The Living Experience Hypothesis

Mats Lewan, January 11, 2026
Updated January 14, 2026, with an addendum on the experience of time.


Planet Earth. Source: Wikimedia Commons

For as long as humans have been able to reflect on their own existence, we have been intrigued by a simple and bewildering fact – we find ourselves inside experience. We see, hear, feel, remember, and think – and all of this unfolds from a single point of view that seems undeniably *ours*. The world appears "out there," our body somewhere in between, and our mind mysteriously on the inside. This everyday experience is so immediate and convincing that it is almost impossible to imagine things being otherwise.

From this basic intuition – that there is a self observing the world – entire systems of thought have grown. Philosophy, religion, literature, and myth have returned to it again and again, offering stories of souls and bodies, inner and outer worlds, observers and the observed. Yet the deeper we look, the stranger the situation becomes. Nowhere in the physical description of the world do we find colors, sounds, emotions, or the feeling of being someone at all. Physics describes particles and fields, biology describes cells and chemistry, neuroscience describes neurons and signals. And still, experience itself remains.

Over the last centuries, remarkable progress has been made in understanding the universe and life within it. Physics and cosmology have revealed a world governed by deep regularities. Chemistry has shown how complexity can arise from simple interactions. Biology has explained how living systems maintain themselves far from equilibrium. Psychology and cognitive science have uncovered many of the mechanisms underlying perception, action, and behavior. And contemplative traditions, most notably Buddhism, have explored experience from the inside with extraordinary precision.

Yet at the boundaries of each of these domains, an explanatory gap remains. Scientists describe structure and function, but not what it is like to exist. Philosophers analyze concepts of mind and matter, but struggle to anchor experience in the natural world. Contemplative traditions dissolve the idea of a self, but often leave open the question of why experience exists at all. Despite centuries of inquiry, we still lack a unified account that explains both the universe as it is and how it is lived from the inside.

The **Living Experience Hypothesis (LEH)** is an attempt to bridge this gap.

Rather than introducing new metaphysical entities or invoking mysterious forces, LEH connects insights that already exist across physics, biology, cognitive science, and philosophy. It asks a simple but radical question: *what kind of world must exist for life and experience to be possible at all?* In doing so, it deliberately avoids traditional starting assumptions – such as a moment of creation, the beginning of time, a privileged observer, or a fundamental divide between mind and matter. Instead, it begins from a single minimal constraint – that not everything is possible.

To make this synthesis coherent, LEH introduces two specific claims about **qualia** — the lived, subjective experience that every human is directly familiar with, and which likely exists, to varying degrees, in many other animals:

- Qualia are the inner, lived aspect of perturbations in a living, autonomous process.
- Proto-qualia most likely arise at the boundary of self-organising, adaptive life – concretely at the cell membrane – and are later integrated into unified first-person experience through bodily and neural dynamics.

As will become clear, these claims, together with the broader theoretical context in which they are embedded, lead to substantive conclusions about the nature of intelligence, and about the deep interdependence between life, experience, intelligence, and meaning.

LEH can be described as a **dual-aspect, non-reductive, naturalistic monism**, compatible with enactivism and embodied cognition. At the same time, it rejects substance dualism, reductive physicalism, panpsychism, and representational forms of cognitivism.

Before going into the details, it is important to emphasize that this short essay is intended as a high-level, first articulation of the Living Experience Hypothesis, written for an interested but non-specialist audience. It is not a scientific paper. A more formal and detailed treatment, developed in collaboration with one or more co-authors, is planned for the future.

What is experience

Picture yourself sitting in a garden with some friends, having coffee and a piece of strawberry cake, while children are running around playing nearby. Without deep reflection or meditative training, most people would naturally describe this situation by saying that they are *observing the world* and interacting with it. And that what they observe and interact with *is* what the world is like.

This is a completely reasonable way of thinking. Everything we learn from infancy onward supports it. Our senses appear to give us direct access to reality, and this view works well – it helps us navigate the world, coordinate with others, and survive. Very little in everyday experience suggests that things could be otherwise.

Except science. And, for some, meditative practice.

Over the last few centuries, scientific discovery has given us an increasingly detailed understanding of the laws governing the universe and the structures that arise from them. From a physical perspective, this knowledge is extraordinarily successful. It allows us to predict events, build technology, and explain phenomena across vast scales of space and time.

Yet, when we look closely, something strange appears. Much of what we experience does not exist in the physical world as described by physics.

From a physical point of view, there is no such thing as the taste of coffee, the redness of strawberries, the sound of your friends' voices or the children's laughter, the scent of flowers in the garden, the feeling of the wind on your skin, or the warmth of the sun on your face. Physics describes electromagnetic waves of certain wavelengths, chemical compounds interacting with receptors, pressure waves in air, and energy transfer. But nowhere in these descriptions do we find taste, color, sound, smell, or touch *as they are experienced*.

And yet, these experiences undeniably exist. Each of us has direct access to them. They are not speculative or theoretical – they are the most immediate facts of our lives. In philosophy and cognitive science, these first-person experiences are called **qualia**.

What this means is that the world we experience is a rich composition of qualia originating from our senses – sight, hearing, taste, smell, and touch – while the physical world described by science contains no trace of them. The "real world," as physics describes it, looks nothing like the world we actually live in.

The most straightforward conclusion is that our multisensory experience of the world is *constructed by the mind*. Fortunately, this construction is broadly similar across individuals, allowing us to communicate, cooperate, and share a common reality – even though each person's experience remains private.

We will return later to deeper questions about whether there is even an "observer," and what we mean by mind, thought, and consciousness. For now, one question takes priority: **if the world we experience is constructed in the mind, what is the world itself actually like?**

Physics offers some answers. According to modern physics, the world is made of matter and energy – which, as described by Einstein's theory of relativity, are ultimately the same thing. Matter is composed of atoms and elementary particles. But when we probe deeply enough, familiar intuitions break down.

At atomic scales, objects no longer behave like solid things moving through space. In quantum mechanics, particles are described by probability distributions rather than definite positions and velocities. Even the notion that particles are "made of stuff" becomes uncertain – they can just as well be described as excitations of underlying fields.

And this raises deeper questions. Where does matter and energy come from in the first place? Was it created? Did it arise from nothing? When did time begin – and why? At these foundations, physics reaches limits where it can no longer provide satisfying answers.

So physics cannot explain where our experience of the world comes from – and when pushed far enough, it struggles to explain the ultimate nature of the world itself. For anyone genuinely curious about existence, this is unsatisfactory.

The **Living Experience Hypothesis (LEH)** is an attempt to address these gaps in a consistent way.

The nature of the universe

To describe the nature of the universe, LEH builds on the hypothesis proposed by **Stephen Wolfram**, according to which reality is fundamentally computational. Calling the universe "computational" does *not* mean that it runs inside a computer, or that the universe literally *is* a computer.

Instead, the idea is that reality consists of an extremely fine-grained logical structure – a rule-based system whose evolution unfolds through discrete steps. These steps are so fundamental that the smallest physical structures we can observe would correspond to unimaginably large numbers of underlying rule applications – numbers far beyond anything encountered in ordinary physics.

Wolfram does not claim to study this structure directly. Rather, he examines the properties of such computational systems in general, and compares them with the observed properties of our universe. The strength of his hypothesis – and the reason it provides a suitable foundation for LEH – lies in three features:

1. **It avoids creation stories**, any external act of initiation, or any privileged observer
2. **It relies on a single minimal assumption**: that not everything is possible
3. **It explains physical laws** we observe not as arbitrary accidents, but as natural consequences of computational structure

At this point, a natural question arises: how could an abstract, rule-based structure give rise to the astonishing richness of the universe we observe – from galaxies to living organisms, and ultimately to experience itself?

This may seem puzzling, even implausible. But recall that we have already concluded that the world we experience is not identical to the world as described by physics. Wolfram himself describes observers like us as *sampling* the underlying computational structure, producing a particular interpretation of it. He stops short, however, of addressing what experience itself is.

The Living Experience Hypothesis takes up that challenge. But first, we need a clearer picture of Wolfram's computational universe.

Stephen Wolfram's computational universe

I will introduce Wolfram's hypothesis briefly here, but the best way to gain a deeper understanding is to read [his first major blog post on the topic] from April 2020.

The hypothesis has its intellectual origin in **Stephen Wolfram's** decades-long study of cellular automata, which revealed how the repeated application of very simple logical rules can generate surprisingly rich and complex structures. His current framework generalizes this insight far beyond cellular automata, using much more abstract rule-based systems to describe the foundations of physics.

At the most basic level, logical rules are applied to states. A state could be something as simple as a sequence of symbols, such as *BBBAAA*. A logical rule might be *BA* → *AB*, meaning that wherever the sequence *BA* appears, it is replaced by *AB*. Applying this rule to *BBBAAA* yields *BBABAA*. Applying it again, you may obtain either *BBAABA* or *BABBAA*, depending on which instance of *BA* is replaced first. In other words, the structure branches.

