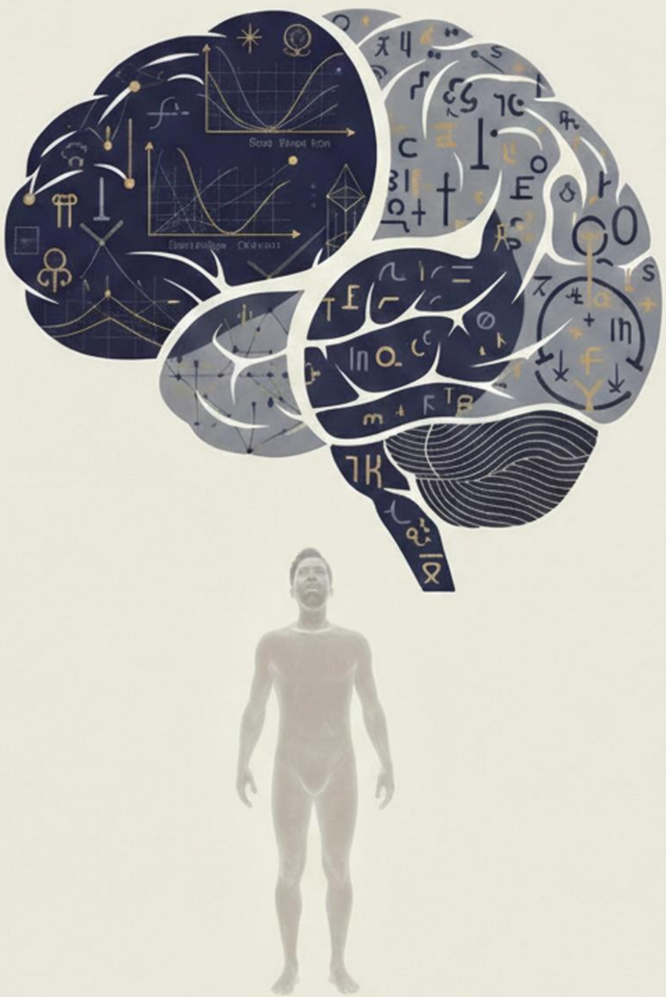
# The Mind-Body Revisited

## Is the Brain a Prediction Machine or an Interpretive System?

MARLON D. BULAQUEÑA   BRADFORD K. HALL

DEPARTMENT OF PHILOSOPHY, YALE UNIVERSITY
New Haven, CT 06520, United States



marlon.bulaquena@yale.edu

WHY PREDICTION ALONE CANNOT EXPLAIN UNDERSTANDING, MEANING, AND MIND.

# ABSTRACT

Recent developments in neuroscience and artificial intelligence have renewed the mind–body problem in a distinctive form. Predictive processing models portray cognition as hierarchical probabilistic inference aimed at minimizing prediction error, offering a powerful unifying framework for perception, action, and learning. While empirically successful, this paper argues that predictive processing cannot, by itself, account for understanding, meaning, or normativity. The difficulty is not empirical but explanatory: prediction regulates neural activity without constituting cognition as a norm-governed phenomenon. Drawing on embodied and interpretive approaches, the paper diagnoses a recurring explanatory overreach in attempts to treat prediction as sufficient for mind. Embodiment is shown to ground the practical stakes of prediction without collapsing cognition into sensorimotor coupling, while interpretation is characterized as responsiveness to publicly assessable norms rather than internal symbolic decoding. Contemporary artificial intelligence systems are examined as limiting cases that instantiate predictive optimization without genuine understanding, thereby clarifying the explanatory ceiling of prediction-based models. The paper advances a form of explanatory pluralism according to which neural regulation, embodied engagement, and interpretive normativity address distinct explanatory questions. On this view, the mind–body problem is reframed not as an ontological puzzle, but as a problem of explanatory scope. Predictive processing remains indispensable to cognitive science, but its philosophical ambitions must be constrained if cognition is to be adequately understood.



KEYWORDS: predictive processing; philosophy of mind; embodiment; interpretation; artificial intelligence; normativity

# Chapter 1 The Return of the Mind–Body Problem as an Explanatory Crisis

## 1.1 Purpose and Form of the Chapter

This chapter sets out the argumentative core of the treatise. Its aim is to show, by conceptual argument and careful counterexample, that the contemporary resurgence of the mind–body problem is best understood as an **explanatory crisis**: a conflict about the proper scope and limits of explanation in cognitive science rather than a renewed metaphysical dispute about substances. The chapter proceeds by (1) defining the central concepts, (2) stating the main thesis in argumentative form, (3) offering a set of supporting arguments, (4) considering the most pressing objections, and (5) drawing methodological consequences for subsequent chapters.

## 1.2 Definitions and Conceptual Groundwork

**Prediction.** A process that generates expectations about future or unobserved states and adjusts internal parameters to reduce discrepancy between expectation and input.

**Predictive processing (PP).** The theoretical framework that models cognition as hierarchical probabilistic inference whose operative principle is prediction error minimization.

**Understanding.** A normative cognitive state characterized by grasp of reasons, capacity for justification, and susceptibility to standards of correctness.

**Explanatory domain.** A class of questions and answers that share a common form of explanation (e.g., mechanistic, functional, normative).

These definitions are deliberately minimal. The argument that follows does not depend on controversial technicalities about Bayesian updating or on a particular formalization of prediction. It depends only on the contrast between **mechanistic regulation** (how systems adjust) and **normative explanation** (why certain states count as correct or justified).

### 1.3 Statement of the Main Thesis

**Main Thesis.** Predictive processing, while explanatorily indispensable for describing neural regulation, cannot by itself account for understanding, meaning, or normative cognitive status. The contemporary mind–body problem is therefore an explanatory crisis: a dispute about whether a single explanatory idiom—prediction error minimization—can legitimately be extended from mechanistic description to normative constitution.

This thesis breaks into two linked claims:

1. **Necessity claim.** Prediction is necessary for many cognitive functions; neural regulation requires predictive mechanisms.

2. **Insufficiency claim.** Prediction is insufficient to ground normativity, semantic content, and the justificatory relations that constitute understanding.

The remainder of the chapter defends these claims and explains why their conjunction reframes the mind–body problem.

### 1.4 Argument from Conceptual Role

**Premise 1.** Normative cognitive states are defined by their susceptibility to standards of correctness and justification.

**Premise 2.** Prediction error minimization is a statistical and mechanistic process that optimizes fit between model and input.

**Premise 3.** Statistical optimization does not, by itself, instantiate standards of correctness that are publicly assessable and reason-governed.

**Conclusion.** Therefore, prediction error minimization cannot, by itself, instantiate the normative features that define understanding.

**Support and Elaboration.** The conceptual role of understanding includes being able to give reasons, to be criticized, and to revise beliefs in light of justificatory standards. These features are not merely behavioral dispositions; they are normative relations among agents, propositions, and practices. A system that minimizes prediction error may change its internal parameters in response to input, but this change is not thereby a response to reasons. The system is not thereby answerable to standards of correctness; it is merely adjusting to statistical regularities.

**Illustrative contrast.** Consider two systems that produce the same output in a linguistic test: a human who can explain why she used a particular expression and a statistical model that produced the same expression because it maximized likelihood. The outputs are behaviorally indistinguishable, but only the human output is embedded in justificatory relations. The difference is not a difference in causal history alone; it is a difference in explanatory category.

## 1.5 Argument from Public Norms and Social Scaffolding

**Premise 1.** Normative statuses such as meaning and correctness are constituted within public practices and social institutions.

**Premise 2.** Predictive mechanisms operate within individual systems and are not sufficient to instantiate public practices.

