

Encoding All the Way Down

A conversation about consciousness, experience, and agency

The dismissal of biological encoding as the foundation of consciousness, the self, and perceived agency may have been premature. This piece doesn't claim to have the answer — it asks why the question was taken off the table, and whether the reasons hold up. It is written for curious lay readers rather than specialists. The arguments can be developed further against technically trained objections — on epiphenomenalism, illusionism, compatibilism, and the combination problem — but that's a different conversation than the one this piece is trying to have.

On the wrong end of Descartes

Something about the standard framing of consciousness has always bothered me. Descartes said "I think, therefore I am" — cognition as the proof of existence. But what if it's backwards? What if existence precedes cognition, and consciousness is simply what a particular kind of biological being feels like from the inside? I am, therefore I think.

On this picture, the hard problem of consciousness is reframed from a different base assumption. Rather than asking how physical process gives rise to experience — which presupposes they're two separate things requiring a bridge — the starting point is that experience may be what a particular kind of biological process is from the inside. Or at least, that's the working hypothesis this piece takes seriously.

I'm not pretending this dissolves every puzzle. I'm asking why we're so quick to assume that physical process and experience must be different kinds of thing in the first place.

The bird

Consider a bird that builds an intricate nest on its first attempt — never taught, never having seen it done. It selects specific materials, arranges and weaves them into a precise structure, corrects errors as it goes. We don't call the nest an illusion because it's encoded. It's a real nest, built by a real process, serving a real function.

We don't know the precise mechanism. We know the output. And we accept the encoding anyway because the output is observable.

Is the bird aware that it's building a nest? We don't know. We can't rule it out. If some form of first-person experience accompanies the behavior, then the same underlying biological process produced both the physical object and whatever awareness there is. From the bird's point of view — if it has one — it would be a first-person experience.

Counter: We don't know that the bird has a point of view at all. You're inferring inner experience from observable behavior.

That's exactly the move I'm questioning. We infer the encoding from the behavior without requiring awareness to explain it. But we also can't rule awareness out. The argument isn't that the bird is definitely conscious — it's that the style of skepticism often used to block biological explanations of consciousness would, if applied consistently, also block our everyday acceptance of biological encoding in cases like nest-building.

The bird analogy isn't claiming that consciousness is "just like" a nest. Its role is epistemic: we already tolerate unknown encoding mechanisms for complex, structured biological outputs. The argument is that we should extend that tolerance to consciousness rather than treating "we don't know the mechanism" as a knockdown objection in one domain but not the other.

Three nested problems

There are three problems that lay people sometimes collapse into one.:

Why does processing feel like anything at all? That's the hard problem proper.

How does feeling produce a unified self? Call that the self problem.

How does a self produce genuine agency — the feeling of being the author of one's actions? That's the free will problem.

The chain, at least phenomenologically, runs: consciousness (awareness) → experience (the felt quality of being) → agency (the feeling of being in charge).

What I'm asking is why we're not seriously considering that a single kind of biological encoding produces all three. Not through three separate mechanisms needing three separate explanations, but as different aspects of the same organized complexity. We still don't know how. Not knowing the mechanism isn't the same as the mechanism not existing.

This doesn't solve the hard problem in Chalmers's sense. It insists that any solution we eventually arrive at will have to live inside biology rather than outside it.

The determinism problem

Determinism is usually invoked to dismiss agency — the causal chain predates the decision, the self is a product of prior causes, therefore the feeling of choosing is an illusion.

But there's an internal consistency question that doesn't get examined often enough. Determinism depends on the causal chain being complete and physical all the way down. If you exempt consciousness from physical explanation — place it outside physics as something categorically different that also does causal work — you've broken the chain. A non-physical consciousness interacting causally with a physical brain isn't pure determinism; it's a form of interactionist dualism. And once you admit that, you've at least complicated the argument you were using against free will.

There are other options — epiphenomenalism (consciousness is non-physical but causally inert) and various non-reductive physicalisms. I don't find them satisfying; they make experience explanatorily idle despite its centrality to our lives. But the narrower point here is that if consciousness does play a causal role, treating it as non-physical undermines the neat "physics explains everything" move often used to dismiss agency. Consciousness is either wholly within the physical story — in which case the biological encoding question is legitimate — or it isn't, in which case the closed-chain version of determinism used to dismiss agency is in trouble.

Counter: Sam Harris invokes causality to dismiss the evolutionary argument. If thoughts arise from prior causes you didn't choose, free will is undermined regardless of whether consciousness is physical.

Causality isn't an argument against the evolutionary thesis. Causality is the mechanism that produces biological encoding — it can't simultaneously be the objection to it. And the causal chain doesn't imprison the self if the self is an expression of that chain. On this view, the causes running through you aren't external constraints imposed on some separate "you." They are you.

Some will recognize this as a compatibilist-adjacent move — and there's truth to that. But it's not quite standard compatibilism, which accepts the determinist picture and redefines free will to survive inside it. The claim here is different: the framing that sets up the problem — a self imprisoned by a causal chain external to it — is what's being rejected. There is no imprisonment because there is no prisoner separate from the process.

The evolutionary argument

Organisms that feel genuine agency — that experience themselves as authors of their actions — may outcompete organisms that don't. An organism that fully experiences itself as a deterministic automaton following a script might not try as hard, plan further ahead, or persist through difficulty as effectively. The felt sense that choices matter is motivationally load-bearing.