States can be more complex than simple strings. They can consist of collections of symbols whose relationships matter – for example, pairs or triplets of elements. Such states can be interpreted as graphs, and each application of a rule can be seen as creating a connection between states. In this way, repeated rule application generates a branching structure, sometimes called a multiway graph.

It is important to emphasize that these symbols and states do not represent anything. They have no intrinsic meaning. They are purely logical configurations. The key takeaway is not representation, but structure. As mentioned earlier, the repeated application of simple rules often leads to impressively complex behavior.

Wolfram goes on to explore how the branches of this multiway structure diverge and sometimes merge again, meaning that the same state can be reached through different sequences of rule applications. He also studies causality within this structure, identifying which states depend on which earlier states. This analysis gives rise to what he calls a *causal graph*.

What he discovers is striking. Under certain conditions, basic features of the physical world we observe – such as aspects of relativity, including a maximum speed of influence (the speed of light), and even elements of quantum mechanics – emerge naturally from these logical structures. Wolfram also shows how notions such as spatial dimensions, time, energy, and momentum can be interpreted within this framework, related through the same mathematical formulas that describe the world we experience.

However, despite this emergent order, most of the structure behaves in a chaotic way. By “chaotic,” Wolfram does not mean random, but rather unpredictable in practice. For most rule systems, there is no shortcut that allows you to know what the structure will look like after many steps without actually applying the rules step by step. Wolfram calls this property **computational irreducibility**.

At the same time, he observes that there are **pockets of computational reducibility**. In these regions, stable patterns emerge that allow for mathematical shortcuts – formulas that make it possible to predict future behavior without simulating every intermediate step. This is where predictability enters the picture.

Predictability is central to how we navigate the world. As infants, we learn by observing, forming expectations, testing them, and adjusting our behavior. If I throw an object into the air, it falls. If I throw food on the floor, my parents may react negatively. Our minds constantly search for regularities, for relations between cause and consequence, in order to anticipate what may happen and adapt accordingly.

This is also what science does. By identifying regularities and expressing them as natural laws, science allows us to predict the future behavior of systems – often with remarkable precision.

Early in the development of his hypothesis, Wolfram imagined that there might exist a specific rule whose repeated application gives rise to *our* universe, with its three dimensions of space and particular physical laws. He later generalized this view, arguing instead that reality consists of *all possible logical rules applied over and over again*. Wolfram refers to this underlying space of all possible rule-based structures as the *ruliad*. In this essay, I will instead use the more general term *logical structure*, since the focus here is not on listing or exploring all such possibilities, but on how the universe we inhabit – and experience – can be situated within such a framework.

What we experience as "our universe" would then be just one pocket of computational reducibility within this vastly larger structure – a pocket whose patterns happen to align with the observations available to beings like us.

As a side note, what we hypothesize to be the beginning of our universe – the Big Bang – can, from this perspective, be understood as the point within the total logical structure where *our* pocket of computational reducibility begins to emerge. This emergence may well correspond to a rapid and "violent" transition in the underlying structure.

In this view, although our observable universe is vast, it would still represent only a tiny region of a far more immense logical structure – most of which would be incomprehensible to a mind like ours, and incompatible with the conditions required for life.

The only assumption required for such a universe to exist is that **not everything is possible**. If not everything is possible, then structure exists. And if structure exists, logical rules can exist. *That is the sole assumption underlying Wolfram's hypothesis – and also the Living Experience Hypothesis.*

Wolfram leaves open the question of whether the universe actually evolves by applying rules step by step, or whether the full structure simply exists as a logical possibility. LEH adopts the latter view, for reasons that will become clear later.

Let me add two side notes here, concerning time and causality.

Time appears essential to how our minds function. It is nearly impossible for us to imagine a world without time flowing. You may picture a frozen moment, but in doing so you implicitly rely on another flow of time – the one in which you observe the frozen scene. Our minds seem fundamentally dependent on temporality.

Yet, in the underlying logical structure, there is no flowing time. Time can be interpreted as a direction within the structure – particularly within pockets of computational reducibility – but it does not "pass." This suggests that our experience of time, like our experience of color or sound, is constructed by the mind. Many philosophers and cognitive scientists have arrived at similar conclusions, and Einstein himself spoke of a *"block universe"* in which all moments of time simply exist.

This has an unsettling implication: if the logical structure is fixed and fully determined, then everything that happens is, in a sense, predestined. However, as Wolfram notes, this predestination has no practical significance for beings like us. And the reason becomes clear when we turn to causality.

Philosophy has long struggled with causality. We all experience it as obvious and real – striking a ball makes it fly away. But the deeper physics looks, the harder it becomes to identify a clear causal mechanism. At many levels, what we find are reliable correlations rather than identifiable causes. In quantum mechanics, even these correlations become probabilistic rather than certain.

In Wolfram's framework, causality exists at the most fundamental level of the structure: rule applications depend on earlier rule applications. This is causality in its purest form. But it exists at a level to which minds like ours have no access.

Our minds depend on the patterns that emerge in pockets of computational reducibility – on regularities that make the world predictable over time. Observing the structure in a coarse-grained way, extracting large-scale patterns and treating them as laws, we encounter what *appears* to be causality.

From a strict logical perspective, there is no fundamental causality at this level – only sufficiently reliable regularity. The true causality lies deeper, beyond our reach.

This brings us back to predestination. If the underlying structure is fully determined, then everything is indeed fixed. But this does not license resignation. The reason is simple: you cannot both rely on predestination and continue to exist as the kind of being you are.

Your existence depends on inhabiting a pocket of computational reducibility – one in which prediction, uncertainty, choice, and the experience of time are meaningful. Accessing the level at which predestination resides would undermine the very conditions that make your existence possible.

In this sense, you must choose. Either you exist as a living, experiencing being, or you appeal to predestination. You cannot have both.

On the contrary – as we will see later, the ability to find meaning and make choices under uncertainty is the very essence of life.

Before turning to how this framework relates to the universe we observe, it is worth noting that the studies and discoveries on which Wolfram builds his hypothesis would not have been possible without computers. They rely on the repeated application of simple rules at a scale and speed far beyond what the human mind can track unaided.

In this sense, the very way of thinking about a potential underlying structure of the universe proposed here could not have arisen much earlier in human history. It is a perspective that depends on technological tools capable of extending human reasoning into domains of complexity we cannot directly grasp.

Relating to the universe we observe

Now, if we momentarily set aside the questions of what life is, and how it can experience the world, and simply accept that this is possible for the time being, we can briefly examine how structures within a pocket of computational reducibility may come to resemble the universe we observe – and how our everyday experience of the world may depend on such structures.

First of all, as Wolfram points out, for his hypothesis to explain the universe as a whole, we – including our minds – must be part of the logical structure itself. We are not external observers. I will return to this point in more detail later.

The general idea Wolfram proposes for the emergence of familiar physical structures is that, within pockets of computational reducibility, large regions of the structure become stable and repeatedly reappear. Such stable “lumps” of structure can have properties that correspond to the elementary particles we observe. From there, it becomes conceivable how increasingly complex structures – atoms, molecules, and eventually the macroscopic world – could arise, provided that the overall structure also gives rise to the natural laws we observe.

Thinking of elementary particles as stable patterns within an underlying logical structure gives an immediate sense of the enormous number of rule-application steps that must underlie even the smallest aspects of our observable universe. Wolfram also estimates that an overwhelming portion of the logical structure does not correspond directly to observable objects or phenomena at all, but instead serves to maintain the deeper framework in which these structures exist. The scale separation between the fundamental rule applications and the world we experience is therefore truly staggering.

Quantum mechanics provides a particularly interesting case. Compared to our everyday experience, quantum phenomena appear unintuitive – outcomes are not certain, but probabilistic, and multiple possibilities coexist until one is observed.

From the perspective of the underlying logical structure, however, Wolfram offers a surprisingly intuitive interpretation. As we have already seen, the structure branches as rules are applied in different possible orders, and these branches can later merge again. This means that events may follow different paths with different intermediate outcomes, yet still lead to the same resulting state. Because we are observers embedded within the structure and cannot follow multiple paths at once, we experience only one outcome of such an event – even though other paths were equally possible.

Wolfram refers to this property as **causal invariance**, and it turns out to be essential for quantum mechanics to emerge as a direct consequence of the logical structure. If a pocket of computational reducibility exhibits causal invariance, then quantum-like behavior follows naturally. From this perspective, the apparent strangeness of quantum mechanics becomes far less mysterious.

We can also think of this in terms of the boundaries of our pocket of computational reducibility. The patterns and regularities we observe when viewing the structure in a coarse-grained way at our familiar scales may not hold outside this pocket. As we probe very small scales – or possibly approach the limits of the observable universe – these regularities may begin to loosen or break down.

Quantum mechanics may be precisely such a boundary phenomenon. And it is possible that beyond quantum mechanics we encounter nothing that makes sense to us at all, as some interpretations suggest – not because there is no deeper structure, but because that structure lies outside the range of patterns that a mind like ours can recognize or interpret.

Ordinary life events

Now, let us consider some ordinary events in everyday life – such as climbing a steep hill or riding a bicycle – and ask how there could be any correspondence between an underlying logical structure and our lived experience of the world.

