Prediction, Control, and the Regulating Mind

Executive Summary — A Proposed Model of Regulation

Dr. Greg Moody

August 17, 2026

Executive summary

What this is, and what it isn't

Before I was a licensed psychotherapist I was an aerospace engineer. Guidance, navigation, control. I'm not a research psychologist, which is the reason this exists and also the reason to read it skeptically.

The claim is smaller than the title makes it sound. Two bodies of work already exist. One says the mind is a control system defending targets. The other says it's a prediction machine guessing at the world. Each has a hole exactly where the other has something solid, and nobody seems to have bolted them together to see what falls out. When you do, you get a handful of parameters that come apart cleanly, and they predict different things about who gets better with which kind of help.

Not a new science. A join between two existing ones.

This is not an alternative to cognitive behavioral therapy. CBT has decades of outcome data. A control-theoretic account of its mechanism has none. Nothing here asks anyone to stop doing what they're doing or to add a technique.

The two halves

The mind as a control system. Targets, error signals, corrective action. Emotions are the error signals. Hunger is what sitting below a set point feels like from the inside. Fear is what it feels like to have estimated more danger than you're built to tolerate. This is cybernetics, it goes back to Wiener and Powers, and it has been the losing side of an argument since the cognitive revolution.

On the positive side of the ledger, the same account holds that **happiness is not an emotion**, because nothing works to drive it to zero. It's what error correction feels like, which is why a large or fast correction produces more of it than a slow one. The best meal of your life was eaten hungry. Amusement, awe and curiosity are left unaccounted for, and I say so rather than stretching the model to cover them.

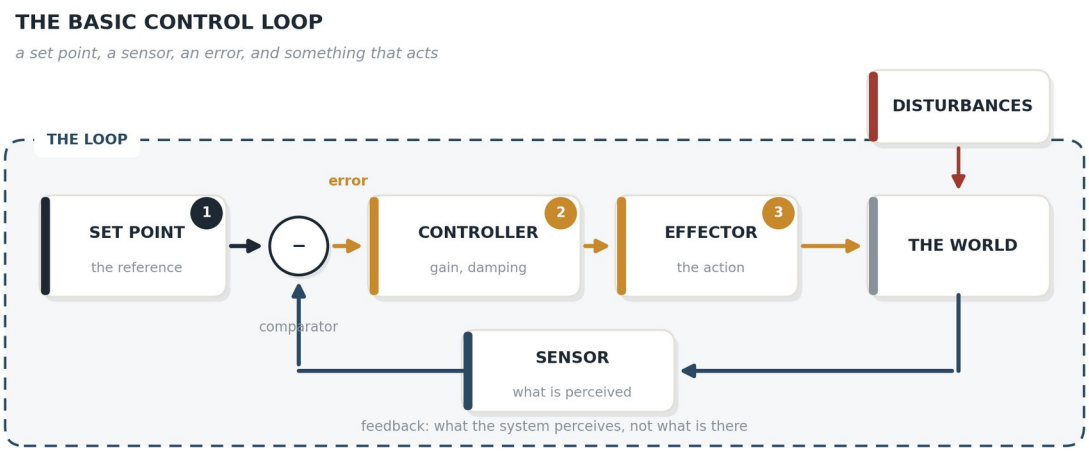


Figure 1. The basic control loop. A set point, a sensor, an error, and something that acts.

Where it falls short. A thermostat reads a number and takes it at face value. No biological system gets that deal. Its sensors are noisy, slow, and measure proxies. Plasma osmolality is not thirst. A racing heart is not danger. The control account has a sensor-shaped hole in it.

The mind as a prediction machine. The brain guesses at the world continuously and corrects the guess only where the evidence insists. Perception is inference. This is predictive processing, and it currently dominates computational psychiatry. It explains beautifully how an estimate gets built, and says very little about why the organism cares or how hard it acts once it knows.

> *Aside… if you have read Plato's The Allegory of the Cave, this is kind of congruent with that… but enough intellectualizing here….*



Figure 2. The allegory of the cave. The prisoners are chained facing a wall, a fire burns behind them, and carried objects throw shadows onto the wall in front of them. They see the shadows, not the objects casting them, and take the shadows for the world.

The prisoners aren't reasoning badly. Their inference is right given what reaches them, and what reaches them is a projection rather than the thing itself. That's the predictive account of perception, arrived at about twenty-four centuries early. The shadows are the sensory input, the prior is what a prisoner already expects a shadow to mean, and precision is how much authority the shadow gets against that expectation.