If behavior and experience aren't cleanly separable — if what the organism does and what the organism feels like doing it are two descriptions of one underlying process — then selecting for the behavior and selecting for the experience aren't two different evolutionary acts.

A critic can reasonably reply that evolution only "sees" causal structure, not phenomenal feel. Granting that: the question then becomes why we shouldn't take seriously the possibility that phenomenal feel just is one aspect of that evolved causal structure. The evolutionary argument doesn't prove the identity claim — it asks why the identity claim isn't the default.

This isn't saying evolution selected for a useful fiction. The claim is stronger: the experience isn't a fiction layered on top of behavior. It's what that behavior is from the inside, if the working hypothesis is right. The nest isn't a useful fiction about nest-building. It is the nest.

Human exceptionalism

As far as I can find, the biological encoding hypothesis was largely set aside for reasons that have more to do with human self-regard than with evidence. The Western tradition — philosophical and religious — placed consciousness in a category uniquely available to humans. Descartes declared animals automata. The soul was uniquely human.

Darwin should at least have shaken this. Evolution doesn't produce metaphysical jumps; it produces continuums, even if there can be sharp transitions in functional space. If consciousness exists in humans and we share common ancestry with everything else, the question isn't whether other organisms have experience but where on the continuum they fall.

The evidence keeps moving the boundary. Orcas demonstrate self-recognition, individual naming, grief-like behaviors, culture, and learned transmission of experience across generations. Their neural architecture is, in some respects, more elaborately developed than ours in areas associated with emotional processing and social cognition — a point worth remembering when we reach for human specialness as the reason to set consciousness aside.

We're not special. We're specialized — in abstract reasoning and symbolic thought. That specialization doesn't place us outside the biological processes that produced everything else. Reasoning is our version of nest-building.

On metaphysics

When the mechanism is unknown, the temptation is to reach outside the physical framework. But invoking the non-physical — soul, pure awareness, non-natural properties — doesn't explain consciousness; it relocates the mystery to a domain where no tools exist to investigate it. We have better odds eventually solving a brutally hard physics and biology problem than making progress on forces we can't define, detect, or constrain.

Metaphysics in the strong, dualist sense and a strict deterministic physics also aren't comfortable companions. Admitting one tends to destabilize the other. If something outside the physical causal chain influences consciousness, the chain is no longer complete — and determinism, in the strong sense often appealed to in free will debates, depends on that completeness.

What follows

If some form of biological encoding produces consciousness, experience, and agency, the search for a separate purpose to existence becomes less pressing. On this view, it's evolution coding for survival. We are what successful encoding looks like after sufficient complexity and time.

If that's even roughly right, there's an irony: we may be coded so aware of ourselves that the awareness becomes the instrument of our undoing. An encoding that produced organisms capable of this level of self-reflection, and this level of destructive

capacity, looks exactly like what natural selection operating without foresight would produce.

Objections examined

The following objections were raised in conversation with an AI during the development of this argument. They are included because they represent the kinds of challenges a philosophically curious reader might bring.

The combination problem: individual neurons don't experience anything. How does combining non-experiencing parts produce an experiencing whole?

This applies a double standard. The process producing a physical nest is equally unknown at the mechanistic level, and we don't withhold acceptance of encoding as the explanation. The objection assumes the separability of physical process and experience — which is precisely the assumption being challenged. If encoding produces behavior and experience as one process described from two vantage points, there are no non-experiencing parts being combined. There is one process. The question of how it produces unified experience is the same question as how it produces unified behavior. Currently unanswered in both cases. That's not an argument against encoding in either.

The hard problem residue: even granting encoding, why does it produce felt quality rather than just processing in the dark?

That question remains open for every framework. Determinism doesn't answer it. Dualism doesn't answer it — it relocates it. The why of subjective experience is everybody's open problem. Raising it specifically against biological encoding is special pleading unless the same standard is applied to all positions equally. The claim here is about the how. The how is encoding. The why stays open — which is the honest position for everyone in this debate.

Encoding errors: if consciousness is encoded, why does encoding produce cognitive biases and mental illness?

Encoding is an imperfect biological process. Developmental anomalies don't argue against encoding as a mechanism — they're features of biological systems operating under real-world constraints. Cognitive biases and mental illness are expressions of encoding operating imperfectly or in mismatched environments. The fact that consciousness fails in predictable, heritable, sometimes correctable patterns is consistent with an underlying encoded mechanism, not evidence against it.

Selected-for doesn't mean real. Evolution selects for useful fictions.

This assumes what it needs to prove: that experience is a representation layered on top of reality rather than part of reality itself. The bird nest isn't a useful fiction about nest-building. It is the nest. The felt experience of agency isn't a useful

misrepresentation of some other underlying process. It is what that process is from the inside, if the working hypothesis holds. The illusion/reality distinction the objection depends on requires the separation between experience and mechanism that the framework rejects.

Behavior and experience are separable. Selection operates on behavior, not internal experience.

This requires arguing for the separability rather than assuming it. If experience and behavior are one process described from two vantage points, selecting for one and selecting for the other aren't distinguishable acts. The objection needs to establish the separation before it can use it.