# Consciousness And Its Spacetime Headset

Spacetime is one headset of many by which consciousness explores itself.

I hold a warm croissant, drizzled with chocolate, sprinkled with powdered sugar, topped with slivers of toasted almond. It's light, mostly air, with a delicate crust. I savor its oven-fresh aroma, then listen to whispered crackling as I bite. Thoughts stop for a moment. Nothing but flavor.

We can taste thousands of flavors, see millions of colors, smell trillions of odors, hear thousands of sounds, and feel myriad bodily sensations. This profusion of experience is crucial data for science, and a sine qua non for human existence.

## Physicalist Theories Do Not Explain Any Conscious Experience

Most scientific theories of conscious experience assume that spacetime and its objects are fundamental. According to such theories, there were no conscious experiences at the Big Bang some 13.79 billion years ago; there was just spacetime, from which emerged the currently observed particles and fields. Then for billions of years these coalesced into the panoply of objects we now observe, including stars, planets, black holes, and galaxies. Earth appeared 4.5 billion years ago, and life 800 million years later. Consciousness and conscious experiences evolved sometime after that. We have no clear evidence yet of life or consciousness elsewhere in the universe. For all we know, we're alone.

In this spacetime framework, conscious experiences somehow emerge from unconscious physical ingredients. But not just any ingredients. Most theories propose that physical systems must have the right causal or functional properties to generate, or instantiate, conscious experiences. A scrambled egg probably doesn't have the right properties, but perhaps a brain or properly programmed computer does.

Researchers have proposed many causal or functional theories. We can't canvas them all. So let's narrow our focus. Which theories are most promising? Which succeed in explaining the most conscious experiences? For instance, which functionalist theory specifies a precise functional architecture that must be the smell of chocolate? What causal theory delineates a causal structure that must be the taste of toasted almonds, or the experience of haptic space?

Put simply, we expect scientific theories to pick specific cases and explain them. There are trillions of experiences. Finding one to explain should be like shooting fish in a barrel.

So, how are physicalist theories of consciousness progressing? How many experiences do they explain?

Zero.

That's right. They fail to explain a single experience. This wouldn't surprise Gottfried Leibniz. In his 1714 book, The Monadology, he argued that this failure is principled. I agree. The scientists seeking physicalist theories of conscious experiences are my friends and colleagues. They are brilliant and hard working. Their failure is principled: No brilliance can square the circle.

This failure might not surprise many high-energy theoretical physicists. Quantum field theory and general relativity together entail that spacetime cannot be fundamental. It ceases to have any operational meaning at the Planck scale: roughly $10^{-35}$ meters and $10^{-43}$ seconds. The reason is straightforward. To resolve smaller objects you must use shorter wavelengths. Shorter wavelengths have greater energy. Greater energy means greater mass. When wavelengths shrink to the Planck scale, the density of mass is so great that it collapses spacetime into a black hole. Trying harder, with smaller wavelengths, just makes the black hole bigger.

So high-energy theoretical physicists now search beyond spacetime for deeper foundations that might spawn spacetime and its objects. There they have found fascinating and puzzling structures such as "causal sets" and "positive geometries."

This failure of spacetime is also predicted by Darwin's theory of evolution. Our senses evolved to guide adaptive behavior, to keep us alive long enough to raise offspring. We might guess that our senses evolved to report relevant truths about the world around us. After all, wouldn't accurate perceptions make us more fit?

No. Remarkably, a mathematical formulation of Darwin's theory, called evolutionary game theory, entails that the probability is zero that any sensory system of any creature has ever been shaped to see any true structure of reality. The basic mathematical idea is remarkably straightforward. The fitness functions that guide evolution are not constrained to be "homomorphisms" of structures in reality. That is, they are not constrained to faithfully preserve such structures. So there are no selection pressures to perceive any true structures of reality. Our perception of space and time, and of objects in space and time, is not a veridical picture of reality. It's more like a VR headset that evolved to let us play the game of life. In Grand Theft Auto, for instance, you appear to drive a car in a world of streets, buildings, and traffic lights. But behind this veneer you are toggling millions of voltages in an unseen computer.