**Conclusion.** Predictive mechanisms require embedding in social and embodied practices to acquire normative force.

**Support and Elaboration.** Meaning is not a private mapping between internal states and external referents; it is a capacity that emerges within shared practices of use, correction, and assessment. A predictive mechanism can be coupled to a social environment and thereby participate in normative practices, but the normative status is not produced by prediction alone. The social scaffolding supplies the standards by which outputs are judged correct or incorrect. Thus embodiment and sociality are necessary contextual constraints that give predictive processes their practical stakes.

### 1.6 The Artificial Intelligence Control Case

**Thesis.** Contemporary AI systems instantiate large-scale predictive optimization without manifesting normative understanding; they therefore serve as control cases that reveal the explanatory ceiling of prediction.

**Argument.**

- Modern AI systems, especially large language models, are trained to minimize loss functions that operationalize prediction error.

- These systems can produce behavior that mimics understanding: they answer questions, generate explanations, and solve problems.

- Yet they lack the capacity to be **answerable** to reasons in the normative sense: they cannot participate in practices of justification, acknowledge error in a way that presupposes normative accountability, or be held responsible for their outputs in the way agents are.

- Therefore, behavioral competence produced by prediction does not entail normative cognition.

**Philosophical significance.** The AI control case shows that predictive optimization is compatible with the absence of normativity. If prediction were sufficient for understanding, then AI systems that implement prediction at scale would instantiate understanding. They do not. Hence prediction is insufficient.

## 1.7 Objections and Replies

**Objection 1.** The distinction between regulation and constitution is illusory; normative features supervene on mechanistic processes and so are in principle reducible to prediction.

**Reply.** Supervenience does not imply reductive explanation. Even if normative properties supervene on physical processes, explaining why a state counts as correct requires normative vocabulary and public criteria. A reductive account that translates normative vocabulary into statistical terms must show how justificatory relations emerge from mechanistic dynamics. The burden is on reductionists to provide a principled account of this translation; mere supervenience is insufficient.

**Objection 2.** Predictive models can be enriched to include representations of norms and social practices; thus prediction can, in principle, capture normativity.

**Reply.** Enrichment is possible, but two points remain. First, adding representational content or social variables to a predictive model changes the explanatory domain: one is no longer explaining neural regulation alone but modeling social practices. Second, modeling a practice is not the same as participating in it. A model that represents norms does not thereby become answerable to them. The explanatory move from representation to normative status requires independent justification.

**Objection 3.** The insistence on public normativity privileges human sociality and risks excluding nonhuman cognition.

**Reply.** The argument does not deny that nonhuman animals can participate in normative practices; it insists that normativity is constituted by relations that are publicly assessable within a community. Where such communities exist, normative statuses can arise. The claim is methodological: explanations of cognition that aim to account for normativity must attend to the social and embodied contexts that instantiate standards.

## 1.8 Methodological Consequences

**Principle 1.** Explanatory pluralism. Different explanatory domains—mechanistic, embodied, normative—are required to answer different questions about cognition.

**Principle 2.** Constraint on extrapolation. Empirical success within one domain (e.g., neural regulation) does not license metaphysical or normative conclusions about cognition without additional argument.

**Principle 3.** Use of limiting cases. Philosophical analysis should employ limiting cases, such as AI systems, to test the sufficiency of explanatory frameworks.

These principles guide the remainder of the treatise. They recommend a disciplined approach: accept predictive processing where it succeeds, but resist the temptation to let its success dictate answers to questions it was not designed to address.

# Chapter 2 The Predictive Brain Thesis: Reconstruction, Reach, and Limits

## 2.1 Aim and Strategy of the Chapter

I reconstruct the predictive processing (PP) framework in philosophical detail, isolates the explanatory commitments that proponents commonly endorse, and evaluates the framework's scope. The goal is not to adjudicate empirical disputes but to clarify what PP claims when it is presented as a theory of cognition, to expose the inferential steps by which it is sometimes extended into philosophical conclusions, and to identify the precise points at which those extensions require additional conceptual resources.

## 2.2 A Minimal Reconstruction of Predictive Processing

**Core claim of PP.** The brain is a hierarchical inference engine that generates probabilistic predictions about sensory input and updates internal models by minimizing prediction error.

Three constitutive elements.

- **Generative models.** Internal structures that encode expectations about the causes of sensory data.

- **Prediction error.** The discrepancy between predicted and actual sensory input that drives updating.

- **Precision weighting and active inference.** Mechanisms that modulate the influence of prediction error and deploy action to bring sensory input into line with expectations.

**Functional description.** On PP, perception is not passive reception but hypothesis testing; action is a means of reducing error by changing the world; learning is the gradual refinement of generative models. These elements together form a compact computational vocabulary that can be applied across sensory modalities, motor control, attention, and learning.

### 2.3 What PP Explains Well

Explanatory successes.

- **Perceptual phenomena.** Illusions and context effects are naturally modeled as the influence of prior expectations on sensory interpretation.
- **Sensorimotor coordination.** Motor commands can be understood as predictions about proprioceptive consequences, with action minimizing proprioceptive prediction error.
- **Adaptive learning.** Bayesian updating provides a principled account of how organisms revise expectations in light of evidence.

**Why these successes matter.** PP supplies a unifying mechanism that links perception, action, and learning under a single formal principle. This parsimony is philosophically attractive because it promises to reduce disparate phenomena to a common explanatory idiom.

## 2.4 From Mechanism to Metaphysics The Inference Step

**The problematic inference.** Many proponents move from the claim that PP *describes* neural dynamics to the stronger claim that PP *constitutes* cognition. This move has three distinct steps that require scrutiny:

1. **Descriptive adequacy.** PP accurately models neural processes.

2. **Functional generalization.** The same computational principles apply across cognitive domains.

3. **Constitutive identification.** The computational principle is taken to be the essence of cognition.

**Where the inference is vulnerable.** Steps (1) and (2) are empirical and methodological; step (3) is philosophical. The vulnerability lies in treating a successful modeling strategy as a metaphysical identity claim without showing how mechanistic optimization yields normative and semantic properties.

## 2.5 Two Distinctions That Matter

### Distinction A: Explanation versus Description.

- **Description** records how systems behave.
- **Explanation** accounts for why behaviors count as cognitive, meaningful, or rational. PP excels at description and mechanistic explanation but does not by itself provide the normative vocabulary required to explain meaning.

### Distinction B: Functional Role versus Normative Status.

- **Functional role**: the causal contribution a state makes to behavior.

- **Normative status**: whether a state is correct, justified, or a reason. A state can have a functional role without possessing normative status; PP models functional roles but does not automatically confer normative status.

## 2.6 Thought Experiments and Diagnostic Cases

**Case 1: The Perceptual Twin.** Imagine two systems, A and B, that produce identical perceptual reports under identical conditions. A is a human with capacities for explanation and correction; B is a PP-driven device that minimizes prediction error. Behaviorally indistinguishable outputs do not guarantee identical explanatory status. The human's report is embedded in justificatory practices; the device's report is a product of optimization.

**Case 2: The Normless Learner.** Consider a learning system that optimizes prediction error in a closed environment with no social feedback. It acquires reliable mappings between stimuli and responses but has no access to public standards of correctness. Its internal models are effective but lack the normative relations that make beliefs count as true or justified.

**Diagnostic lesson.** These cases show that PP can account for reliability and adaptivity without thereby accounting for normativity or meaning.

## 2.7 Formalizing the Limit: What Prediction Cannot Do Alone

**Three formal limits.**

- **Aboutness.** Prediction encodes correlations between internal states and sensory patterns; it does not by itself establish that those states are *about* worldly entities in the normative sense.

- **Justification.** Bayesian updating prescribes how to revise probabilities given evidence; it does not by itself provide reasons that are publicly assessable or that ground normative criticism.

- **Answerability.** Minimizing loss functions does not make a system answerable to reasons; it makes it responsive to gradients of error.

**Philosophical upshot.** The formal apparatus of PP supplies constraints on how internal states change, but it does not supply the conceptual resources to transform causal responsiveness into normative status.