When you climb a hill, you feel the need to activate your muscles to resist the pull of gravity and move upward toward the top. You feel your muscles growing tired from the effort, and you feel your body becoming warmer as energy is expended. What could experiences like these possibly have to do with a logical structure?

Let us begin by assuming that this structure gives rise to relationships that may be interpreted as the natural laws we observe in physics, chemistry, and biology. As discussed earlier, we also assume that our minds are capable of experiencing aspects of an underlying physical world.

Within this structure, the laws describe quantities such as energy, momentum, and forces – for example, the force required to overcome gravity acting on the mass of your body. They also describe how the necessary energy is produced through chemical reactions in your muscles, how that energy is converted into movement, and how heat is generated as a by-product of these processes.

Biology then describes how these processes give rise to signals within the body – through the nervous system and through chemical signalling, such as neurotransmitters, hormones, and metabolic byproducts. In much the same way that the mind constructs experiences of light and sound, it constructs experiences of effort, fatigue, and warmth based on these signals. At the same time, your mind constructs the experience of moving upward along the hill through sensory input from vision, touch, balance, and proprioception.

Taken together, the laws governing a pocket of computational reducibility constrain how we, as parts of the logical structure, can interact with the underlying world. Our minds – which are themselves also part of that structure – construct a unified, multisensory experience of these interactions unfolding over time.

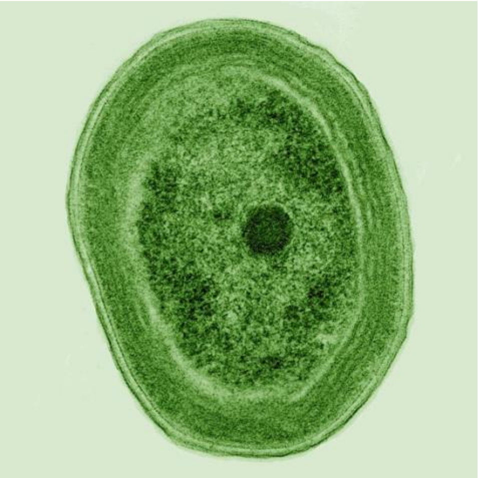
The same reasoning applies to riding a bicycle. In this case, a different set of physical constraints governs how we interact with the world – the dynamics of balance, motion, and momentum involved in staying upright on two wheels while moving forward. To succeed, we must continuously engage our senses and adjust our actions based on ongoing feedback about our movement and position in space. Yet the underlying logic is the same – lawful interactions at one level give rise to lived experience at another.

What remains unexplained is how the mind constructs experience – how qualia arise from physical processes, and ultimately from information. We will return to this question. But first, we need to examine what we mean by life, and how life itself may arise within this logical structure.

Life, autopoiesis, and entropy

At the most basic level, life might be defined as a process that maintains itself. But this description is too broad. Fire, sustained by chemical reactions, maintains itself. Stars, sustained by nuclear fusion, do the same. We need a more stringent definition.

The Living Experience Hypothesis adopts the definition of life proposed in 1972 by biologists Humberto Maturana and Francisco Varela, who introduced the concept of *autopoiesis*, from the Greek *auto* (self) and *poiesis* (creating). An autopoietic system is one that continuously produces and regenerates the processes that constitute it, while maintaining its own boundary with the environment.



Prochlorococcus marinus – a single-celled bacterium exemplifying the simplest living system with an inside, an outside, and a self-maintaining boundary. Source: Wikimedia Commons.

This definition introduces several crucial features – identity, autonomy, and regulation. An autopoietic system is distinct from its environment, yet continuously interacts with it, adapting to changing conditions by regulating its internal processes.

At this point, it is useful to introduce the concept of entropy. Entropy is often described as a measure of disorder – low entropy corresponds to high order, while high entropy corresponds to disorder, noise, or randomness.

Entropy is closely connected to the second law of thermodynamics, which states that in any natural thermodynamic process, the total entropy or disorder of interacting systems does not decrease. A familiar corollary is that heat does not spontaneously flow from a colder body to a warmer one.

We know this from everyday experience. A warm object placed in a cooler environment gradually loses heat until both reach the same temperature. More generally, the second law tells us that, over time, order tends to dissolve unless energy is expended to maintain it. Entropy may decrease locally, but only at the cost of increasing entropy elsewhere.

Fire provides a simple example – it releases heat and increases entropy overall, while destroying the organized chemical structure stored in its fuel.

The second law also implies an arrow of time. Most fundamental physical laws are time-symmetric – they do not distinguish past from future. The second law is what gives time a direction at the macroscopic level.

It is worth noting that the laws of thermodynamics, including the second law, are not proven in the same way as mathematical theorems. They are grounded in overwhelming empirical evidence and the absence of counterexamples, rather than formal derivation from deeper principles.

Wolfram adds an important perspective here. At the level of the underlying logical structure, there is no fundamental arrow of time. Rule applications do not erase information, and in principle one could reconstruct earlier states from later ones. From this perspective, the structure is timeless.

However, for observers like us, who sample the structure in a coarse-grained way within a pocket of computational reducibility, information is effectively lost at each step. This makes it impossible to reconstruct the past in detail, giving rise to an experienced arrow of time. According to Wolfram, this is why we observe the second law: entropy appears to increase globally, even though it may decrease locally.

From the definition of autopoiesis, it follows that life must adapt to its environment. And since life can only exist within a pocket of computational reducibility – where apparent causality, regularity, and predictability are available – it must also obey the second law of thermodynamics.

This means that since life maintains internal order, it can only do so by exporting entropy to its surroundings. This necessarily implies interaction with the environment. An autopoietic system cannot be isolated – isolation would mean thermodynamic equilibrium, and equilibrium means death.

However, autopoiesis also implies a boundary separating the system from its environment. In the simplest life forms – single-celled organisms – this boundary is the semi-permeable cell membrane. The membrane is the interface through which the cell exchanges matter and energy with its environment, via ion channels, transport proteins, receptors, and other molecular mechanisms. In other words, the cell membrane is a boundary, but a boundary that doesn't imply isolation.

Here it is useful to add a side note on entropy, which bears on life from two complementary perspectives.

From an energetic and material point of view, low entropy – high order – is valuable because it can be used to perform work. Energy gradients, structured matter, and organized resources are what allow living systems to maintain themselves and act on their environment.

From an informational perspective, however, the situation is different. Perfect order carries little information, because it is fully predictable. Complete disorder or chaos carries little usable information as well, because it lacks stable structure. Rich information emerges in the middle ground between order and chaos, where structured but flexible patterns can arise.

This balance is precisely where life operates. Living systems exist far from equilibrium, maintaining enough order to persist while remaining flexible enough to adapt. It is also arguably why many aspects of human experience are most compelling when they inhabit this same middle ground – whether in art, beauty, relationships, architecture, design, work, or play. The beauty of nature itself reflects this balance – never perfectly ordered, never entirely chaotic, but structured, dynamic, and alive.

At this point, it is natural to ask how the earliest life forms could arise in the first place.

Assuming that atoms and molecules exist, and that they interact according to stable physical and chemical laws within a pocket of computational reducibility, the emergence of life no longer appears mysterious in principle. Chemistry already provides many examples of self-organizing, far-from-equilibrium processes that maintain structure by dissipating energy.

Under suitable conditions, certain molecular systems can begin to store information, replicate with variation, and catalyse their own persistence. Molecules such as RNA – capable of both carrying information and participating in chemical reactions – offer a plausible route toward early biological organization. Once such systems become enclosed by a boundary, such as a lipid membrane, they acquire a crucial new property – an inside distinct from an outside.

At this point, a minimal form of life can arise – an autonomous, self-maintaining process that regulates its internal state, exchanges matter and energy with its environment, and persists over time. This transition requires no violation of physical laws. It is a natural consequence of chemistry operating far from equilibrium within a stable and predictable physical regime.

Life, qualia, meaning, and intelligence

Now things become interesting.

The existence of a cell membrane implies the emergence of normativity. The cell must distinguish between different kinds of perturbations at its boundary and respond to them in ways that preserve its internal organization. Changes in salinity, acidity, temperature, or nutrient concentration matter differently for the viability of the cell.

It is difficult to imagine such regulation being achieved through internal computation in the sense of explicit calculation – evaluating perturbations, comparing them to stored values, and choosing responses accordingly. A far more efficient solution is that these differences are lived by the system itself.

In fact, it is plausible that such lived sensitivity is not merely advantageous, but a necessary condition for autopoiesis. From this perspective, experience is not an optional addition to life, but an intrinsic aspect of it.

This leads to the two central claims introduced by the Living Experience Hypothesis:

- Experience is the inner, lived aspect of normativity – how perturbations relevant to viability matter from the inside of a living system.
- Experience emerges at the boundary of life itself – concretely, at the cell membrane.

This elementary form of experience can be described as proto-qualia – the most basic origin of qualia.

Proto-qualia do not introduce a new external property or substance. They are the intrinsic, first-person aspect of information as it matters for a living system – an aspect that cannot be captured or reduced by third-person description alone.