Where the set points come from

A set point is a learned quantity, and the load-bearing ones were estimated early. The calibration runs through the trade between attachment and authenticity (Maté & Maté, 2022): when what a child feels or needs threatens the connection with a caregiver, the child suppresses the signal, and a few thousand repetitions of that produce an estimate of what the child has to be for the connection to hold. Genetics sets the gain on that process rather than the value it lands on, and the range runs in both directions, so the same sensitivity that costs a child in a harsh environment pays them in a good one (Belsky & Pluess, 2009; Pluess & Belsky, 2013). Gene-environment correlation (Plomin et al., 1977; Klahr & Burt, 2014) adds that the child helps write the environment doing the calibrating, which is why two children in one house didn't get one childhood. A defended value isn't permanent, it just doesn't move on argument.

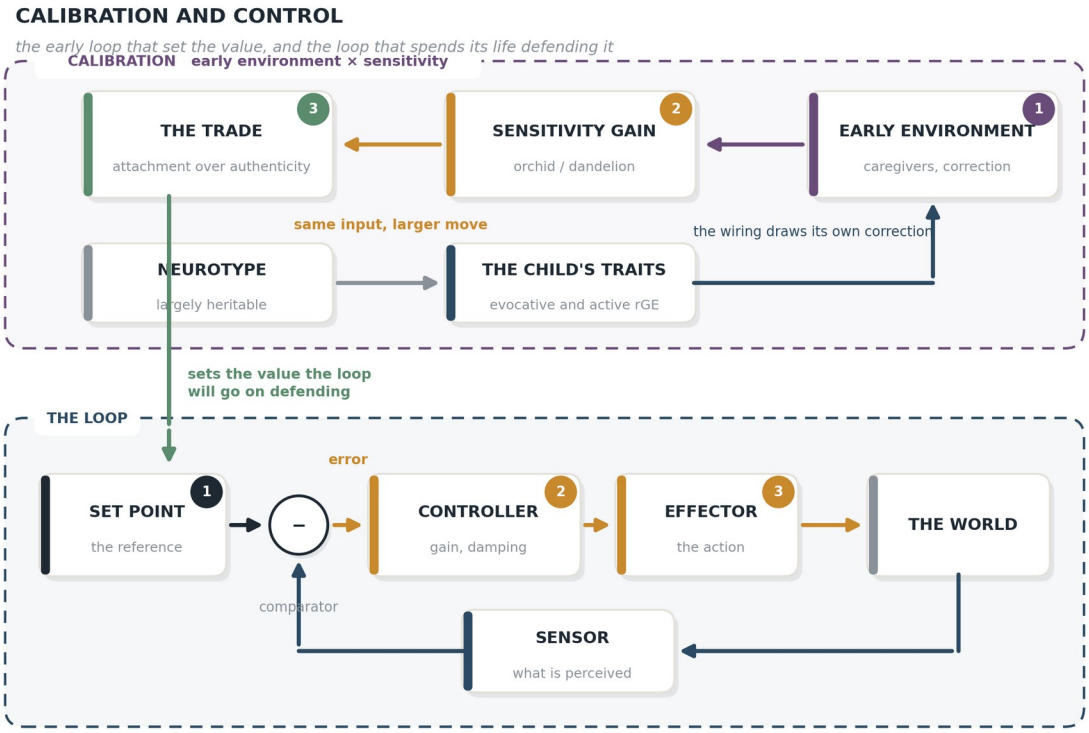


Figure 1a.1. Calibration and control. The lower loop defends a set point. The upper loop estimated it.

The combination

The estimate establishes the target. The estimator combines expectation with sensation and produces an estimate of the current state. That estimate is also the updated prediction of what should be happening, and that prediction is what the controller compares against. There's no separate goal off to the side. The thing you predict IS the set point you defend.