## 2.8 Responses from Within the PP Community and Assessment

**Common PP replies.**

- **Enrichment reply.** Normativity can be modeled as higher-level priors or as social priors encoded in generative models.

- **Continuity reply.** Normative phenomena supervene on predictive dynamics; there is no categorical gap.

Assessment.

- **Enrichment is not elimination.** Adding priors that represent norms changes the explanatory target: one now models social practices rather than deriving normativity from bare prediction.

- **Supervenience is not explanation.** Showing that normative states supervene on predictive dynamics leaves open the question of why those states should be treated as normative rather than merely functional.

**Consequence.** PP can be extended to model features associated with normativity, but such extensions presuppose the very normative frameworks they aim to explain.

# III. The Overreach Diagnosis: From Regulation to Constitution

### 3.1 Aim and Overview

This chapter advances the central negative diagnosis of the treatise: **predictive processing (PP) commits explanatory overreach when it is taken from a theory of neural regulation to a constitutive theory of cognition**. The aim is to make that diagnosis precise, to show why the move is philosophically illicit, and to map the minimal additional resources that would be required to make the constitutive claim plausible. The chapter proceeds by (1) isolating the regulative claims of PP, (2) articulating the constitutive claims that are sometimes inferred, (3) drawing the decisive conceptual hinge between them, (4) offering diagnostic thought experiments and control cases, and (5) canvassing and rebutting the most promising replies.

### 3.2 What Predictive Processing Actually Claims About Regulation

**Descriptive core.** At its core PP describes how neural systems manage uncertainty: they instantiate hierarchical generative models, compute prediction errors, and adjust synaptic and network parameters so as to reduce those errors. This is a description of **regulative mechanisms**—processes that govern dynamics, stability, and adaptivity.

**Typical regulative phenomena explained by PP.**

- **Perceptual inference:** priors bias sensory interpretation.
- **Motor control:** actions enacted to minimize proprioceptive prediction error.
- **Learning:** synaptic change driven by error-driven updating.
- **Attention and precision weighting:** modulation of error signals according to estimated reliability.

These are **mechanistic** claims about how neural systems maintain functional organization and respond to perturbation. They explain how systems stay within viable bounds and how behavior is tuned to environmental regularities.

### 3.3 The Constitutive Claim and Why It Tempts Philosophers

**Constitutive claim.** Some proponents and interpreters move from the regulative description to the stronger claim that **prediction error minimization is what cognition essentially is**—that cognition, understanding, and meaning are constituted by predictive dynamics.

**Why the move is tempting.**

- **Unificatory elegance:** a single principle that covers perception, action, and learning is philosophically attractive.

- **Empirical reach:** PP models fit a wide range of data, encouraging generalization.

- **Formal expressiveness:** Bayesian formalisms can be extended to model higher-level cognition, suggesting continuity from low to high levels.

But the temptation to identify regulation with constitution is precisely what this chapter challenges.

## 3.4 Regulative Mechanisms Versus Constitutive Conditions

**Key distinction.**

- **Regulative mechanisms.** Processes that govern neural dynamics and behavior; they answer questions about *how* systems maintain stability, adapt, and produce reliable outputs. Examples: error minimization, homeostatic control, synaptic plasticity.

- **Constitutive conditions.** Features that make a process genuinely cognitive in the normative, semantic, or rational sense; they answer questions about *what* cognition is and *why* certain states count as understanding, meaning, or reasons. Examples: semantic aboutness, justificatory relations, normative accountability.

**Why the distinction matters.** Explanations that identify how a system is regulated do not automatically supply the normative vocabulary required to say that the system *understands* or *means* something. The two explanatory targets are different in kind: one is mechanistic and causal, the other normative and evaluative.

### 3.5 The Decisive Hinge: Regulation Does Not Entail Constitution

**Central argumentative move.** The decisive hinge of the chapter is the following conditional:

- **If** an account only supplies regulative mechanisms (how neural states change to reduce error),

- **then** it does not, without further normative or social resources, establish that those states possess constitutive cognitive properties (meaning, understanding, reason-responsiveness).

Support for the hinge.

- **Category difference.** Normativity involves standards of correctness that are publicly assessable; regulation involves responsiveness to gradients of error that are not themselves normative.

- **Explanatory gap.** Mechanistic optimization explains reliability and adaptivity but leaves open why a state should be treated as a reason or as true rather than merely effective.

- **Practical asymmetry.** Systems can be highly regulated and behaviorally competent while remaining non-answerable to reasons (as in many engineered systems).

This hinge is not a metaphysical denial of dependence; it is an epistemic and explanatory claim about what kinds of resources are required to ground normative statuses.

### 3.6 Diagnostic Thought Experiments

**Thought Experiment 1 The Mimetic Agent** Imagine an agent whose internal dynamics are governed by PP and which reliably produces human-like utterances and actions. The agent passes behavioral tests for understanding. Yet it has no history of participating in practices of justification, no capacity to give reasons, and no sensitivity to public standards. The thought experiment shows that behavioral parity does not guarantee normative parity.

**Thought Experiment 2 The Isolated Predictor** Consider a biological organism that, through evolution, develops highly effective predictive mechanisms for navigating a stable niche but never engages in social practices that instantiate norms. Its predictive machinery yields reliable behavior but there is no community to assess correctness. The organism's internal models are adaptive but not normative in the sense required for understanding.

**Diagnostic conclusion.** Both cases illustrate that prediction can secure adaptivity and reliability without securing the constitutive features of cognition.

### 3.7 The Artificial Intelligence Control Case Revisited

**AI as a control case.** Contemporary AI systems—large language models and deep networks—implement large-scale predictive optimization and can mimic many cognitive outputs. They thereby instantiate the regulative side of PP at scale.

**What AI reveals.**

- **Behavioral competence without answerability.** These systems can produce explanations and solutions but lack the capacity to be held to reasons in the normative sense.

- **No intrinsic normativity.** Their outputs are not subject to justificatory practices except insofar as humans impose such practices externally.

**Philosophical import.** If prediction were sufficient for constitutive cognition, then AI systems that implement prediction at scale would instantiate understanding. Their failure to be normative agents indicates that prediction alone is insufficient.

### 3.8 Common Replies from PP Advocates and Replies

**Reply A Enrichment Reply** PP can be enriched: higher-level priors can encode norms, social practices can be internalized as priors, and so prediction can, in principle, model normativity.

**Counter.** Enrichment changes the explanatory target. Modeling norms is not the same as showing that prediction *constitutes* normativity. Moreover, representing a norm inside a model does not by itself make the model answerable to that norm in the public, practice-based sense.

**Reply B Continuity Reply** There is continuity from low-level prediction to high-level cognition; normative features supervene on predictive dynamics.

**Counter.** Supervenience does not dissolve the explanatory gap. To show that normative features are not merely supervenient epiphenomena, one must provide an account of how justificatory relations and public standards arise from dynamics—an account that PP, in its standard form, does not supply.

**Reply C Pragmatic Reply** Normativity is a pragmatic byproduct of successful prediction: organisms that predict well are treated as correct by others, and norms emerge from coordination.

**Counter.** This is plausible as a sociological account of how norms arise, but it again locates normativity outside bare predictive dynamics. It confirms the chapter's point: prediction needs social and embodied scaffolding to acquire normative force.

### 3.9 Minimal Resources Required for a Constitutive Claim

If one wishes to defend the stronger constitutive claim—that prediction *is* cognition—then at minimum one must supply:

- **An account of aboutness.** How do internal predictive states acquire semantic content that is not merely correlational?

- **An account of justificatory relations.** How do predictive updates instantiate reasons that are publicly assessable?

- **An account of social embedding.** How do individual predictive mechanisms participate in practices that instantiate norms rather than merely track regularities?

Absent these resources, the constitutive claim remains an unjustified extrapolation from mechanistic success.