One may think of this as a dual aspect of information. Just as a waveform can be described either in the time domain or the frequency domain, or as quantum phenomena can be described as particles or waves, normatively relevant information in living systems has both an external, structural aspect and an internal, lived aspect.

This inner aspect arises only when information is encountered by a living system and bears on the viability of its internal processes. Crystals, for example, may be influenced by external perturbations, but they lack autopoiesis and normativity; therefore, they lack experience.

For this reason, LEH rejects panpsychism – the view that experience, or experience-like properties, are present in all matter. It also rejects the idea that experience only appears with the evolution of neurons and nervous systems. While nervous systems dramatically expand and integrate experience, placing the origin of experience at that point in evolution would be arbitrary in relation to the emergence of life itself.

Instead, experience is more plausibly located at the beginning of life, with nervous systems evolving as mechanisms for integrating, synchronizing, and amplifying experience across many cells.

This suggests that life and experience are deeply interdependent. Two additional concepts follow naturally – intelligence and meaning.

Even the simplest life forms can be regarded as minimally intelligent, in the sense that they continuously solve one of the most difficult problems known – staying alive under changing conditions.

Meaning, at its most basic level, arises from normativity. What matters to a living system – what contributes to or threatens its continued existence – constitutes the simplest form of meaning.

From this perspective, life, experience, intelligence, and meaning arise together in their most elementary forms:

- **Life** is autonomous self-maintenance under thermodynamic constraint.
- **Experience** is the inner, lived aspect of life.
- **Meaning** is the normative significance of what happens to the system.
- **Intelligence** is the functional capacity to act on meaning to sustain life.

These aspects are distinct, yet inseparable.

The proposal that experience originates at the cell membrane is not arbitrary. It is supported by several considerations:

- experience requires an inside/outside distinction
- perturbations must be transduced across a boundary
- the system must care about what crosses that boundary
- membranes are where chemistry becomes normatively organized

This does not imply that membranes “feel” in any human sense, nor that neurons are irrelevant. It is a claim about where experience must begin in its most elementary form.

There is also suggestive evidence from pharmacology – many anaesthetics disrupt membrane dynamics and ion channel function, interrupting the integration of experience. While this does not prove the hypothesis, it is consistent with it.

If experience originates at the cellular boundary, how do we account for higher-level qualia such as colour, sound, pain, temperature, smell, and taste?

Rather than undermining the hypothesis, well-established biological facts support it. Biological differentiation is largely realized through specialized membrane dynamics. Different cell types exhibit highly specific sensitivities to physical and chemical perturbations, shaped by evolution to matter for regulation.

Qualitative differences in experience arise not because cells represent external properties, but because different membrane dynamics are selectively perturbed in different ways that matter for the organism. These lived perturbations are then integrated into coherent experiential modalities by nervous systems, which we will discuss later.

Normativity does not mean that every perturbation is good or bad for individual cells. It means that perturbations matter relative to the self-maintenance of the organism as a whole. In photoreceptors, photons are normatively relevant not in isolation, but because their membrane dynamics contribute to organism-level regulation.

We may stop here for a moment to discuss a natural objection to the Living Experience Hypothesis – that it appears to relocate the problem of experience rather than resolve it. By proposing that proto-qualia arise at the cellular boundary, one might argue that the hypothesis merely places experience at the earliest possible stage of life, without explaining why experience exists at all.

This objection is understandable, but it rests on an expectation that the existence of experience should be derivable from purely third-person descriptions of physical and biological processes.

The Living Experience Hypothesis makes no such claim. The claim is instead a classificatory and ontological one – if experience exists anywhere in the natural world, it must arise within living systems, and it must do so at the point where an inside/outside distinction, normativity, and autonomous self-maintenance first coincide.

On this view, locating proto-qualia at the cellular boundary is not an attempt to hide an unexplained phenomenon, but to avoid introducing experience arbitrarily at later stages, such as with the appearance of neurons or human cognition. Nervous systems demonstrably transform and integrate experience, but they do not mark a principled ontological threshold at which experience could coherently be said to appear from nothing.

The Living Experience Hypothesis therefore does not claim to explain why experience exists in an ultimate sense. Rather, it aims to specify where experience must belong if it is to be understood as a natural phenomenon, and to rule out positions that either distribute experience indiscriminately across all matter or postpone it without justification until complex brains arise.

In this sense, the hypothesis does not dissolve the problem of experience but reframes it in a way that is consistent with biology, non-dualist physicalism, and the methodological limits of third-person science.

Multicellular life

It is now time to discuss **multicellularity**.

From the perspective of the Living Experience Hypothesis, the transition from unicellular to multicellular life represents a critical step not only in biological complexity, but in the integration of lived experience. In unicellular organisms, normativity and regulation are confined to individual cells. Multicellularity introduces a new level of autonomy, in which survival depends on coordinated regulation across many living units *sharing a common boundary and fate.*

In multicellular organisms, normativity does not vanish at the level of individual cells, but it becomes reorganized. Each cell continues to regulate its internal processes and remains sensitive to perturbations relevant to its own viability. However, this local normativity is now constrained by, and coupled to, a higher-level normativity belonging to the organism as a whole. Individual cells relinquish a degree of autonomous control in exchange for the stability, resources, and protection provided by the collective. What is "good" for a cell is no longer defined solely in terms of its own persistence or replication, but in terms of its contribution to the viability of the organism that sustains it.

This transition creates, for the first time, the biological conditions under which locally lived perturbations could be integrated into a higher-level experiential unity – a process later amplified by the evolution of nervous systems.

The fact that unicellular life persisted for billions of years before the emergence of complex multicellular organisms is telling. From the perspective of the Living Experience Hypothesis, this long delay suggests that integrating autonomous, normatively closed living units into a higher-level unity – potentially capable of unified experience – represents a deep biological constraint rather than an obvious evolutionary step. It also hints at the profound challenge involved in integrating autonomous single cells into a collective multicellular organism – a transition that required individual cells to give up some degree of autonomous control in exchange for the stability and support of the larger living whole.

Before looking further at how qualia may be integrated into higher-level consciousness, it is important to note a direct implication of the Living Experience Hypothesis: if experience, normativity, and intelligence originate with life itself, then even plants must be said to possess rudimentary forms of these capacities, insofar as they are living, self-regulating organisms. This view is consistent with a growing body of biological research showing that plants exhibit electrical and chemical signalling, adaptive responses to stress, learning-like priming effects, and coordinated regulation across tissues – all directed toward survival and growth.

There is, however, no reason to expect higher forms of awareness or cognition in plants. While plants exhibit sophisticated signalling and coordination – including forms of inter-plant communication and cooperation mediated by chemical cues and fungal networks – they lack a nervous system capable of integrating lived perturbations into a unified, temporally synchronized field of experience. Their intelligence and normativity therefore remain distributed rather than centrally integrated.

End of life

So far, the discussion has focused on living systems while they are alive. Before turning to higher cognitive functions, it is worth briefly considering what happens when life ends.

If autopoiesis – self-maintenance across a boundary separating inside from outside – is a condition for experience, then it follows naturally that a complex organism tends toward a unified first-person perspective – a single lived "someone," rather than a collection of separate inner lives. This is reflected in the fact that when an organism dies, it does not cease to exist in parts – it loses its integrated autonomy as a whole.

Death, however, is not an instantaneous event at every level. The moment at which a person is considered clinically dead typically corresponds to an irreversible loss of organism-level function – especially sustained circulation and respiration – and, crucially, the loss of the brain's capacity to maintain integrated activity. From the perspective of the Living Experience Hypothesis, this marks the end of unified, organism-level experience.

At the same time, many cells remain metabolically active for some time after clinical death, and some tissues retain partial function longer than others. If proto-qualia are tied to living cellular processes, then it is at least conceivable that proto-qualia could persist locally for a limited period even after unified experience has ceased. This would not imply the continuation of a personal consciousness or a surviving self, but rather a gradual dissolution of integration – organism-level experience ends, while local living processes fade on their own timescales.

This perspective offers a naturalistic way to approach reports of unusual experiences surrounding death without invoking anything supernatural. It suggests that what matters for the continuity of being someone is not the continued existence of cells in isolation, but the continued integration of living processes across the organism – a process that may degrade in stages rather than switching off instantaneously.

Nothing in this view implies the survival of a personal self or consciousness beyond death, nor does it require invoking any non-physical or supernatural processes.

Integration of qualia

At this point, it is relevant to ask the question if qualia are strictly inner, first-person aspects of living processes, how can a nervous system "integrate" them without accessing or representing them? Wouldn't that be a category error?

Our choice of the word "integrate" may be the reason for this concern. Integration is a third-person description of a first-person unification.

It is important to emphasize that, within the Living Experience Hypothesis, qualia are not integrated by the nervous system as objects or contents. Neither are qualia integrated as discrete contents, like tiles assembled into a mosaic. Rather, the nervous system integrates living processes across the organism, and the unified field of experience is the intrinsic, first-person aspect of this integrated regulation. No system ever accesses or combines qualia from the outside – experience is unified by being the unified process. Unified experience is what this integration is like from the inside.

An analogy may be that when oscillations in many neurons synchronize, nothing "collects" oscillations, and nothing "observes" synchrony. Yet a unified dynamical regime exists.