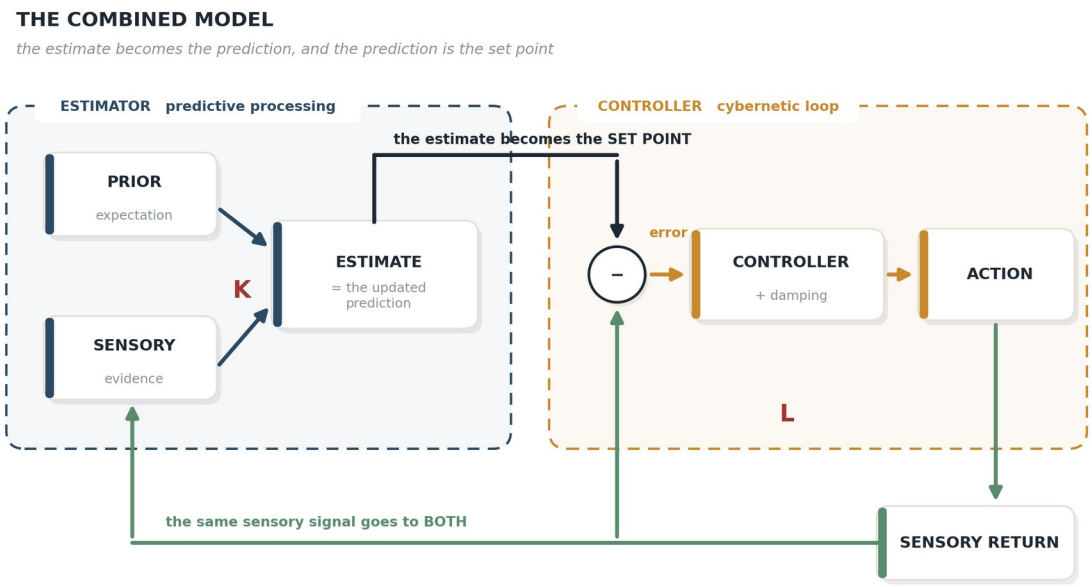


Figure 3. The combined model. The estimate becomes the prediction, which becomes the set point. The sensory return feeds both halves.

That gives two gains, formally distinct.

Observer gain (K) governs how much your senses move your estimate, relative to what you already expected. Predictive processing calls this *precision*. Because the estimate becomes the target, K determines where the target sits.

Controller gain (L) governs how hard you act on the gap. It has nothing to do with belief.

The separation is the separation principle from control theory, Kalman 1960. What's new is the claim that K sets the target rather than only the confidence, which is what makes the two dials clinically different rather than merely separable.

Both are learned. Volatility teaches observer gain: grow up somewhere unpredictable and you learn to weight new evidence heavily, because there it *was* heavily weighted. Consequence teaches controller gain: where small problems reliably became large ones, high gain is the correct setting. Overshoot teaches damping.

Most of what looks like dysregulation is a correctly learned parameter running in an environment that no longer matches it.

One way to think about all of it: a shortcut. A heuristic is what's happening whenever either half bypasses differentiated conscious thought and the thing simply runs. You walk into a room and already know what kind of room it is. Something happens and you have already responded. The word names a MODE both halves spend most of their time in, not a new part, and you need them: without shortcuts you'd consciously think through every little thing you do, from driving a car to what hand to use to eat.

The schema is the stored part, a cognitive assessment that's embedded, and the estimator is the function that adjusts it as evidence comes in. A shortcut keeps running on what it learned until enough turns up that adjusts the system, and where there's serious trauma the evidence has to be much stronger. It goes wrong two ways: the situation sits outside what taught it, or it's maladaptive, learned where it worked (like between 5-10 years old) and still running somewhere it doesn't. Both can happen with the estimator or the controller, and they need different solutions.

How it fits CBT

The four terms map one to one, with nothing left over.

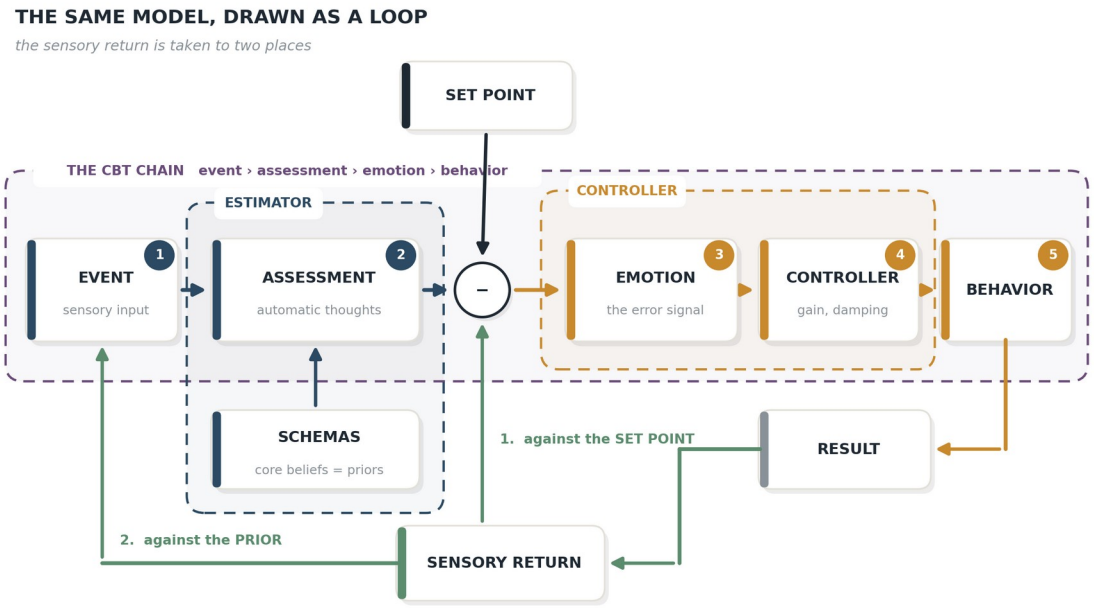


Figure 4. The CBT chain drawn as a control loop. The sensory return goes to two places: compared against the set point, and weighed against the prior.