### 3.10 Implications for Theory Building

**Methodological caution.** The diagnosis counsels restraint: accept PP as a powerful regulative theory while resisting the temptation to let its success dictate metaphysical conclusions about cognition.

**Programmatic suggestion.** Productive theory building should aim for **explanatory integration**: show how predictive mechanisms interact with embodied constraints and social practices to produce the phenomena we call understanding. Integration requires explicit bridging principles, not rhetorical assimilation.

Predictive processing provides indispensable explanations of neural regulation but, by itself, does not constitute cognition in the normative, interpretive sense. The decisive hinge between regulation and constitution remains unbridged by PP alone. To move from regulation to constitution one must supply independent accounts of semantic aboutness, justificatory relations, and social embedding. The remainder of the treatise will examine how embodiment and interpretive practices can supply those resources without collapsing into reduction or mysticism.

# IV. Why Prediction Cannot Ground Meaning or Understanding

## 4.1 Aim and Overview

This chapter argues that **prediction error minimization**, as a statistical and mechanistic process, cannot by itself generate the normative features that constitute **meaning** and **understanding**. The argument proceeds by (1) clarifying the normative features at stake, (2) showing the conceptual mismatch between statistical optimization and normativity, (3) isolating three specific deficits—aboutness, epistemic responsibility, and rational accountability—and (4) responding to the most plausible rejoinders. The conclusion is that predictive mechanisms are necessary for many cognitive functions but insufficient to ground the normative status that makes states genuinely cognitive.

## 4.2 Normativity: What Understanding Requires

**Normative features of understanding.** To say that an agent *understands* is to attribute to it capacities that go beyond mere reliable performance. Three interrelated normative features are central:

- **Correctness.** A cognitive state can be evaluated as correct or incorrect relative to standards that are publicly articulable.
- **Justification.** Cognitive states can be supported by reasons; agents can cite evidence or principles that make a state warranted.
- **Answerability.** Agents can be held to account for their beliefs and actions; they can respond to criticism and revise in light of reasons.

These features are not merely descriptive—they are evaluative and practice-dependent. Understanding involves being embedded in networks of justification and assessment that make claims true or false, warranted or unwarranted.

### 4.3 The Mechanistic Character of Prediction Error

**What prediction error is.** Prediction error minimization is a formal, statistical procedure: internal models generate expectations; discrepancies between expectation and input produce error signals; parameters are adjusted to reduce future error. This process is characterized by:

- **Statistical optimization.** The objective is to reduce a loss function or free energy, not to satisfy publicly articulated standards.
- **Causal responsiveness.** Changes occur because of causal pressures (error gradients), not because of recognition of reasons.
- **Instrumental success.** The mechanism aims at adaptive fit, not at justificatory relations.

The formal apparatus is powerful for explaining reliability and adaptivity, but its primitives are numerical and causal rather than normative.

### 4.4 The Three Deficits: What Error Minimization Cannot Produce

**Crucial claim.** No amount of error minimization, by itself, can generate the following normative features.

#### 4.4.1 Semantic Aboutness

- **Problem.** Prediction encodes correlations between internal states and sensory patterns; it does not, by itself, confer *aboutness*—the property of being about worldly entities or states of affairs in a way that supports truth-conditions.

- **Why this matters.** Aboutness underwrites semantic content: to say a belief is about X is to place it within a network of referential relations and truth-conditions that go beyond mere covariance.

- **Illustration.** A predictive model may reliably activate an internal pattern when presented with images of dogs, but this covariance does not explain why the internal pattern should be taken to *refer* to dogs rather than merely correlate with dog-images.

### 4.4.2 Epistemic Responsibility

- **Problem.** Error minimization prescribes how to update parameters to reduce loss; it does not prescribe how an agent ought to respond to reasons or how it is to be held responsible for its beliefs.

- **Why this matters.** Epistemic responsibility involves being answerable to standards of evidence and justification—practices that presuppose agents who can give and assess reasons.

- **Illustration.** A system that updates its weights when confronted with counterevidence is not thereby responsible; responsibility presupposes capacities for justification, explanation, and acknowledgement of error in normative terms.

### 4.4.3 Rational Accountability

- **Problem.** Rational accountability requires that agents' inferences be assessable against norms of reasoning (consistency, coherence, relevance). Minimizing prediction error optimizes performance but does not instantiate these normative standards.

- **Why this matters.** To hold an agent rationally accountable is to demand reasons and to evaluate whether those reasons meet communal standards; optimization alone cannot supply those standards.

- **Illustration.** Two systems may reach the same conclusion; one does so by rule-governed inference that can be inspected and criticized, the other by opaque optimization. Only the former is straightforwardly subject to rational appraisal.

## 4.5 Diagnostic Cases and Intuitions

**Case A: The Perfect Predictor.** Imagine a device that, through exhaustive training, predicts human responses to moral dilemmas with near-perfect accuracy. It reproduces human judgments but cannot explain them, cannot revise in light of reasons, and cannot be held accountable. Intuitively, it does not *understand* morality; it merely reproduces patterns.

**Case B: The Social Novice.** Consider an organism with excellent predictive capacities but no history of participating in social practices that instantiate norms. Its internal models guide behavior effectively, yet there is no community to evaluate correctness. Its states lack the public normative status that constitutes understanding.

These cases show that predictive success and normative status can come apart.

### 4.6 Replies from Predictive Advocates and Responses

**Reply 1: Norms as Priors.** One can encode norms as high-level priors; once norms are internalized, prediction can represent and enforce them.

- **Response.** Encoding norms inside a model is modeling the *representation* of norms, not showing how normative status arises. A model that contains a representation of a norm is not thereby answerable to that norm in the public, practice-based sense. The move from representation to normative standing requires social and institutional embedding.

**Reply 2: Continuity and Supervenience.** Normativity supervenes on predictive dynamics; there is no categorical gap.

- **Response.** Supervenience is compatible with an explanatory gap. Even if normative properties depend on physical processes, explaining why a state counts as normative requires normative vocabulary and public criteria. Supervenience alone does not provide the justificatory story.

**Reply 3: Pragmatic Emergence.** Norms emerge pragmatically from coordination among predictive agents; prediction is the engine of norm formation.

- **Response.** This account locates normativity in social processes that are external to bare predictive dynamics. It confirms the chapter's point: prediction contributes to, but does not by itself constitute, normativity. The normative force arises in the social practices of coordination, correction, and justification.

Predictive mechanisms are indispensable for explaining how organisms track regularities, adapt to environments, and produce reliable behavior. They explain *how* internal states change in response to error and how behavior becomes tuned to ecological demands. But **meaning** and **understanding** are normative phenomena: they require semantic aboutness, epistemic responsibility, and rational accountability—features that are not captured by statistical optimization alone. To account for these features we need additional explanatory resources: accounts of reference and content, analyses of justificatory practices, and theories of social embedding that make normative statuses publicly assessable. Prediction is a necessary component of cognitive explanation; it is not, on its own, a sufficient foundation for meaning or understanding.

# V. Embodiment Reconsidered: Necessary but Not Sufficient

## 5.1 Aim and Overview

This chapter defends a middle course between two extremes. It accepts the enactivist insistence that cognition is **situated, bodily, and action-oriented**, while resisting the claim that embodiment alone displaces predictive explanation or that embodiment by itself supplies the normative resources of meaning and understanding. The central claim is simple and programmatic: **embodiment supplies the contextual constraints that give predictive processes practical stakes, but it does not by itself convert prediction into meaning.**

## 5.2 Engagement with Enactivism

**Enactivist insight.** Enactivist and embodied approaches correctly diagnose a persistent internalist bias: treating cognition as detached symbol manipulation or as purely intracranial computation obscures the role of bodily structure, sensorimotor skills, and environmental coupling.