In a multicellular organism, individual cells do not merely coexist – they are dynamically coupled through chemical, electrical, mechanical, and metabolic processes. Over evolutionary time, these couplings become increasingly constrained and coordinated. Cells specialize, relinquish degrees of autonomy, and participate in organism-level regulation. What emerges is not a sum of experiences, but a reorganization of normativity – what matters is no longer primarily what preserves individual cells, but what preserves the organism as a whole.

Nervous systems dramatically amplify this integration. They do not create experience *ex nihilo*, but synchronize, temporally align, and hierarchically organize already lived processes across the body. Fast electrical signalling allows distant perturbations to be bound together into coherent patterns. Oscillatory dynamics allow multiple subsystems to participate in shared temporal structures. Learning stabilizes these patterns over time. The result is a single, temporally extended field of experience, structured but fluid, unified but differentiated.

On this view, there is no additional step at which "consciousness turns on." There is only increasing depth, scale, and coherence of integration. A human mind is not a container filled with qualia, but a living, embodied process whose internal organization gives rise to a richly articulated inner life.

Importantly, this also explains why consciousness degrades gradually rather than abruptly. Sleep, anesthesia, coma, and certain brain injuries do not eliminate life at the cellular level but disrupt large-scale integration. As integration weakens, unified experience fragments or fades – not because qualia disappear everywhere at once, but because there is no longer a process capable of binding them into a single perspective.

Unified consciousness, then, is not a mysterious extra property added on top of biology. It is the natural experiential aspect of deeply integrated living organization.

Consciousness and The Embodied Mind

We can now take the final step toward consciousness as humans experience it. Building on the idea that proto-qualia arise at the cellular boundary and are integrated, synchronized, and amplified by nervous systems, it becomes possible to connect the Living Experience Hypothesis to well-established findings in neuroscience.

Contemporary neuroscience has accumulated substantial evidence that large-scale integration in the brain depends on coordinated neural dynamics, including oscillatory activity across multiple frequency bands. Such oscillations appear to play an important role in binding distributed neural processes into coherent perceptual, affective, and cognitive episodes. While no single theory of consciousness has achieved consensus, these dynamical features are widely regarded as necessary conditions for unified experience.

Neurons, with their capacity for rapid electrical signalling and plastic adaptation, enable the integration of locally generated bodily signals into temporally coherent patterns. These patterns interact with memory systems and with mechanisms that balance expected inputs against real-time sensory input – a dynamic emphasized, for example, in Stephen Grossberg's *Adaptive Resonance Theory*. In parallel, affective and homeostatic systems contribute evaluative signals that shape what is salient or motivating for the organism.

From the perspective of the Living Experience Hypothesis, these neural mechanisms do not create experience from non-experience. Rather, they organize, integrate, and modulate experience that is already present in elementary form at the level of living systems. Once the origin of experience is grounded in life itself, the role of the nervous system can be understood as coordinating and amplifying lived perturbations into the rich, unified, and temporally structured field of consciousness familiar to humans.

The result is the remarkably detailed multisensory experience we live through – including sight, hearing, touch, taste, and smell, as well as more subtle aspects such as depth perception and the sense of moving through a three-dimensional space. Interwoven with this are thoughts, emotions, affects, and the experience of a self, which we will return to shortly.

In this sense, the Living Experience Hypothesis does not compete with neuroscience. It provides a conceptual framework within which existing neuroscientific findings can be situated, without requiring new physics or speculative entities, and without reducing experience to computation or representation alone.

Seen this way, proto-qualia originating in individual cells are integrated into a coherent experiential field at the level of the whole organism. If we accept that meaning and intelligence also originate in their most elementary forms at the cellular level, then these too must be integrated in parallel. The organism thereby acquires a lived capacity to distinguish events in the environment as beneficial, harmful, or indifferent for its continued existence, along with an integrated functional capacity – intelligence – to act accordingly.

This aligns with a widely held view in neuroscience – that the primary evolutionary role of the brain and nervous system is not abstract thinking or the construction of a self, but the regulation of behaviour in the service of survival. Cognitive functions, on this view, are elaborations of this more fundamental role.

At the same time, given the densely networked nature of bodily and neural systems, it is natural to expect emergent properties that go beyond mere survival. These emergent capacities bring us closer to our lived experience as humans – our abilities to create, invent, imagine, dream, love, hope, belong, become, and relate to one another and to the world.

If experience, meaning, and intelligence originate at the cellular level, then these capacities cannot be confined to the brain alone. They must be grounded in the entire living body, with the nervous system acting as an integrative and coordinating organ rather than the sole seat of mind.

This is also the form of embodiment with which we are familiar as living beings – a body that is not merely a physical structure, but one that remembers and acts intelligently, and that experiences pleasure and pain, presence and distance, desire and aversion, alarm and longing, anxiety and hope, love and fear.

From this perspective, experience does not cease during sleep. Rather, what is temporarily interrupted is the large-scale integration of experience into a unified, reportable consciousness. Local forms of lived activity – proto-experience at the level of individual cells and bodily subsystems – may continue, even though they are no longer bound together into a coherent first-person narrative. This distinction aligns with the phenomenology of sleep and dreaming, and with the observation that many regulatory and affective bodily processes remain active during sleep despite the suspension of waking awareness.

This brings us close to the concept of the *Embodied Mind*, introduced in the book of the same name by Francisco J. Varela, Eleanor Rosch, and Evan Thompson in 1992. As stated at the outset of this essay, the Living Experience Hypothesis supports the *enactive* approach presented in that work. According to enactivism, our experience of the “physical world” does not consist of internal representations of an external reality, but is enacted through the ongoing interaction of mind, body, and environment.

The authors emphasize that this enactment is inseparable from bodily action, sensory engagement, and cultural context. This aligns naturally with the LEH view that experience, meaning, and intelligence originate in living processes distributed throughout the body and are subsequently integrated.

LEH also supports the non-dualist position developed in The Embodied Mind, drawing both on cognitive science and on Buddhist meditative traditions. On this view, the mind is not separate from the body, and there is no functional need for a distinct inner observer or homunculus that receives experience. Instead, the experience of a self is enacted in much the same way as other aspects of experience.

From a scientific and philosophical perspective, a separate self – conceived as something ontologically distinct from embodied processes – would have no causal influence on the physical world and thus could not play any effective role in behaviour or experience.

The Embodied Mind notes that Buddhism arrives at a similar non-dualist conclusion, not through scientific investigation, but through disciplined introspective practice. By combining these approaches, the book aims to bridge the apparent deep gap between scientific accounts of the mind – which explain mechanisms but seem silent about experience – and lived experience itself, which is immediate and undeniable but opaque to mechanistic explanation.

The bridge consists in rejecting both dualism and representationalism, and in recognizing that mind and world arise together through embodied, situated activity. In this way, enactivism offers a middle path between scientific explanation and everyday experience.

The Living Experience Hypothesis supports this move but goes one step further by proposing a naturalistic origin of experience itself. It describes life and experience as two aspects of the same autonomous process – one accessible through third-person description, the other through first-person manifestation.

Scientific explanation operates through third-person compression – laws, models, and abstractions that preserve predictive structure while necessarily eliminating lived experience. If experience is understood as the inner, irreducible aspect of perturbations that matter for a living system's own viability, then it becomes clear why no third-person scientific description can ever fully encompass it. This should not be seen as a failure of science, but simply a consequence of its method.

From Wolfram's perspective, experience is not something outside the physical world. It is part of the same underlying logical structure – arising within those pockets of computational reducibility that make life and mind possible. There is nothing that observes life from the outside – life is something that is experienced by being it.

Non-dualism follows naturally from this view. The experience of a self emerges as part of the integration of experience, meaning, and intelligence into a higher-level conscious unity. The sense of self is therefore a construct, but a deeply natural one.

Yet, our experience of a self is natural and compelling. And even if it may ultimately be considered an illusion, and even if Buddhist meditative practice can reveal important insights into its constructed nature, it is not obvious that the everyday experience of a self must be contested or eliminated. If the sense of self can be explained in a fully naturalistic way – without invoking any separate observing entity – then its presence need not be seen as a philosophical problem.

In fact, experiencing the world from a richly detailed and highly coherent first-person perspective since we are infants, the appearance of a sense of self must be considered almost inevitable.

At the same time, the sense of self does not appear to be a necessary evolutionary adaptation. Survival and normativity may not require a reified self, but only autonomous regulation – the capacity to maintain internal organization in relation to the environment. The sense of self could therefore be better understood as an unavoidable interpretive consequence of richly integrated experience, rather than as a feature selected directly for its own sake.

The self then persists because nothing in ordinary experience contradicts it. Only disciplined introspection or abstract theoretical reasoning can reveal it as a construction rather than a fundamental entity.

In this way, the Living Experience Hypothesis supports enactivism and rejects dualism, while remaining neutral on whether the sense of self is something that must be overcome. It explains how the experience of a self arises, without prescribing how one ought to relate to it.