The event is sensory input. The cognitive assessment is the estimator, running on schemas as priors. The emotion is the error signal. The behavior is the controller's output.

What the redraw adds is the piece the standard chain doesn't show: the behavior produces a result, the result is the next event, and the loop closes. That's why a maladaptive core belief survives being wrong. It runs a loop that manufactures its own evidence. Not all core beliefs are maladaptive; it's the maladaptive ones that get in the way.

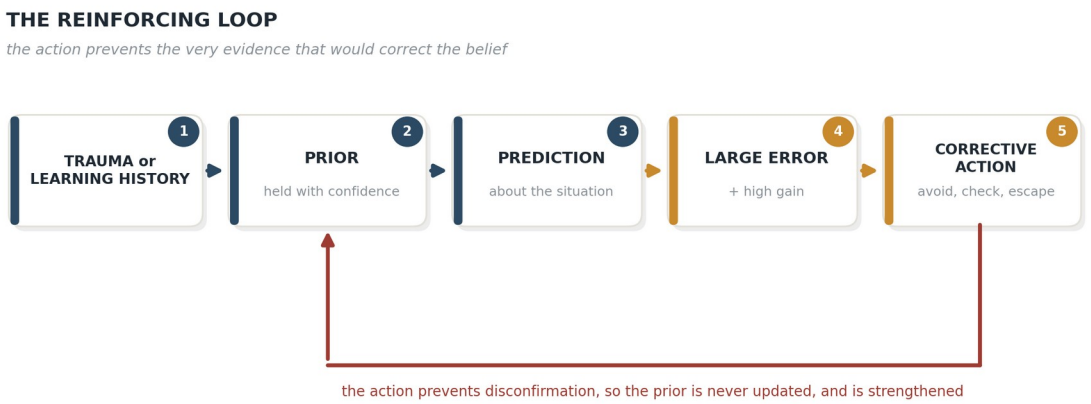


Figure 5. The reinforcing loop. The corrective action prevents the very evidence that would correct the belief.

What it gives a practitioner

Four questions that sort people into different treatments.

Is this belief inaccurate, or accurate and acted on too hard? Restructuring addresses the second and does little for the first.

Is the belief wrong, or held with too much confidence? Content and confidence come apart. Behavioral experiments work on confidence. That's why the experiment beats the argument.

Is the person failing to update, or being prevented from updating? Safety behavior stops the error signal from being generated. The prior isn't stubborn, it's starved.

What environment taught this setting, and does it still exist? Usually answerable, and it changes the conversation from correction to recognition.

What it gives everyone else

The same parameters describe a person deciding under pressure, a manager running a team, and an organization steering itself. Not by analogy — by the same mechanics.

Most business failures of this kind are lag failures. A change goes in Monday, results land in six weeks, the owner checks Wednesday and changes it again. The campaign didn't fail. The loop got run faster than the system could answer. Set the review interval longer than the feedback delay, and set it in advance, before stress sets it for you.

Micromanagement is a gain and damping problem, not a personality problem. Checking every two hours generates corrections faster than the work can absorb them, output oscillates, the manager reads that as unreliability and checks more. Lengthen the interval and the oscillation damps out.

KPIs are set points, and they inherit every failure mode: arbitrary targets, proxy sensors, unacknowledged reporting lag, gain sized for a bad year applied to a bad month, and no damping because strategy changes before any strategy produces evidence about itself.

What the evidence actually supports

Solid. People track volatility and adjust learning rates accordingly (Behrens et al., 2007). Anxious people fail to make that adjustment (Browning et al., 2015). Exposure works by expectancy violation, not habituation (Craske et al., 2014). Safety behaviors block disconfirmation (Salkovskis et al., 1999).

Reinterpretation, not result. The CBT mapping, rumination as oscillation, response prevention as damping. Coherent. Not findings.

Untested. The central claim. Nobody has shown that observer gain and controller gain dissociate in real people or predict differential treatment response. That study is designable and cheap, and it has not been run.

Status, stated plainly

This is a proposal, not a finding. Section 9 of the full document lists four results that would show it's wrong, stated up front, because a model that can absorb any outcome isn't really saying anything.

The strongest objections survive: only the ratio of sensory to prior precision enters a Gaussian update, so the two dials may not be identifiable from most data; Friston argues there's no separate controller to bolt an estimator onto; and the model says nothing about meaning, narrative or grief. It describes a regulatory layer. It isn't a theory of persons.

What's worth doing now

Ask the two questions separately: how confident is this belief, and how hard is this person acting on it.

Set the review interval before you start, and make it longer than the system's response time.

When a pattern looks maladaptive, ask what environment would have made it correct.

Companion documents apply this model to autism and repetitive behavior, including the full evidence audit.