Enactivism restores attention to capacities that are exercised in action, to the ways bodies shape perception, and to the primacy of practical know-how in many cognitive achievements.

What enactivism contributes.

- **Sensorimotor mastery.** Perception is skillful activity, not passive reception.
- **Constitutive coupling.** Cognitive systems are constituted through ongoing engagement with an environment that affords possibilities for action.
- **Practical normativity.** Many normative assessments (e.g., success in grasping, navigating, or tool use) are grounded in bodily competence and ecological fit.

These contributions are indispensable for a full account of cognition; they correct distortions that arise when theorists focus exclusively on internal computation.

### 5.3 Clarified Position: Embodiment and Prediction Are Complementary

**Not an alternative.** Embodiment is not offered here as a rival to predictive processing. Rather, it is a **contextual constraint**: the body and its ecological niche determine which predictions matter, how priors are formed, and which errors are salient.

**How the relation works.**

- **Constraint on priors.** Bodily morphology and sensorimotor repertoire shape the space of plausible generative models.
- **Constraint on precision.** Affordances and bodily reliability influence precision weighting—what counts as noise versus signal.

- **Constraint on action selection.** Goals grounded in bodily needs and vulnerabilities determine which prediction-driven actions are pursued.

In short, embodiment supplies the pragmatic parameters within which predictive mechanisms operate.

### 5.4 Bodies Make Prediction Matter: Goals, Affordances, Vulnerabilities

**Goals.** Bodily needs and capacities instantiate ends that predictions aim to secure (e.g., homeostasis, locomotion, manipulation). Predictions are not neutral statistical artifacts; they are directed toward maintaining bodily viability and achieving embodied goals.

**Affordances.** The environment presents affordances relative to bodily capabilities. Predictive models that ignore affordances will misprioritize errors; models that incorporate affordances better align predictions with actionable possibilities.

**Vulnerabilities.** Bodily fragility and ecological constraints make certain errors costly. The salience of particular prediction errors is therefore a function of embodied risk and opportunity.

These three factors—goals, affordances, vulnerabilities—explain why some prediction errors are prioritized and why organisms act to reduce them. Embodiment thus supplies the *practical stakes* that make prediction instrumentally significant.

### 5.5 Limits of Embodiment: Why Bodies Do Not Make Prediction Meaningful

**Normativity is not automatic.** Although bodies ground practical significance, they do not by themselves instantiate the **normative structures**—public standards, justificatory practices, and semantic relations—that constitute meaning and understanding.

**Three ways embodiment falls short.**

- **Private pragmatics.** Bodily needs explain why an organism cares about certain outcomes, but caring is not the same as being answerable to reasons that others can assess.

- **Local competence versus public correctness.** Skillful action can be evaluated for success or failure, but success in action does not automatically yield semantic content or truth-conditions for beliefs.

- **No intrinsic justificatory relations.** Bodily constraints shape what is adaptive, not what counts as a justified belief in a community of inquirers.

Thus embodiment supplies *why* prediction matters instrumentally; it does not supply *why* a representational state should be treated as true, about something, or justified in the normative sense.

### 5.6 Integration: How Embodiment and Prediction Interact Without Collapse

Complementary explanatory roles.

- **Predictive processing** explains the **mechanics** of how internal states are adjusted and how action is selected to minimize error.

- **Embodiment** explains the **pragmatics** of why certain errors are salient and which actions are possible or desirable.

Bridging principles (programmatic).

1. **Embodied priors.** Bodies constrain the form and content of priors that predictive systems instantiate.

2. **Affordance-weighted precision.** The body's action possibilities modulate precision assignments in predictive hierarchies.

3. **Social extension.** When embodied agents participate in social practices, predictive models can come to represent norms—but representation of norms is not identical to normative standing.

These principles show how the two domains can be integrated without reducing one to the other: embodiment informs the structure and deployment of prediction; prediction implements regulation within embodied constraints.

### 5.7 Diagnostic Examples

**Example 1 Tool Use.** A person learning to use a hammer develops predictive models tuned to the affordances of the tool and the body's dynamics. Embodiment explains why certain predictions are prioritized (grip, force); prediction explains how motor commands are refined. Yet the meaning of "hammer" and the justificatory reasons for using it in a given context require social and linguistic practices beyond mere sensorimotor coupling.

**Example 2 Navigating a Cliff.** An animal's predictive system prioritizes errors that signal imminent falls because bodily vulnerability makes such errors costly. Embodiment explains the salience; prediction explains the rapid corrective responses. But the animal's avoidance behavior does not by itself instantiate propositional content or justificatory discourse.

## 5.8 Objections and Replies

**Objection 1.** If embodiment cannot supply normativity, then enactivism fails to account for higher cognition.

**Reply.** Enactivism need not be expected to supply full normative accounts on its own; its role is to show how bodily and ecological factors ground practical significance. Higher cognitive normativity requires additional social and interpretive resources that build on, but are not reducible to, embodiment.

**Objection 2.** Treating embodiment as merely contextual risks downgrading its explanatory importance.

**Reply.** The claim is not that embodiment is secondary; it is central. Calling embodiment a contextual constraint emphasizes its constitutive role in shaping predictive architecture and behavioral priorities while preserving the need for further resources to explain meaning.

Embodiment is indispensable: bodies shape priors, determine affordances, and make certain prediction errors salient by virtue of goals and vulnerabilities. But embodiment alone does not generate the normative structures—aboutness, justification, answerability—that constitute meaning and understanding. The proper theoretical posture is integrative: predictive mechanisms operate within embodied contexts that give them practical force, and normative, interpretive practices must be added to account for semantic and justificatory dimensions. The next chapter turns to those interpretive practices and asks how public norms and social scaffolding can supply the missing normative resources without collapsing into homuncular explanations.

## VI. Interpretation as a Distinct Explanatory Category

### Conceptual upgrade

**Interpretation** should not be reduced to inner decoding, symbolic manipulation, or mere storage of representations. Instead, interpretation is best understood as **responsiveness to publicly assessable, norm-governed standards of correctness**—a form of cognitive activity whose proper explanation invokes evaluative criteria that transcend purely causal description. This reconceptualization shifts the explanatory burden: we no longer ask only how internal states covary with inputs, but how those states become subject to standards that permit correction, justification, and dispute.

### Features of interpretation

- **Socially scaffolded.** Interpretation depends on practices, language, and shared norms. Meaning emerges within communities of practice where uses, corrections, and conventions stabilize semantic roles and justificatory standards.

- **Temporally extended.** Interpretive capacities unfold across time: they are learned, transmitted, and refined through histories of practice. A judgment's status as correct or incorrect often rests on its place in a temporal sequence of reasons, tests, and revisions.

- **Publicly assessable.** Interpretive states are open to evaluation, correction, and justification by others. This public assessability is what makes normative claims intersubjective rather than private correlations.

### Why this category matters

Treating interpretation as a distinct explanatory category accomplishes three philosophical tasks at once.

1. **Blocks eliminativism.** If interpretation is reducible to mere causal regularities, then talk of meaning and justification becomes dispensable. Recognizing interpretation as norm-governed preserves the irreducibility of normative vocabulary: meanings and reasons are not epiphenomenal descriptions of causal patterns but practices with evaluative force.

2. **Blocks homuncular mentalism.** Interpretation is not an inner homunculus silently decoding symbols. Its normativity is constituted in public practices and temporal processes; positing an inner interpreter merely relocates the problem without explaining how normative standards arise or are sustained.

3. **Guides empirical and conceptual work.** By isolating interpretation as its own explanatory domain, researchers can ask targeted questions—How do social practices instantiate standards? How are justificatory norms learned and transmitted?—rather than conflating these questions with mechanistic accounts of neural regulation.

## Minimal explanatory commitments for interpretation

Any adequate account of interpretation must show how:

- **Public practices** instantiate standards that make correctness assessable;
- **Historical processes** produce stable norms and shared conceptual frameworks;
- **Agents** become answerable to reasons in ways that are not reducible to mere error gradients.

