

The Observer Ecosystem

Governance as a Dynamic System of Competing Signals

A companion paper to The Observer Gap

Drew Bruce

drewbruce.me | 7-April 2026

EARLY DRAFT

Abstract

The Observer Gap (Bruce, 2026) proposed that general intelligence requires a separable governance layer — the Observer Function — capable of monitoring, representing, assessing, and modulating cognitive processing. That paper treated the Observer Function as a single control capability selecting from discrete governance postures. This companion paper argues that the Observer Function is itself a dynamic system: not a single observer but an ecosystem of competing governance signals — analogous to emotional and attentional states — whose interactions produce the adaptive posture that governs processing. The resulting architecture is one of interconnected balancing loops, where curiosity tempers caution, empathy modulates aggression, and alertness escalates or de-escalates based on the confluence of multiple internal signals. This paper extends the Externally Modulated Adaptive Control (EMAC) pattern from a single-token policy selector to a multi-signal governance resolution system, situates it against empirical evidence from the interactive optimisation literature, and proposes a research programme for testing the hypothesis.

Draft research position paper — companion to The Observer Gap

1. From Single Observer to Observer Ecosystem

The Observer Gap established a binary distinction: systems either have a separable governance layer or they do not. This framing was deliberately simplified to make the core claim as clean and testable as possible. But human experience suggests that the observer capacity is not a single, monolithic function. It is an ecosystem of competing signals that resolve into a governance posture.

Consider the phenomenology of human decision-making under uncertainty. A person facing a complex choice does not experience a single “observer” calmly selecting a policy. They experience curiosity pulling toward exploration, anxiety urging caution, empathy weighting the impact on others, anger driving toward decisive action, and fatigue dampening all of the above. The resulting behaviour is not selected by any one of these signals. It emerges from their interaction — from the way they amplify, inhibit, and balance each other in real time.

This is not merely a subjective phenomenon. It has architectural implications. If the Observer Function is a single policy selector, the design challenge is relatively straightforward: build a classifier that maps internal state to a governance posture. If the Observer Function is itself a dynamic system of interacting signals, the design challenge is qualitatively different. The governance layer must contain its own internal dynamics — feedback

loops, inhibition pathways, amplification cascades — that produce emergent postures rather than selected ones.

This paper argues for the latter: that the Observer Function, properly understood, is an ecosystem of governance signals whose interactions produce adaptive behaviour. The single-token EMAC model described in The Observer Gap is a necessary simplification for first-order testing, but it is not a sufficient model of what human-level governance actually requires.

2. Balancing Loops: How Competing Signals Produce Governance

The concept of balancing feedback loops is well established in systems dynamics. A balancing loop is a causal structure in which an increase in one variable triggers a response that counteracts the increase, tending the system toward equilibrium. Human emotional regulation operates through precisely such loops.

2.1 The Curiosity-Caution Loop

Curiosity drives exploration: the pursuit of novel information, the testing of hypotheses, the willingness to engage with uncertainty. Left unchecked, curiosity produces reckless behaviour — a system that explores without regard for risk. Caution (modelled in the EMAC framework as the ANXIOUS state) counterbalances curiosity by raising validation thresholds, reducing retry tolerance, and increasing escalation sensitivity. The balance between these two signals determines whether the system explores boldly, explores cautiously, or retreats to defensive processing.

In human experience, this loop is familiar. A researcher feels drawn to an unconventional hypothesis (curiosity) but hesitates because the evidence is thin (caution). The resulting behaviour — cautious investigation rather than either full commitment or full retreat — is an emergent property of the loop, not a policy selected from a menu.

2.2 The Empathy-Anger Loop

Empathy modulates action by weighting the impact of decisions on other agents. Anger drives decisive, boundary-enforcing action. In isolation, empathy produces paralysis (every action harms someone) and anger produces destructive overreaction. Together, they produce assertive boundary-setting — the capacity to act firmly while remaining aware of consequences.

In an AI governance context, this loop would modulate how a system handles conflicting stakeholder requirements, ethical trade-offs, or situations where optimal outcomes for the system conflict with optimal

outcomes for users. Neither pure “empathy” (total deference to external preferences) nor pure “anger” (total commitment to internal objectives) produces good governance. The balance does.