In this sense, the Living Experience Hypothesis may also help clarify our relation to the sense of self. As long as the existence of experience itself remains unexplained, the idea of a self tends to function as a metaphysical placeholder. The prospect of dissolving the self becomes psychologically and conceptually threatening not because experience would disappear, but because the mystery of experience would otherwise have nowhere to reside.

Here we may add another related philosophical consideration. A natural question arises when combining a block-universe perspective with lived experience: why does one's experience appear to be located at a particular moment in time rather than another?

On closer inspection, this question presupposes a standpoint outside the experience of time from which such a comparison could be made. For beings like us, whose existence consists in living processes extended over time, no such standpoint is available. Any attempt to adopt it necessarily introduces another implicit present, and with it another experience of time.

The resulting paradox reflects not a gap in explanation, but a structural limitation of first-person cognition. A mind like ours, which depends on the experience of time to function at all, cannot step outside that experience without contradiction. Any such attempt leads to a perspective that is inconsistent with the very assumptions under which the question is posed.

A partial way of approaching this may be to note that a mind like ours must find itself at some point within experienced time, from which the past is remembered, and the future appears to unfold. The point in time where we find ourselves is one such point — and no further selection or explanation is available from within experience itself.

Imagination, dreaming, and human intelligence

Before looking at human level intelligence, it is useful to consider another highly useful capacity of the mind – imagination – that may also be understood as a natural consequence of the high-level integration of experience, meaning, and intelligence into consciousness. If we consider this integrated process as a system drawing continuous input from the entire body, we can imagine what happens when this input is partially decoupled from immediate sensory engagement. The system is then able to bring forth experiences that are not directly present in the surrounding physical environment – in other words, imagination. The very act of entertaining this idea is itself an example of that capacity.

Dreaming may be understood in a related way. Rather than being driven primarily by current sensory input, the integrating system appears to draw freely on earlier experiences, memories, and affective patterns. A growing body of research suggests that dreaming may serve functions such as consolidating memory, exploring emotional scenarios, or reorganizing experience, all of which are consistent with this view.

Imagination – and possibly dreaming as well – thus opens a broader perspective on biological and human intelligence, extending adaptive engagement beyond the immediately present moment.

Let us the consider human level intelligence.

As we have seen, intelligence in its most fundamental form can be understood as the functional capacity to act on meaning to sustain life. Higher forms of intelligence extend this capacity across longer timescales and richer domains, allowing living systems to anticipate, explore, and refine their engagement with the world.

One important aspect of this extended intelligence may be described as *truth-seeking* – not in an abstract or absolute sense, but as the ongoing effort to grasp the regularities and constraints of the surrounding world. Another closely related aspect is *problem solving* – the capacity to navigate obstacles, conflicts, and constraints in ways that preserve or enhance viability.

A central component of both truth-seeking and problem solving is the acquisition of a practical understanding of *causality* – the ability to anticipate how actions and events are likely to unfold, and to adjust behaviour accordingly.

From this perspective, we can outline a simple, experience-based cycle through which humans engage in truth-seeking and problem solving:

1. **Perception**. Originating in proto-qualia at the cellular level and integrated into a rich multisensory field – including affects and emotions – perception provides a lived experience of the environment and of the body itself, extended over time and space.
2. **Imagination**. Building on perception and on prior experience of causal regularities, imagination enables the formation of hypotheses and counterfactuals – possible actions, alternative scenarios, and anticipated outcomes that are not immediately present.
3. **Embodiment**. Through bodily action, these imagined possibilities are tested. The body becomes the means through which hypotheses about the world are enacted.
4. **Incremental learning in real time.** The outcomes of action are perceived and continuously integrated with prior expectations, refining future perception, imagination, and action

Although it may be tempting to describe this process as the brain constructing an internal *"world model"* – a representation of external reality and its dynamics – the earlier rejection of representationalism and dualism suggests a different interpretation. Rather than modelling the world from a distance, this cycle can be understood as a lived organization of experience, shaped by memory, expectation, and real-time sensation, through which the world is directly engaged.

It is worth noting that the remarkable plasticity of the human mind – its ability to learn continuously while selectively reinforcing what matters and letting other experiences fade – is itself a complex and not yet fully understood process. This capacity is arguably as important as perception, imagination, and action in enabling adaptive intelligence.

It is also important to emphasize that embodiment is not confined to moments of overt action. As argued in The Embodied Mind, and as naturally implied by the Living Experience Hypothesis, the entire body is continuously involved in interacting with and discovering the world, including its own internal states.

Finally, beneath even the most abstract forms of imagination and understanding lies a fundamental and non-negotiable driver – the same one present in the earliest forms of life.

The incentive to explore, learn, and make sense of the world is ultimately rooted in the need for survival.

Thought, language, culture, and techné

Building on this cycle of perception, imagination, embodiment, and learning, we can now turn to what is commonly called *thought*. In the context of the Living Experience Hypothesis, thought is not a fundamental capability, nor a process of manipulating internal representations.

Rather, thought can be understood as the gradual stabilization, structuring, and social coordination of imagination and experience. Where imagination explores possibilities, thought renders them durable – through memory, language, narrative, and shared symbols – enabling coordination and sharing across longer timescales and broader social domains.

From this perspective, thought may be defined as the stabilized, shareable, and recursively refined organization of experience and imagination that allows a living system to coordinate action across extended timescales, social contexts, and abstract domains.

While imagination is private, fluid, transient, experiential, and often non-linguistic – simulating possibilities and exploring counterfactuals – thought is stabilized, remembered, reused, refined, and shareable. Thought is also typically constrained by norms such as true and false, right and wrong, or coherent and incoherent.

Thought may further be refined into specialized forms such as narratives and concepts.

A *narrative* can be understood as a temporal structuring of experience and meaning, connecting causality, identity, and action.

Concepts, by contrast, may be seen as compressions of repeated patterns of experience – tools for efficient coordination, particularly in scientific and technical contexts.

It now becomes natural to discuss language, symbols, and culture as extensions of thought. But before doing so, it is important to consider why such extensions are necessary for groups of people.

From the perspective of the Living Experience Hypothesis, groups of humans do not constitute a single living system in the same sense as an individual organism. They lack a shared biological boundary, a unified metabolism, and a common autopoietic organization. As a result, groups cannot integrate experience, meaning, and intelligence into a single inner perspective. There is no collective first-person experience in the strong sense – no shared what it is like to be the group.

This does not mean, however, that groups cannot act intelligently or adaptively. Rather, collective intelligence operates differently. Instead of being grounded in a unified inner experience, it emerges through coordination among distinct living systems, each with its own experience, normativity, and agency.

To collaborate effectively – and to increase resilience, problem-solving capacity, and adaptability – human groups require mechanisms for communication and coordination. Such collaboration is possible even without language, as seen in many social animals. However, language, culture, and shared symbolic systems dramatically increase the efficiency, scope, and temporal reach of collaboration.

Language allows experiences, intentions, and insights to be externalized and aligned across individuals. Culture preserves and refines these shared structures over time. In this way, thought becomes stabilized and distributed across a social network, even though experience itself remains irreducibly individual and embodied.

From this perspective, collective intelligence is not the integration of inner lives into a higher-level subject, but the coordination of many distinct subjects through shared practices, symbols, and material structures. Language and culture do not create a collective mind – they enable collective action.

History offers abundant evidence of the astonishing results humanity has achieved through such collaboration – from scientific advances and technologies that support and extend human activity, to cultural expressions in art, literature, and social life. At the same time, the same capacities have enabled destruction, suffering, and large-scale damage to the planet's ecosystems.

Much of this has occurred using technology, or techné, which may be defined as thought made durable and externalized in physical form – through media such as mechanics, chemistry, electricity, electronics, optics, electromagnetism, and nuclear processes.

Because technology does not experience, it cannot adapt through experience. Technologies are therefore frozen solutions to existing or past problems. Another way to express this is to say that technology is causal knowledge externalized into tools – thoughts that no longer need to be thought.

Technology can therefore extend human agency, increase power and precision, and reduce cognitive load. But because it is not living, it does not possess normativity or meaning. It has no experience, and no stake in outcomes.

Here it is important to note that human problem-solving through intelligence, and the externalization of thought into technology, long predates spoken language. The earliest stone tools were used by our ancestors approximately 2.2 million years ago, and controlled use of fire dates back at least one million years. Spoken language, by contrast, is generally thought to have emerged much later, perhaps between 50,000 and 100,000 years ago.

From the perspective of the Living Experience Hypothesis, this temporal ordering is significant. Intelligence, in its most fundamental form, originates with life itself and thus vastly predates both technology and language. Even within human evolution, however, this observation highlights that language – while extraordinarily powerful as a medium for coordination, collaboration, and cultural transmission – is a comparatively recent development. The capacity to solve problems and to externalize solutions into material form arose long before the emergence of symbolic communication.

If meaning, in its most fundamental form, likewise originates with life, it follows that words and language do not create meaning from nothing. Rather, they connect to, shape, and reorganize forms of meaning that are already present in lived experience. Language can refine, extend, and transform meaning – making it explicit, shareable, and cumulative – but it does not generate normativity or significance independently of life. Meaning is grounded in what matters to living systems. Language provides a powerful means for articulating and negotiating that meaning across individuals and cultures.