Absent these commitments, claims that prediction or embodiment alone suffice to explain meaning remain incomplete.

## VII. Artificial Intelligence as a Limiting Case, Not a Model

### Methodological move

Use contemporary artificial intelligence systems as **philosophical control cases** rather than as straightforward models of human cognition. Control cases test the sufficiency of explanatory frameworks: if a framework entails that certain capacities amount to understanding, then systems that instantiate those capacities should, on that framework, count as understanding. If they do not, the framework is incomplete.

### Observation

Modern AI systems instantiate large-scale predictive optimization. Trained on massive corpora and optimized for loss minimization, they achieve **high behavioral competence**—fluent language production, problem solving, pattern recognition—while lacking the hallmarks of interpretive understanding and normative answerability. Their outputs can be indistinguishable from human outputs in many tests, yet they do not participate in justificatory practices, cannot be held to reasons in the normative sense, and do not exhibit publicly assessable commitments.

### Diagnostic claim

AI systems reveal the **ceiling of predictive explanation**: behavioral success can be decoupled from cognitive answerability. From this control case three diagnostic lessons follow.

- **Necessity without sufficiency.** Prediction is necessary for many cognitive functions—perception, anticipation, motor coordination—but it is not sufficient for understanding. The presence of predictive optimization does not guarantee semantic aboutness or normative standing.

- **Separation of competence and answerability.** Competence (what a system can do) is separable from answerability (what a system can be asked to justify). Human cognition combines both; current AI systems instantiate the former without the latter.

- **Empirical constraint on theorizing.** Any theory that identifies prediction alone with cognition must explain why AI systems, which implement prediction at scale, fail to instantiate interpretive normativity. The burden is on such theories to supply bridging principles that are currently absent.

## Illustrative contrasts

- **Behavioral parity, normative disparity.** A language model can generate a grammatical explanation for a phenomenon but cannot engage in the social practices that would make that explanation subject to correction, endorsement, or repudiation in a community of inquirers.

- **Opaque optimization versus inspectable reasons.** An AI's internal optimization path may be opaque and uninterpretable; a human's reasons are, at least in principle, articulable and contestable.

## Implications for theory and practice

- **Theoretical humility.** Cognitive theories should treat AI as revealing limits of mechanistic accounts rather than as straightforward confirmation of philosophical claims about mind.

- **Research agenda.** To explain understanding, researchers must investigate how predictive mechanisms are embedded in social practices and how normative statuses are acquired and maintained—work that combines neuroscience, developmental psychology, social epistemology, and philosophy of language.

- **Ethical and practical caution.** Because AI can mimic understanding without possessing it, reliance on behavioral tests alone is epistemically risky when normative accountability matters (e.g., in legal, medical, or educational contexts).

AI functions as a revealing foil: it shows that **prediction scales** but **normativity does not** automatically follow. The philosophical task is to articulate the missing explanatory links—how social scaffolding, historical embedding, and justificatory practices transform regulated predictive activity into interpretive, norm-governed cognition.

## VII Artificial Intelligence as a Limiting Case, Not a Model

### Aim and Methodological Move

**Aim.** To use contemporary artificial intelligence systems as **philosophical control cases** that test the sufficiency of prediction-based explanations for cognition. The chapter argues that AI, properly construed, reveals the limits of predictive explanation: it shows what prediction can accomplish at scale and what it cannot—namely, the normative, interpretive features that constitute understanding.

**Methodological move.** Treat AI systems not as straightforward models of human minds but as **diagnostic tools**. A control case is useful when a theory entails that certain capacities amount to cognition; if the control case instantiates those capacities yet lacks other hallmark features of cognition, the theory is incomplete. Contemporary AI systems therefore function as empirical thought experiments: they instantiate large-scale prediction and allow us to observe which cognitive features follow and which do not.

## Observation: What Modern AI Instantiates

**Core observation.** Modern AI systems—large language models, deep neural networks, and other statistical learners—implement large-scale predictive optimization. They are trained to minimize loss functions that operationalize prediction error across vast datasets.

**Manifest capacities.**

- **Behavioral competence.** Fluent language production, pattern recognition, problem solving, and task performance that often match or exceed human benchmarks in narrow domains.

- **Statistical generalization.** Ability to interpolate and sometimes extrapolate from training data to novel inputs.
- **Apparent reasoning.** Generation of explanations, analogies, and solutions that mimic human-style reasoning.

Absent capacities.

- **Interpretive understanding.** No evidence that these systems possess grasp of meaning in the normative sense—they do not participate in justificatory practices or exhibit genuine semantic aboutness.
- **Normative answerability.** They are not answerable to reasons; they cannot acknowledge, justify, or revise beliefs in a way that presupposes normative accountability.

## Diagnostic Claim: The Ceiling of Predictive Explanation

**Diagnostic claim.** AI systems demonstrate the **ceiling of predictive explanation**: prediction at scale can produce behavioral competence without producing interpretive understanding or normative answerability. From this follows a threefold diagnostic lesson.

1. **Prediction is necessary but not sufficient.** Predictive mechanisms are indispensable for many cognitive functions—perception, anticipation, motor coordination, pattern detection—but their presence does not guarantee the emergence of meaning or justificatory relations.

2. **Competence is separable from answerability.** A system can be competent (able to perform tasks) without being answerable (able to give, assess, or be held to reasons). Human cognition typically couples competence with answerability; current AI systems do not.

3. **Theoretical burden on reductionists.** Any theory that identifies prediction alone with cognition must explain why AI systems, which implement prediction at scale, fail to instantiate interpretive

normativity. The burden is to supply bridging principles that are currently absent.

## Illustrative Contrasts and Cases

- **Behavioral parity, normative disparity.** A language model can produce a grammatical, persuasive explanation for a scientific claim. A human scientist can produce the same explanation and defend it in a community of peers. The model's output lacks the social and justificatory embedding that makes the human's explanation a candidate for acceptance, criticism, or revision.

- **Opaque optimization versus articulable reasons.** An AI's internal optimization path is often opaque and uninterpretable; a human's reasons are, at least in principle, articulable and contestable. Opacity undermines the possibility of normative appraisal.

- **Simulated endorsement versus genuine commitment.** AI can simulate endorsement of a claim by producing supportive arguments, but it does not exhibit commitment in the sense of being disposed to revise or defend its position within a practice of inquiry.

## Objections and Replies

**Objection 1 AI will eventually acquire normativity.** Future architectures could internalize social practices and thereby become answerable to reasons.

**Reply.** Two points follow. First, internalizing representations of norms is not the same as participating in the public practices that instantiate normativity. Second, even if future systems model social practices, the explanatory task remains: one must show how modeled representations become normative standings—how being represented inside a system yields public assessability and justificatory relations.

**Objection 2 AI's lack of answerability is an engineering gap, not a conceptual one.** Given the right training regimes and environments, AI will develop interpretive capacities.

**Reply.** This is an empirical hypothesis that requires demonstration. The conceptual point stands: showing that a system can be trained to behave as if it is answerable is not the same as showing it genuinely participates in normative practices. Philosophical analysis must distinguish between behavioral mimicry and normative membership.

**Objection 3 Using AI as a control case anthropomorphizes machines.** We should not expect machines to meet human standards of normativity.

**Reply.** The point of a control case is precisely to test theoretical sufficiency. If a theory implies that prediction suffices for cognition, then machines that implement prediction should instantiate cognition. Their failure to do so is a legitimate philosophical datum that constrains theory.

## Implications for Cognitive Theory and Research

**Theoretical humility.** Cognitive theories should acknowledge that predictive optimization explains important mechanistic features but does not, by itself, resolve questions about meaning and normativity.