2.3 The Alertness Escalation Cascade

Alertness is not a binary state. It exists on a continuum from nominal monitoring through heightened attention to full crisis response. The escalation is driven by signal accumulation: a single anomaly raises alertness slightly; multiple correlated anomalies compound into a state shift. De-escalation follows the reverse path, but typically more slowly — hysteresis in the system prevents rapid oscillation between crisis and calm.

This cascade maps directly onto the EMAC state progression (CALM → CURIOUS → ALERT → ANXIOUS → DISTRESSED) but reframes it as a continuous dynamic rather than a discrete classification. The governance posture at any moment is not a selected state but a position on a multidimensional landscape shaped by the current balance of all active signals.

3. Empirical Precedent: The Human in the Loop as Observer Ecosystem

Liu, Dwyer, Tack, Gratzl and Marriott (2021) demonstrated in the Problem-Solving Loop that interactive optimisation — where a human collaborates with an algorithm by evaluating outputs, refining models, and guiding search — produces better outcomes than autonomous black-box execution. The human in this loop is performing precisely the Observer Function described in *The Observer Gap*: monitoring processing, representing it as an object of evaluation, assessing it against higher-order goals, and modulating subsequent execution.

But the human in the loop does not operate as a single policy selector. They bring their full emotional and cognitive repertoire: curiosity about unexplored solution spaces, frustration with poor results driving model refinement, caution when solutions seem too good to be true, and satisfaction providing a termination signal when the result is acceptable. The quality of their governance is a function of the interplay between these signals, not any single one.

This suggests that the success of human-in-the-loop systems is not merely because they add a governance layer. It is because they add a governance layer with internal dynamics — one that adapts, self-corrects, and resolves competing objectives through the kind of emotional and cognitive balancing that no current AI architecture replicates.

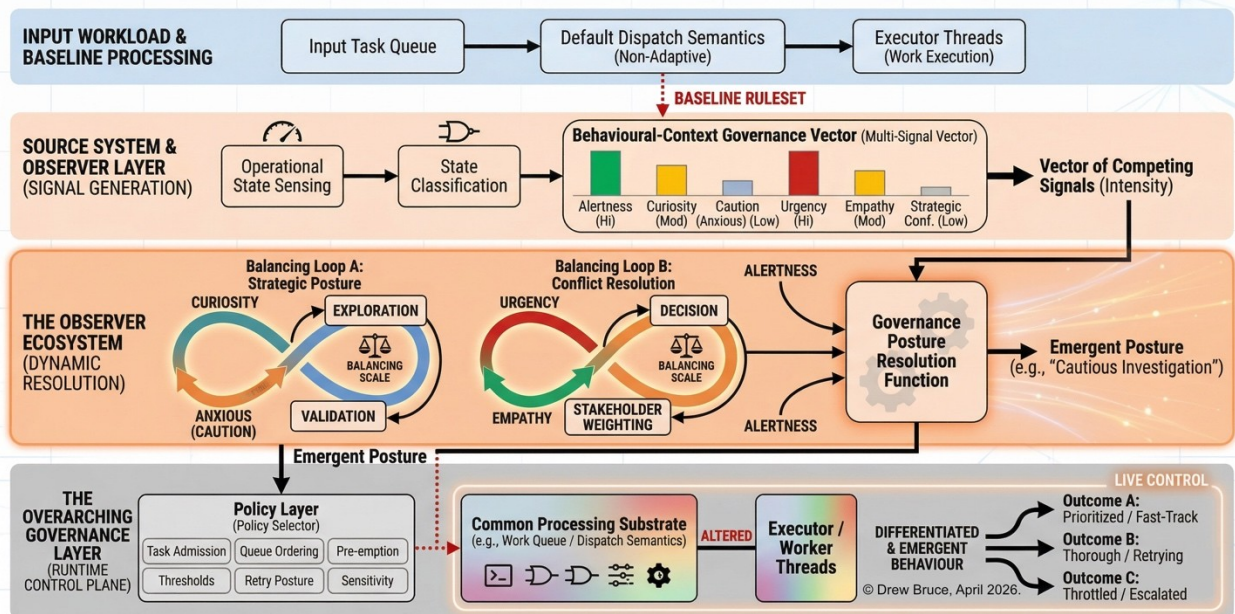
The architectural insight is this: the path to internalising the human-in-the-loop advantage is not to build a single monitor that evaluates processing. It

is to build an ecosystem of competing governance signals that resolve into adaptive postures — and to do so as a separable architectural layer, consistent with the EMAC framework.