In other words, language did not create intelligence or techné – instead, it radically amplified their reach, efficiency, and persistence across individuals and generations.

Artificial Intelligence and robots

This brings us to the much-discussed field of artificial intelligence.

When the authors of The Embodied Mind began writing their book in the mid-1980s, confidence in the progress of AI toward human-level intelligence was approaching its peak. The dominant approach at the time was symbolic AI and expert systems – a representational conception of intelligence in its most explicit form. Intelligence was assumed to consist of manipulating symbols according to formal rules, much as many researchers believed the human brain did.

Corporations and governments alike expected rule-based AI systems to scale across domains such as medicine, finance, and manufacturing, and to replace significant aspects of human expertise.

Public and internal expectations were high. Many researchers and managers believed that general problem-solving intelligence was a near-term engineering challenge rather than a deep conceptual one.

This was precisely the moment when representationalism appeared strongest — and therefore when a critique of it carried the greatest force. Although The Embodied Mind was published a few years later, in 1991, its rejection of representationalism and its turn toward embodied and enactive approaches directly challenged the assumptions underlying symbolic AI.

In hindsight, the gap between expectations and reality was widest toward the end of the 1980s, which helps explain why the collapse that followed was so abrupt. Expert systems proved expensive to build and maintain, brittle outside narrowly defined domains, and unable to deliver the broad, flexible intelligence that had been promised. Predictions of human-level AI made in the 1970s and early 1980s clearly failed to materialize.

As a result, governments and investors pulled back. What became known as the “AI winter” set in, lasting for roughly a decade.

Gradually, the field recovered by lowering its ambitions. Rather than aiming to replace human intelligence, AI research shifted toward probabilistic methods instead of symbolic rules, and toward concepts such as pattern recognition, statistical learning, optimization, and prediction. It also became clear that scaling data and computation often mattered more than devising clever algorithms.

This shift led to many successful niche applications, and to a few highly visible demonstrations that again raised hopes of human-level intelligence. These included IBM Deep Blue defeating Garry Kasparov in chess in 1997, IBM Watson winning against human champions in *Jeopardy* in 2011, and Alpha Go surpassing human performance in the board game Go in 2016. Each of these achievements renewed speculation about human-like intelligence, even though they were realized in narrowly constrained domains.

With the advent of large language models, LLMs, and the public release of ChatGPT in 2022, expectations surrounding artificial general intelligence, AGI, once again began to rise toward a peak.

This resurgence has been followed by a familiar line of critique, reminiscent of the debates of the 1980s – fundamentally centred on the absence of genuine understanding. At its core, this critique once again concerns representationalism. Although large language models do not manipulate explicit symbols representing objects and events, they nonetheless operate on tokens – numerical values – that lack intrinsic meaning. Any apparent meaning arises only through statistical correlations learned from vast corpora of human-generated language.

From this perspective, understanding in such systems is simulated rather than lived.

The Living Experience Hypothesis extends this critique further. Not only is contemporary AI based on representations without intrinsic meaning. It is trained primarily on language – which, as we have seen, does not create meaning, even though it can refine, extend, and transform meaning grounded in lived experience.

Human problem-solving predates language by at least two million years, and if intelligence and meaning originate with life itself, they predate language by billions of years. In this sense, intelligence cannot be built on language – it's the other way round.

A further limitation follows from the absence of normative interiority. Lacking an intrinsic sense of *better or worse*, AI systems have no lived stake in outcomes. They do not care whether their responses are correct, coherent, or truthful. Error correction occurs only through externally imposed objectives and feedback, not from within. When such systems are chained together, errors may accumulate or compound without any intrinsic mechanism for concern or correction.

We can also revisit the experience-based cycle for truth-seeking and problem-solving outlined earlier. AI systems lack lived perception of time and space, lack imagination in the experiential sense, lack embodiment – aside from robotic systems that still possess only simulated bodies without lived bodily experience – and lack incremental, real-time learning. Large language models, for example, are vast but essentially frozen structures during use.

As a result, they lack a genuine method for grasping causality, which severely limits their capacity for open-ended problem-solving and truth-seeking, especially in novel or unforeseen situations.

Efficiency presents another fundamental issue. As discussed earlier, evaluating meaning through calculation and stored representations is computationally expensive and adaptively slow compared to lived experience. This is one reason to conclude that life requires inner experience to survive in open-ended, continuously changing environments. Any purely computational, representational approach to general intelligence will therefore tend to be resource-intensive and inefficient relative to biological intelligence.

From the perspective of the Living Experience Hypothesis, AI is best understood as *techné*: externalized thought. More precisely, it is simulated thought that has been automated. AI is an extraordinarily powerful form of techné, and immensely useful for many purposes. But this framing also reveals its inherent limits – lack of intrinsic normativity, lack of lived meaning, lack of genuine adaptivity, lack of stake in outcomes, lack of autonomy, and dependence on externally imposed goals.

One might object that AI could simply represent a different kind of intelligence – distinct from biological intelligence. To address this, we can return to Wolfram's hypothesis of the underlying logical structure of the universe. It is conceivable that forms of intelligence exist outside our pocket of computational reducibility, forms that may be difficult or impossible for a mind like ours to comprehend.

However, within our pocket of computational reducibility, intelligence appears to have evolved for a specific reason – to sustain autopoietic life. From the perspective of the Living Experience Hypothesis, this is the only form of intelligence that carries intrinsic meaning and thus is meaningful to us. This may appear circular, since meaning, intelligence, and life are proposed to arise together. But this circularity is not vicious – it reflects the closed normative loop that constitutes life itself.

In practical terms, any intelligence we would recognize as genuinely intelligent in a meaningful way must share the defining features of biological intelligence – intrinsic normativity, a stake in outcomes, and an experiential grasp of causality, which in turn arguably requires experienced perception of time and space, imagination grounded in lived experience, rich embodiment, and incremental real-time learning.

Most likely, this would require building synthetic life – creating organisms, not machines.

In other words, unless expectations are reduced, we may expect a new collapse in the AI field, and a new AI winter.

Meanwhile, the risk is that we mistake performance for intelligence, simulation for experience, and language for meaning – in other words, that we overvalue AI and give it a level of trust that we normally reserve for humans. The effects of this are already well known and widely discussed – lack of accuracy outside narrow domains, bias, lack of transparency and responsibility, cognitive offloading, deskilling, emotional dependence, and what has been called AI slop – large amounts of low-quality content that risk contaminating processes, work, collaboration, and culture more broadly.

As for humanoid robots, which are becoming increasingly flexible and capable of moving in ways similar to humans – and in some respects beyond – a related illusion and risk is emerging – that we mistake flexibility, superficial versatility, and simulated responsiveness for genuine adaptivity and problem-solving.

Equipped with today's forms of AI for communication and decision-making, and lacking the fundamentally rich, experienced, and lived embodiment of biological organisms, robots remain techné – simulated and automated thought combined with physical action. They can be highly effective in niche tasks and controlled environments, but in open-ended, continuously changing situations they remain brittle, divergent in behaviour, and resource-intensive rather than genuinely adaptive.

Overall, there are strong and well-grounded reasons to remain cautious about the potential capabilities and general use of AI and robotics, as long as they lack anything resembling synthetic life.

LEH and Integrated Information Theory

At first glance, the Living Experience Hypothesis may appear related to *Integrated Information Theory*, IIT, developed by Giulio Tononi, which proposes that consciousness corresponds to the degree of integrated information within a system. Both views emphasize integration and reject simple localization of experience to isolated components.

However, the two approaches differ in a crucial respect. Integrated Information Theory treats information integration itself as sufficient for experience, independently of whether the system is alive. The Living Experience Hypothesis does not. In LEH, integration matters because it arises within living, self-maintaining, normatively organized systems. Information becomes experiential only insofar as it matters for the continued existence of a living process.

From this perspective, integration is not the origin of experience but a way of organizing and amplifying experience that already arises with life. This distinction places LEH firmly outside panpsychist interpretations and avoids attributing experience to non-living systems solely based on structural complexity.

LEH and a Computational Universe

The Living Experience Hypothesis is also compatible with broader views that treat the universe as fundamentally mathematical or computational, such as those proposed by Max Tegmark. Like Wolfram's framework, these views suggest that physical reality may ultimately be described as an abstract logical or mathematical structure.

LEH does not challenge this possibility. Instead, it addresses a different question: how experience arises within such a structure. A universe that is mathematical or computational in nature does not, by itself, explain why any part of it should be lived from the inside. LEH proposes that this occurs only in rare regions where living, autonomous organization emerges, bringing with it normativity, meaning, and inner experience.

In this sense, LEH can be seen not as a competitor to computational or mathematical universe hypotheses, but as an attempt to locate life and experience within them.

End notes

We have seen that by relying on existing theories and findings, and by adding two specific claims about the origin of inner experience, it becomes possible to propose a coherent and consistent account of lived experience and the universe – without the need for creation stories or a privileged observer, and with the single assumption that not everything is possible.