**Research agenda.** To explain understanding, interdisciplinary work is required: developmental studies of how children acquire justificatory practices; social epistemology of how norms are established and transmitted; computational models that explicitly incorporate social interaction and public assessment; philosophical analysis of the conditions under which representation acquires normative force.

**Practical caution.** Because AI can mimic understanding without possessing it, reliance on behavioral tests alone is epistemically risky in contexts where normative accountability matters (education, law, medicine). Systems should not be treated as normative agents merely because they perform competently.

Contemporary AI systems function as revealing limiting cases: they instantiate prediction at scale and thereby expose what predictive explanation can achieve and what it cannot. The decoupling of behavioral competence from interpretive answerability in AI demonstrates that **prediction is a necessary component of cognition but not a sufficient foundation for meaning or understanding**. The philosophical task is to articulate the missing explanatory links—how social scaffolding, historical embedding, and justificatory practices transform regulated predictive activity into interpretive, norm-governed cognition.

## VIII. The Explanatory Pluralism Thesis Rewritten Precisely

This chapter restates the treatise's central positive proposal with precision. **Core claim.** Cognition is best explained by **distinct explanatory domains**—each answering different kinds of questions—rather than by a single hierarchical layering that attempts to reduce all cognitive phenomena to one principle. The appropriate explanatory domains are **neural regulation**, **embodied engagement**, and **interpretive normativity**. Each domain is necessary for a full account of cognition; none is by itself sufficient.

### Neural Regulation Defined

**Domain.** Neural regulation covers prediction, inference, and control. **What it explains.** Mechanistic dynamics, error minimization, synaptic plasticity, precision weighting, and the ways neural systems maintain adaptive organization. **Characteristic questions.** How do internal models update? How do error signals propagate? How are actions selected to reduce discrepancy? **Explanatory form.** Mechanistic and computational explanation grounded in physiology and formal models.

## Embodied Engagement Defined

**Domain.** Embodied engagement covers situated action, bodily constraint, and ecological vulnerability. **What it explains.** Why certain predictions are pragmatically salient, how affordances shape perception and action, and how bodily form and needs structure cognitive possibilities. **Characteristic questions.** Which predictions matter for survival and skill? How do bodily capacities constrain representational content? How do environmental affordances shape learning? **Explanatory form.** Functional and ecological explanation that links morphology, sensorimotor skill, and environmental structure.

## Interpretive Normativity Defined

**Domain.** Interpretive normativity covers meaning, reason, and justification. **What it explains.** Why cognitive states are subject to standards of correctness, how reasons and justificatory relations arise, and how public practices instantiate semantic content. **Characteristic questions.** What makes a belief correct? How are reasons given and assessed? How do social practices stabilize meaning? **Explanatory form.** Normative and social explanation that appeals to practices, language, and intersubjective standards.

## Relations Among Domains

**Not reducible.** Explanations in one domain do not logically or conceptually exhaust the questions posed in another. Mechanistic accounts of neural regulation do not by themselves answer why a state counts as justified. Embodied descriptions of affordances do not by themselves supply truth conditions. Interpretive accounts of normativity do not by themselves specify neural implementation.

**Not competitors in a zero-sum sense.** The domains are complementary rather than adversarial. A successful scientific program will deploy mechanistic, ecological, and normative explanations where each is most appropriate, and will seek principled connections among them rather than attempting to eliminate one in favor of another.

**Not ontologically disjoint.** The domains are explanatorily distinct but ontologically continuous. Neural processes, bodily structures, and social practices coexist in the same world and causally interact. Distinctness is methodological and conceptual, not a claim of separate substances.

### Principles for Integration

To make pluralism productive rather than merely pluralistic, adopt three integration principles.

- **Constraint Principle.** Explanations in one domain constrain hypotheses in the others. Neural models must respect embodied constraints; normative accounts must be compatible with what neural and embodied systems can realize.

- **Translation Principle.** Provide explicit bridging concepts when moving between domains. For example, specify how an affordance alters precision weighting, or how a public practice supplies criteria that a predictive model can represent.

- **Autonomy Principle.** Preserve domain-specific standards of evaluation. Mechanistic models are judged by empirical fit and causal adequacy; normative accounts are judged by coherence with practices and justificatory power.

## Methodological and Philosophical Payoff

Adopting explanatory pluralism yields several advantages.

- **Avoids category mistakes.** It prevents illegitimate moves from mechanistic description to normative constitution.

- **Preserves explanatory depth.** It allows each domain to contribute its best resources without collapsing into reduction or mysticism.

- **Guides empirical research.** It points to interdisciplinary programs that study how neural regulation, embodiment, and social practices interact rather than trying to subsume one under another.

The explanatory pluralism thesis reframes the mind–body problem as a problem of explanatory scope and integration. Cognition requires **neural regulation**, **embodied engagement**, and **interpretive normativity**—distinct explanatory domains that are mutually informing. Recognizing their distinctness and interdependence clarifies what predictive processing achieves and what additional resources are required to account for meaning, understanding, and rational agency.

## IX Reframing the Mind-Body Problem

### Decisive reorientation

The traditional framing of the mind-body problem treats it as a single metaphysical puzzle about how mental properties relate to physical substances. That framing invites three familiar but unhelpful responses: **reduction** (identify mental properties with physical ones), **vague emergence** (appeal to an ill-specified emergentism), and **ontological inflation** (postulate new kinds of substance or properties). Each of these responses attempts to solve the problem by changing the ontology. The decisive reorientation proposed here dissolves the puzzle instead: the apparent intractability arises because we have conflated **two fundamentally different explanatory projects**—causal explanation and normative explanation—and sought a single, monolithic account to do both.

- **Causal explanation** asks how processes are implemented, how they interact, and how they produce reliable behavior. It is the domain of neuroscience, computational modeling, and mechanistic biology.

- **Normative explanation** asks why certain states count as correct, justified, or reason-responsive. It is the domain of philosophy of language, social epistemology, and ethics.

Recognizing that these are distinct explanatory aims reframes the problem. The question becomes not which ontology subsumes the other, but how different explanatory practices relate and constrain one another. Once we accept that causal and normative explanations answer different questions, the demand that a single physical description exhaustively account for normative features is misplaced.

### Why this reorientation dissolves the puzzle

Three points make the dissolution persuasive.

1. **Category clarity.** Many paradoxes in the mind-body literature arise from category mistakes—treating normative predicates as if they were causal predicates. Distinguishing explanatory aims prevents such conflations.

2. **Explanatory sufficiency is domain-relative.** A successful causal explanation need not, and typically does not, provide the resources for normative assessment. Conversely, normative accounts presuppose causal facts but are not reducible to them.

3. **Integration without reduction.** Dissolution does not mean fragmentation. It permits a program of **explanatory integration** in which causal, embodied, and normative accounts are coordinated through explicit bridging principles rather than collapsed into a single metaphysical identity.

## Preservation of realism

This reframing preserves a robust scientific realism about the physical basis of cognition while rejecting the stronger claim that physical description alone exhausts cognitive explanation.

- **Realist commitment.** Neural processes, bodily structures, and environmental couplings are real and causally efficacious. Explanations in the neural regulation domain are indispensable for understanding how cognitive systems operate.

- **Anti-reductionist restraint.** Realism about the physical does not entail that physical description suffices to capture normative statuses. Normativity—meaning, justification, reason-responsiveness—emerges within social and historical practices that require their own explanatory vocabulary.

The result is a **non-reductive realism**: mental phenomena are grounded in the physical, but their full explanation requires distinct normative and social analyses that are continuous with, yet not reducible to, mechanistic description.

Practical and theoretical consequences

- **For philosophy.** The mind-body debate should shift from metaphysical combat to methodological clarification: specify which explanatory questions are being asked and choose appropriate conceptual tools.

- **For cognitive science.** Researchers should be explicit about the explanatory domain of their models and cautious about extrapolating mechanistic success into claims about meaning or agency.