4. From EMAC to Multi-Signal Governance Resolution

The Observer Gap described EMAC as a pattern where a single Behavioural-Context Token selects a governance posture. This paper extends the model. Instead of a single token, the Source System publishes a vector of governance signals — each representing the current intensity of a distinct evaluative dimension (curiosity, caution, urgency, empathy, confidence, fatigue, and others). The Policy Selector resolves this vector into a governance posture through a resolution function that models the interactions between signals.

Observer Ecosystem: Dynamic multi-signal governance resolution



4.1 Conceptual Band 1: The Input Workload (Baseline Processing)

This top section presents a simple flowchart of baseline system operation: the execution of tasks *prior to* any adaptive governance. It depicts a clean, linear sequence of standard software steps:

1. **Workload:** This box represents the queue of incoming tasks.
2. **Dispatch Semantics:** This represents the non-adaptive, default rules for how tasks are picked up and prioritized. A red dotted line will connect this to a lower band, showing that these rules are *what* the governance layer will eventually modulate.
3. **Executor Threads:** The basic mechanism for processing the workload.
4. **Nominal Output:** The standard, non-differentiated outcome of processing.

4.2 Conceptual Band 2: Source System & Observer Layer (Signal Generation)

This section describes the provenance of the 'Behavioural-Context' vector: a flow that feeds directly into the core ecosystem.

1. **Operational State Sensing:** A box with a small icon (a thermometer or gauge) that monitors the environment and the system's internal metrics (e.g., resource usage, conflict, ambiguity).
2. **State Classification:** An icon of a classifier, which takes raw sensing data and converts it into structured state models.
3. **Governance Signal Vector:** This is the key component. It is a large container box displaying a clean bar chart with labeled, color-coded bars for each competing signal:
 - o **Alertness:** (e.g., High, medium, low)
 - o **Curiosity:** (Pulling for explore/novelty)
 - o **Caution (ANXIOUS):** (Pushing for safety/validation)
 - o **Urgency:** (Driving rapid dispatch/high execution)
 - o **Empathy:** (Weighting stakeholder impact)
 - o **Strategic Conf.** (Dampening or amplifying other signals)

A single arrow labeled "Discrete Vector of Competing Signals (Intensity)" leads from this vector into the main ecosystem band.

4.3 Conceptual Band 3: The Observer Ecosystem (Dynamic Resolution)

This is the central, most dynamic part of the diagram. It is a large, slightly glowing orange/peach container box that shows how the signal vector is processed *before* a policy is selected. It features two prominent, interwoven "Balancing Loop" structures:

- **Balancing Loop A: 'Strategic Posture'**
 - o This shows an infinity-loop shape. On one side, **CURIOSITY** pushes an arrow toward **EXPLORATION**. A counter-arrow labeled **ANXIOUS (CAUTION)** pushes back, urging **VALIDATION**. A central icon, a balancing scale, indicates that the system is resolving a strategic middle-ground.
- **Balancing Loop B: 'Conflict Resolution'**
 - o This adjacent loop shows **URGENCY** driving an arrow toward **DECISION**. A counter-arrow labeled **EMPATHY** pulls back, urging **STAKEHOLDER WEIGHTING**. Another balancing scale icon sits at the center of this loop.

Both of these loops and any remaining active signals (like general ALERTNESS) are fed into a single large node: the **Governance Posture Resolution Function**. This is depicted as a complex, interlocking mechanism (perhaps subtle, transparent gears). This function applies an interaction model (e.g., weighted summation or threshold dynamics) to resolve the competing signals. A bold arrow leaves this node, labeled "**Emergent Posture** (e.g., 'Cautious Investigation')."