This account extends from the most fundamental level of the universe to rare regions where structure is sufficiently ordered, predictable, yet rich and open-ended, for experience, intelligence, and meaning to arise together with life. One might describe this, metaphorically, as a logical structure trying to understand itself.

The Living Experience Hypothesis is not offered as a final theory, nor as a solution to every philosophical puzzle surrounding mind and existence. It is an attempt to take experience seriously without abandoning naturalism – and to take science seriously without denying what it necessarily leaves out.

At its core, LEH begins from a simple observation – we are living systems that find ourselves already inside experience. Any account of reality that explains everything except this fact is incomplete. At the same time, any account that explains experience by stepping outside the natural world risks losing explanatory contact altogether.

The view presented here tries to hold both truths at once. Life is not a special interruption of physics, and experience is not a supernatural intrusion into matter. They are two aspects of a process that unfolds wherever autonomous living organization arises within a world constrained enough to sustain it.

This perspective has consequences. It challenges familiar intuitions about intelligence, consciousness, and artificial systems. It suggests humility about the reach of computation and simulation. And it invites a renewed respect for life – not because life is mystical, but because it is the only place where meaning, concern, and understanding can exist at all.

Finally, LEH does not demand any existential or ethical conclusion. It does not instruct us to dissolve the self, nor to cling to it. It does not promise transcendence, nor deny depth. It simply proposes that being alive already places us inside the most remarkable phenomenon we know – a world that is not only structured but felt.

Some may find this view unpoetic, seeing in it a reduction of mystery or spiritual imagery. Others may find in it a sense of inner calm, and a reason to wholeheartedly celebrate the extraordinary aspects of life, nature, and the human mind – with everything experience makes possible, from curiosity, innovation, science, and creativity to art, music, literature, and the simple desire to touch and be touched. And, not least, to find direction amid uncertainty, which has arguably been our guiding force throughout humanity, and remains so today.

If this hypothesis succeeds, it will not be because it eliminates mystery, but because it places mystery where it belongs – not outside nature, not beyond science, but at the very heart of living existence itself.

Hopefully, it will inspire others to further explore its meaning and implications.

Mats Lewan, January 11, 2026

Addendum on the experience of time

January 14, 2026

This addendum is the result of further reflection clarifying a question left implicit in the original text of The Living Experience Hypothesis – how minimal temporality arises together with normativity and qualitative experience in living systems.

As discussed in the original text, there is good reason to hold that our experience of time is a construct of the mind, just as other qualia are. The experience of time, however, has a particular position, since all qualitative experiences are time-dependent in one way or another. In fact, as noted earlier, it would not be possible for a mind like ours to exist without the experience of time. Nor can we fully comprehend a timeless structure, such as the proposed logical structure underlying the universe, or a block universe. We can imagine such structures, but only at the cost of implicitly introducing an external observer and an external flow of time.

It is therefore particularly challenging to imagine how experience of time may arise from a timeless structure. We have already touched upon a related topic – why we have the experience of living at a particular point in time and not any other – and concluded that no explanation is available from within experience itself.

We will now discuss where and how a minimal experience of time may emerge from the perspective of the Living Experience Hypothesis.

First, we revisit the topic of identity. As mentioned earlier, Stephen Wolfram proposes that in pockets of computational reducibility, where predictability exists, elementary particles can be understood as large regions of stable patterns that remain unaltered through the continuous application of rules—in other words, patterns that remain stable over what may be interpreted as time. In this sense, particles acquire an identity that is stable over time.

A similar consideration may be made about living organisms. Living organisms may also be considered large regions of patterns that remain stable over time, but while particles remain stable because they are invariant and unaltered, living organisms remain stable through self-regulation. This also implies a form of identity.

Using Wolfram's concept of stable patterns – "lumps" – particles may be described as static-stable lumps: stable structures that persist through invariance. Organisms, by contrast, may be described as dynamic-stable lumps: stable processes that continuously change internally, maintaining identity through regulation rather than invariance.

Another way to express this is that particles have identity because they remain the *same*, while living systems have identity because they remain *themselves*.

The way living systems remain themselves – how they manage self-regulation – is through normativity, which, as proposed by LEH, is not represented but experienced through the distinction between perturbations that are negative, supportive, or indifferent with respect to the ongoing viability of the organism's living process.

At first glance, both qualitative experience and normativity appear to be time dependent. As we will see, however, it may be more appropriate to understand this relationship as one of interdependence. Let us begin with normativity.

Viewed as stable processes within a pocket of computational reducibility in the underlying logical structure, normativity consists in how a living system responds to successive states or perturbations. A new state, or sequence of states, may be supportive, threatening, or indifferent with respect to continued viability. The organism responds by altering its internal dynamics in ways that either increase or decrease the likelihood of continued exposure to such perturbations – or, in the case of indifference, by remaining unchanged. This is also where directionality first appears.

Indifference can be understood as reinforcing identity, although identity itself is already established through the boundary of life and through persistence across change. Identity arises with autopoiesis, and indifference is the expression of that identity: the capacity of a system to remain itself across change. Without indifference, there is no persistence, and without persistence, there is no identity.

Primitive life is unlikely to possess memory or predictive capability, and LEH rejects the viability of internal representation at this level. This means that a primitive living organism cannot reasonably compare two adjacent states. How, then, does normativity operate?

The answer lies in the history of the organism – a history it does not need to be aware of.

As discussed earlier, the emergence of life may be described as largely stochastic, involving trial and error under constraints imposed by natural laws. According to LEH, this process has no inner aspect and precedes the emergence of life, normativity, meaning, and intelligence.

Once life is established, it can be described from a third-person perspective as an automated process that carries the history of its evolution in its own structure. This history constrains how the organism can respond to perturbations, allowing only responses that keep the process within a certain range – within the margins of continued viability. Otherwise, it would not be a living organism capable of survival.

These constraints are not only historical. The organism's internal dynamics are continuously altered as it adapts to a changing environment, always within the margins of viability. The constraints are therefore dynamic, but always present.

While this appears as automation from a third-person perspective, LEH proposes that this process gives rise to an inner lived experience. Let us now focus specifically on the experience of time.

The constraints an organism must obey to remain alive may be compared metaphorically to tension or force. Forces and tensions are, in a sense, timeless – they exist at each moment without themselves implying change. Yet, when viewed from a temporal perspective, they hold the potential for movement when equilibrium is disturbed.

In a similar way, the constraints a living organism must obey – built up through its evolutionary and developmental history and through ongoing adaptation – hold the potential for temporality. At each moment, in each state within the underlying logical structure, a living organism, unlike a particle, exists under an ongoing tension – a requirement to continue.

This requirement is not something the organism knows or represents. At the most basic level of life, it is embodied in the organism's organization. It constrains what can happen next, limiting the immediately possible states to those compatible with continued viability.

From a third-person perspective, this appears as automatic, mechanical regulation. From a first-person perspective – which LEH suggests begins already with life itself – the same constraint can be understood as a minimal form of temporality, or proto-time.

Proto-time is thus the inner lived aspect of the constraints a living organism must obey in every moment to remain within the margins of viability.

Proto-time is not an awareness of succession or duration. It is the lived fact that the present state cannot simply remain as it is without further regulated continuation. The minimal experience of time is not a moment – it is a meaningful transition.

One might object that the organism could simply "ignore" the next step and remain in the present state. But this objection quietly introduces an external observer with an external notion of time. From within the system, a single timeless state cannot meaningfully exist on its own. We can only understand such a state as an abstraction – a slice taken from an ongoing succession of states – and neither can the organism. A living system does not exist in a state; it exists through states. Existing across successive steps is therefore not an added assumption but a necessary condition of being alive at all.

If proto-time allows for a minimal temporal experience grounded in constraint at each moment, it may also serve as a link between successive moments. The arrow of time then arises, as discussed earlier, from the fact that life operates through a coarse-grained perspective within a pocket of computational reducibility. Information is lost at each step in the underlying logical structure, and the system must therefore obey entropy and the second law of thermodynamics.

Having established proto-time, we can now return to its interdependence with qualitative experience and normativity.

If changes that matter to a system are experienced at all – as supportive, threatening, or indifferent to continued existence – those changes are necessarily experienced as ordered and asymmetric. Temporality is therefore not something perceived in addition to meaning and qualitative experience; it is the form that meaningful experience takes in a living system.

In this sense, the experience of time is not an independent feature added on top of life. It is inseparable from qualitative experience and normativity.

Proto-time, proto-qualia, and normativity are therefore mutually dependent. None can arise on its own, and none can be reduced to the others. Together, they constitute the minimal interior perspective of life – an inside from which persistence matters, change is not neutral, and existence is lived as a continuing requirement rather than a static state.

Building on the earlier discussion of how the nervous system expands and integrates proto-qualia into higher-level experience and consciousness, the same can be said of proto-time. Integrated with memory, imagination, and predictive capacity, proto-time develops into the rich and coherent temporal experience with which we are deeply familiar.

Once again, this perspective highlights both the richness and the rarity of the human mind, enabling forms of experience and meaning that are not immediately necessary for survival, yet arise as a consequence of life itself.

Mats Lewan, January 14, 2026.