model and all existing heads.

We study two environment types:

- **Stationary**: a single world is active for all 40 phases.
- **Switching**: four different worlds are active in sequence, each for 10 phases (world changes at phases 11, 21, and 31).

2.2 Feelings and value weights

For each phase we compute three scalar feelings:

- **Mismatch**: mismatch $= \text{acc}_{\text{scratch}} - \text{acc}_{\text{best}}$, where $\text{acc}_{\text{scratch}}$ is the accuracy of the scratch model and acc_{best} is the accuracy of the best existing head on the phase batch.
- **Competence**: comp $= \text{acc}_{\text{best}}$.
- **Complexity load**: $n_{\text{heads}}^{\text{norm}} = |\mathcal{H}|/10$, where $|\mathcal{H}|$ is the number of active heads, normalised by 10.

We introduce a value-weight vector

$$W = (w_{\text{conf}}, w_{\text{eleg}}, w_{\text{comp}}),$$

with one weight for each feeling family:

- w_{conf} controls how strongly mismatch and uncertainty push the agent toward growth.
- w_{eleg} controls how strongly complexity is penalised.
- w_{comp} controls how strongly low competence drives structural change.

These weights define a scalar *spawn score*

$$\text{score} = w_{\text{conf}} \cdot \text{mismatch} + w_{\text{comp}} \cdot (1 - \text{comp}) - w_{\text{eleg}} \cdot n_{\text{heads}}^{\text{norm}}.$$

The structural policy is:

- If no heads exist, **SPAWN** a head initialised from the scratch model.
- Otherwise, if score > 0.05, **SPAWN** a new head from the scratch model.
- Otherwise, **UPDATE** the current best head on the phase data.

MERGE and FORGET are disabled in this experiment; we focus on the trade-off between growth and minimality.

Environment	Weights W	Mean acc.	Mean # heads	Meta-reward
Stationary	$(1, 1, 1)$	0.987	1.10	0.965
Stationary	$W^\star \approx (3.13, 2.47, 0.14)$	0.987	1.00	0.967
Switching	$(1, 1, 1)$	0.976	2.40	0.928
Switching	$W^\star \approx (2.12, 1.10, 0.11)$	0.976	2.30	0.930
Switching	"too elegant" $(0.5, 3.5, 0.5)$	0.959	1.58	0.928
Switching	"growth-happy" $(3.5, 0.1, 0.5)$	0.978	5.26	0.873

Table 1: Experiment 7: Adaptive Inner Voice. The meta-learned value weights $W^\star$ differ between stationary and switching environments. In the stationary case, the optimal inner voice is highly elegant and maintains a single head while preserving maximal competence. In the switching case, the optimal inner voice increases the influence of confusion and relaxes elegance relative to the stationary setting, supporting a modest but non-trivial number of heads.

2.3 Meta-reward and learning the inner voice

Within an episode, we evaluate *oracle competence* by selecting, at each phase, the head with highest accuracy on a held-out batch from the active world. We also record the number of active heads.

We define a structural meta-reward

$$R_{\text{meta}} = \text{mean accuracy} - 0.02 \times \text{mean number of heads},$$

which trades off long-horizon competence against structural complexity.

For each environment type (stationary vs switching) we perform a simple meta-search over value-weight vectors:

1. Sample candidate weights $W = (w_{\text{conf}}, w_{\text{eleg}}, w_{\text{comp}})$ from a bounded range.
2. For each candidate W, run three episodes in the chosen environment and estimate R_{meta}.
3. Select the $W^\star$ with highest meta-reward as the learned "inner voice" for that environment.

We compare these learned weights to a baseline with fixed weights $W = (1, 1, 1)$, and also include two hand-crafted ablations in the switching environment: a "too elegant" setting with very high elegance weight and a "growth-happy" setting with very low elegance and high confusion weight.

2.4 Results

Table 1 summarises the results, averaged over 10 evaluation episodes for each setting and environment type.

In the **stationary** environment, the meta-learned weights $W^\star$ place strong emphasis on both confusion and elegance and down-weight competence. This combination keeps competence at its maximum while gently discouraging unnecessary spawning, resulting in a single active head and slightly higher structural reward than the baseline $W = (1, 1, 1)$.

In the **switching** environment, the optimal inner voice changes. The learned $W^\star$ still values elegance, but less aggressively than in the stationary case, and increases the weight on confusion. This leads to a modest number of heads (around 2–3) that can track the changing worlds without exploding in size. The "too elegant" ablation uses fewer heads but suffers a clear drop in accuracy, while the "growth-happy" setting attains strong accuracy at the cost of uncontrolled head growth and a lower meta-reward.

2.5 Discussion

This experiment demonstrates that even a simple value-weight mechanism is sufficient for EGWM to adapt its internal priorities to the environment. The same governor G_ψ—the same mapping from weighted feelings to actions—behaves very differently depending on the learned inner voice W:

- In stable worlds, EGWM learns a quiet, elegant inner voice that compresses structure into a single head without sacrificing competence.
- In non-stationary worlds, EGWM learns a growth-tolerant inner voice that keeps confusion loud and relaxes elegance just enough to maintain high competence with a small but non-trivial number of heads.

In other words, the agent is no longer constrained by a single, hand-chosen trade-off between confusion, competence, and elegance. Instead, it can learn how much each feeling should matter in different regimes, moving one step closer to an architecture that not only *acts* on feelings but also *manages* its own emotional priorities.