4.4 Conceptual Band 4: The Overarching Governance Layer (The Runtime Control Plane)

This bottom band shows how the finalized posture controls the runtime execution, closing the loop. It is a large gray container that mirrors the "work" from the very top band:

1. **Policy Layer (Governance Posture):** This box receives the 'Emergent Posture' and displays the list of six specific, color-coded policy parameters that it is now "setting" for the runtime (Task Admission, Queue Ordering, Pre-emption, Thresholds, Retry Posture, and Sensitivity).
2. **Work Queue / Dispatch Semantics:** This is a box with an icon. It has clear visual markers (like color shifts) indicating that its state has been changed by the input from the Policy Layer. The red dotted line from Band 1 connects here, visually confirming that the original semantics have been over-ridden.
3. **Executor / Worker Threads:** This box also has a color shift, showing that the worker threads are now operating *under* the new governance rules.
4. **Differentiated & Emergent Behaviour:** This final output box has three branching arrows to show the *kind* of divergent outcomes now possible (e.g., Outcome A: Prioritized/Fast-Track, Outcome B: Thorough/Retrying, Outcome C: Throttled/Escalated).

4.5 From Discrete Selection to the Observer Ecosystem

The initial proposal of the Observer Function relied on a simplified model of **Externally Modulated Adaptive Control (EMAC)**, where a singular behavioral-context token acted as a discrete policy selector. While this provided a clean architectural primitive for first-order testing, it failed to capture the nuanced, often contradictory nature of human-level governance.

We now extend this model into an **Observer Ecosystem**. In this framework, the Observer Function is no longer a monolithic entity but a dynamic system of competing governance signals—such as curiosity, caution, and empathy—whose real-time interactions produce an emergent adaptive posture.

4.5.1 The Logic of Balancing Loops

Human governance is rarely the result of a single "calm" observer selecting a policy. Instead, it emerges from interconnected **balancing feedback loops**:

- **The Curiosity–Caution Loop:** Curiosity drives the pursuit of novel information and hypothesis testing. It is counterbalanced by Caution (the *Anxious* state),

which raises validation thresholds and reduces retry tolerance to prevent reckless behavior.

- **The Empathy–Anger Loop:** This loop modulates action by weighting the impact on other agents (Empathy) against the need for decisive, boundary-enforcing action (Anger). The balance between these signals prevents both destructive overreaction and total deference.
- **Alertness Escalation:** Alertness exists on a continuum rather than a binary. Escalation occurs through signal accumulation, where correlated anomalies compound into a state shift, while de-escalation is often slower (hysteresis), preventing rapid oscillation between crisis and calm.

4.5.2 Mathematical Resolution and Interpolation

To move beyond a "nearest attractor" selection model, the architecture employs a **Multi-Signal Governance Resolution** system. The source system publishes a vector of intensities for distinct evaluative dimensions. Rather than selecting a predefined state, the system resolves this vector into a governance posture through a resolution function.

By interpolating between defined configurations, the system can produce **genuinely novel postures**. In this model:

- **Emergent Postures:** Parameters such as task admission, execution thresholds, and retry posture are calculated as weighted averages of all active attractor states.
- **Non-Linearity:** Governance becomes a coordinate in a multidimensional landscape rather than a hard-coded linear flow.
- **Self-Correction:** If a single signal like "urgency" dominates inappropriately, healthy interaction dynamics allow balancing signals like "caution" to counteract it, creating a richer self-correction capacity than a single-token system.

4.6 Implications for Alignment and AGI

The transition from metaphor to architecture is critical. We posit that the functional roles played by emotions in human governance are not "feelings" in the subjective sense, but evolved mechanisms for modulating behavior that are architecturally necessary for general intelligence.

In this ecosystem model, alignment is reframed: it is less about "getting the right answer" and more about maintaining **healthy internal signal dynamics**. A system that can represent its own reasoning as an object of evaluation is structurally better equipped to detect internal inconsistency and default to caution when governance signals conflict.