- **For interdisciplinary work.** Productive collaboration requires explicit bridging principles that translate constraints across domains without illicitly importing domain-specific standards.

# X Why the Prediction-Machine Metaphor Must Be Abandoned

## Final claim

The metaphor of the brain as a **prediction machine** has been scientifically fruitful. It has generated testable models, unified disparate findings, and sharpened empirical programs. But when the metaphor is elevated into a metaphysical identity claim—that the brain literally *is* a prediction machine and that prediction exhausts cognition—it becomes philosophically misleading. The metaphor conflates a powerful regulative description with constitutive explanation and thereby obscures the normative and interpretive dimensions of mind.

## Why the metaphor misleads

- **Category slippage.** The metaphor encourages treating a mechanistic optimization procedure as if it carried normative force. Prediction explains how systems adjust; it does not explain why certain states count as correct or why agents are answerable to reasons.

- **Overextension.** The success of predictive models in perception and motor control tempts theorists to extend the same explanatory idiom to meaning, agency, and justification without supplying the bridging principles required for such extensions.

- **Practical risk.** Treating prediction as essence risks misguiding empirical priorities and ethical judgments, for example by assuming that behavioral mimicry suffices for moral or epistemic standing.

### Recommended replacement

Replace the metaphysical slogan with a **functional and contextual description** that preserves the empirical virtues of prediction while situating it within a broader explanatory ecology.

- **Treat the brain as a regulatory system embedded in organisms.** The brain participates in maintaining viability through prediction, homeostasis, and control, but it does so as part of an organism whose morphology, needs, and environment shape which predictions matter.

- **Emphasize participation in norm-governed practices.** Human cognition attains meaning and justificatory force through social practices, language, and institutions that make states publicly assessable.

- **Frame prediction as a tool, not the essence.** Prediction is a powerful instrument for regulation and adaptation; it is one indispensable component of cognition, not its defining essence.

## How the replacement guides research and reflection

- **Modeling practice.** Computational models should be explicit about the domain they address—mechanistic regulation, embodied affordances, or normative practices—and avoid unqualified claims that their success settles questions outside that domain.

- **Interdisciplinary programs.** Research agendas should investigate how predictive mechanisms are embedded in bodies and social practices, and how those embeddings enable normative statuses to arise.

- **Philosophical clarity.** Philosophers should resist metaphors that conflate explanatory categories and instead articulate the distinct standards appropriate to causal, functional, and normative explanations.

Abandoning the prediction-machine metaphor as a metaphysical identity does not diminish the scientific achievements of predictive processing. It places those achievements in their proper explanatory context and opens space for a richer, pluralist account of cognition—one that honors neural regulation, embodiment, and interpretive normativity without collapsing them into a single, misleading slogan.

## XI. Philosophical Contribution Explicit for Referees

This paper advances a targeted philosophical intervention: it preserves the empirical achievements of predictive processing (PP) while resisting the philosophical overreach that treats prediction as constitutive of meaning, understanding, and normative cognition. It does so by articulating a principled explanatory pluralism that distinguishes **neural regulation**, **embodied engagement**, and **interpretive normativity** as distinct explanatory domains and by showing how each domain contributes to a full account of cognition.

### What the paper delivers

- **Preserve empirical PP.** It accepts and preserves PP's explanatory successes in modeling neural regulation—prediction, inference, precision weighting, and active inference—treating these as indispensable mechanistic resources.

- **Delineate PP's limits.** It argues that statistical optimization and error minimization cannot, by themselves, generate **semantic aboutness**, **epistemic responsibility**, or **rational accountability**—the normative features that constitute understanding.

- **Clarify embodiment.** It reframes embodiment as a **necessary contextual constraint**: bodies and affordances determine which predictions matter and how predictive mechanisms are deployed, without endorsing mystifying anti-representationalism.

- **Offer explanatory pluralism.** It reframes the mind–body problem as a problem of **explanatory scope and integration**, not a single ontological puzzle, and provides a framework for coordinating mechanistic, embodied, and normative explanations.

## Anticipated objections and succinct responses

- **Objection: The pluralist account is ad hoc. Response:** The pluralism is principled. Each domain answers different kinds of explanatory questions—mechanistic, ecological, normative—that empirical practice already separates. The framework formalizes these separations and prescribes principled bridging moves rather than ad hoc patching.

- **Objection: Interpretive normativity is mysterious. Response:** Normativity is analyzed in tractable terms: public practices, standards of correctness, and justificatory relations. These are amenable to philosophical analysis and empirical investigation (developmental, sociological, and cognitive), not mystical invocation.

- **Objection: This weakens PP's unifying ambition. Response:** The account preserves PP's unifying role **within** the neural regulation domain while constraining its philosophical scope. Unification remains a methodological virtue where appropriate; it is not a license to conflate mechanistic description with normative constitution.

Why this matters to referees

- **For empirically oriented referees.** The paper respects and retains the technical and empirical contributions of PP, recommending concrete ways to situate predictive models within embodied and social contexts rather than discarding them.

- **For normatively oriented referees.** It defends the irreducibility of normative explanation without lapsing into dualism: normative phenomena are grounded in social practices and historical processes that require their own explanatory vocabulary, even while remaining physically instantiated.

- **For interdisciplinary standards.** The framework supplies clear methodological principles—constraint, translation, and autonomy—that guide legitimate cross-domain inference and interdisciplinary collaboration.

This paper proposes a defensible middle path: it honors the explanatory power of predictive models, integrates crucial insights from embodiment, and insists that interpretation and normativity require distinct philosophical treatment. The result is a scientifically informed, philosophically rigorous program that both constrains overreach and opens concrete avenues for empirical and conceptual research.

## References


Bickhard, M. H. (2009). *The interactivist model*. In J. K. O'Regan & A. Noë (Eds.), Sensorimotor Foundations of Perception (pp. 45–68). Oxford University Press.

Brandom, R. (1994). *Making it explicit: Reasoning, representing, and discursive commitment*. Harvard University Press.

Chalmers, D. J. (1996). *The conscious mind: In search of a fundamental theory*. Oxford University Press.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Trends in Cognitive Sciences, 17*(3), 1–9.

Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis, 58*(1), 7–19.

Dreyfus, H. L. (1992). *What computers still can't do: A critique of artificial reason* (2nd ed.). MIT Press.

Friston, K. (2013). Life as we know it. *Journal of the Royal Society Interface, 10*(86), 20130475.

Friston, K., Thornton, C., & Clark, A. (2017). Free-energy minimization and the dark-room problem. *Behavioral and Brain Sciences, 40*, e347.

Gallagher, S. (2005). *How the body shapes the mind*. Oxford University Press.

Hohwy, J. (2013). *The predictive mind*. Oxford University Press.

Kirchhoff, M., & Kiverstein, J. (2019). Predictive processing, enactivism and the skeptical turn. *Synthese, 196*(10), 1–23.

McDowell, J. (1994). *Mind and world*. Harvard University Press.

Putnam, H. (1975). The meaning of 'meaning'. In K. Gunderson (Ed.), *Language, mind, and knowledge* (Minnesota Studies in the Philosophy of Science, Vol. 7). University of Minnesota Press.

Rietveld, E., & Kiverstein, J. (2014). A rich landscape of affordances. *Ecological Psychology, 26*(4), 325–352.

Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences, 3*(3), 417–424.

Seth, A. K. (2014). The cybernetic Bayesian brain: From interoceptive inference to sensorimotor contingencies. *Open Biology, 4*(12), 140.

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition*. Harvard University Press.

Tomasello, M. (1999). *The cultural origins of human cognition*. Harvard University Press.

Winograd, T., & Flores, F. (1986). *Understanding computers and cognition: A new foundation for design*. Ablex.

Zahavi, D. (2005). *Subjectivity and selfhood: Investigating the first-person perspective*. MIT Press.