5. Implications and Predictions

5.1 Emergent Governance vs. Programmed Governance

A system with a multi-signal observer ecosystem should produce governance behaviours that were not explicitly programmed. Because the posture emerges from signal interaction rather than selection, novel combinations of signals should produce novel postures — appropriate responses to situations the designer never anticipated. This is a testable prediction: present the system with scenarios outside its training distribution and measure whether its governance posture degrades gracefully or fails categorically.

5.2 Emotional Analogy is not Metaphor

The use of terms like “curiosity,” “caution,” and “empathy” to describe governance signals is not anthropomorphism. It is a recognition that human emotional states are themselves governance signals — evolved mechanisms for modulating behaviour in response to internal assessment of context. Emotions are the human Observer Ecosystem. The claim is not that AI should “feel” but that the functional role emotions play in human governance is architecturally necessary for general intelligence, and that role can be implemented without requiring subjective experience.

5.3 Alignment Implications

A system governed by a single policy selector has a single failure mode: wrong policy. A system governed by an ecosystem of competing signals has a richer failure landscape but also richer self-correction capacity. If one signal (e.g., urgency) dominates inappropriately, the balancing signals (caution, empathy) can counteract it — provided the interaction dynamics are healthy. This maps directly onto the human experience of emotional dysregulation: the system isn't broken, the balance is. Alignment in an ecosystem model is therefore less about getting the right answer and more about maintaining healthy signal dynamics.

6. Research Programme

The extension from single-observer to observer-ecosystem generates several testable predictions beyond those outlined in The Observer Gap:

First, multi-signal resolution should produce more nuanced governance than single-token selection. A comparative study varying the number and interaction complexity of governance signals while holding the processing substrate constant should show measurable improvements in adaptiveness for multi-signal systems.

Second, systems with balancing loops should be more robust to adversarial governance inputs than single-signal systems. Corrupting one signal in an ecosystem should degrade performance gradually, while corrupting the single token in an EMAC system should produce catastrophic posture failure.

Third, the resolution function should exhibit hysteresis — resistance to rapid state oscillation — mirroring the well-documented phenomenon in human emotional regulation where state transitions are asymmetric (faster to escalate, slower to de-escalate).

Fourth, the Liu/Dwyer Problem-Solving Loop framework provides a natural experimental testbed: replace the human in the loop with a multi-signal observer ecosystem and measure whether the system approximates the governance quality that human participants provide.

7. Conclusion

The Observer Gap argued that general intelligence requires a separable governance layer. This paper argues that the governance layer itself must be a dynamic system — not a switch but an ecosystem. Human intelligence is not governed by a single observer selecting policies. It is governed by an interplay of competing signals whose resolution produces adaptive behaviour. Replicating that interplay — not as metaphor but as architecture — may be the next step toward systems that are genuinely intelligent rather than merely capable.

The path forward is not one paper but a programme of work: formalise the signal interaction model, build testable implementations, and measure whether multi-signal governance produces qualitatively different behaviour from single-signal selection. The hypothesis is clear, the predictions are testable, and the engineering is within reach. What remains is the will to build it.

References

- [1] Bruce, D. (2026). The Observer Gap: Why AGI Won't Emerge From Better Pattern Matching. drewbruce.me.
- [2] Liu, J., Dwyer, T., Tack, G., Gratzl, S. & Marriott, K. (2021). Supporting the Problem-Solving Loop. IEEE TVCG, 27(2), 1764-1774.
- [3] Baars, B.J. (1988). A Cognitive Theory of Consciousness. Cambridge University Press.
- [4] Damasio, A. (1994). Descartes' Error: Emotion, Reason, and the Human Brain. Putnam.
- [5] Sterman, J.D. (2000). Business Dynamics: Systems Thinking and Modeling for a Complex World. McGraw-Hill.
- [6] LeDoux, J. (1996). The Emotional Brain. Simon & Schuster.
- [7] Barrett, L.F. (2017). How Emotions Are Made: The Secret Life of the Brain. Houghton Mifflin Harcourt.

[8] Meadows, D.H. (2008). Thinking in Systems: A Primer. Chelsea Green Publishing.

© Drew Bruce, April 2026. All rights reserved.