This view of evolution may seem self refuting. Darwin's theory assumes that there are organisms in space and time that compete for mates and resources. But I just claimed that Darwin's theory entails that spacetime and its objects are almost surely not fundamental reality. So one might argue that I've shot myself in the foot, that I used Darwin's theory to disprove Darwin's theory. I should learn basic logic.

This objection misunderstands science and its evolution. Each scientific theory starts with assumptions. It says, “Please grant these assumptions. Then I can explain all these interesting things.” A competent theory indeed explains those things. But it cannot explain its assumptions. One can propose new, deeper, assumptions and construct a new theory that explains old assumptions. But this new theory cannot explain its new assumptions.

This is good news. It means there’s infinite job security in science. There is not now, and never will be, a final Theory of Everything. Our best theories will always explain zero percent of reality. This does not mean theories are worthless. Each competent theory has some explanatory scope. But each also has explanatory limits. A competent theory gives you tools to explore its scope. A great theory gives you tools to pinpoint its limits—the explanatory confines of its assumptions.

Darwin’s theory is great. When made formal, with evolutionary game theory, it reveals the limits of its foundational concepts: organisms competing for resources in space and time. These concepts are not the ultimate nature of reality. But we knew that already. No theory starts with ultimate assumptions that entail the final account of reality.

We see this play out in high-energy theoretical physics. Quantum field theory and general relativity assume that spacetime and its objects are fundamental. They then prove that spacetime falls apart at the Planck scale. Great news. We knew it had to happen somewhere. The theory is deep enough to tell us where: $10^{-35}$ meters. Now we can explore new theories beyond spacetime, such as causal sets and positive geometries.

Spacetime ends at the Planck scale. Since spacetime is not fundamental, neither are objects within spacetime. Most theories of consciousness ignore this fact of physics. They assume that spacetime and physical objects, such as neurons and microtubules, are fundamental, and then propose that objects with the right properties cause, or are, conscious experiences. This proposal is false, as Leibniz understood centuries ago and physics underscores today.

Spacetime is a headset. Objects in spacetime are icons that appear in the headset. We render them when we look. They cease to exist when we look away. This is counterintuitive to most people, and raises obvious objections.

For instance: If you think that speeding truck is just an icon, why don’t you step in front of it?

I wouldn’t for the same reason I wouldn’t carelessly drag a file icon on my desktop to the trash icon. Not that I take the icons literally. My file is not the color and shape of its icon. But I do take the icons seriously. If I’m careless I could lose my file.

That’s a lesson from Darwin. Our senses evolved to guide adaptive behavior. Take them seriously. If I see a crocodile, don’t hug it. If I see a cliff, don’t jump off.

But don't take them literally. The probability is zero that what I see resembles anything in objective reality, whatever that reality might be.

There are many similar objections. What about a car that hits me from behind? Won't it kill me even if I don't see it? And doesn't that prove it's real? Or how about a photograph? If my photo shows a car, even though I didn't notice the car, surely that proves it's there when I don't look? To answer such questions, simply imagine a VR version of the question. In no case will you find that your perceptions must match or resemble reality.

Darwin's theory provides good reasons to conclude that our beliefs and our perceptions of objects in spacetime are not necessarily the truth. Objects are just useful icons in our headset. But this idea that objects are just icons is not new. Plato sketched a low-tech version in his allegory of the cave.

So scientific theories of consciousness that start with causal or functional properties of objects in spacetime, such as neurons, fail for principled reasons. That project is impossible. Objects in spacetime do not exist when they are not perceived. Like objects in a VR headset, they are rendered when we look, and disappear when we glance away. Neurons don't create consciousness because they don't exist when they're not perceived.

This means we need more funding for neuroscience, not less. Neuroscience is far more complicated than we now imagine. We think that the brain is complex, with its 86 billion neurons and untold trillions of synapses. But the brain is just an interface depiction of a reality beyond spacetime that is infinitely more complex.

There are many correlations between neural activity and conscious experiences. But they do not entail that neural activity causes conscious experiences any more than correlations between rooster crowing and sunrise entail that crowing causes sunrise. Yet these neural correlates are essential data if we are to reverse engineer our headset, for one key reason: The brain is our headset's representation of how the headset is engineered. So cognitive neuroscience is critical to understand this engineering.

## Theories Of Consciousness Fail To Explain Spacetime Or Matter

Physicalist theories fail to explain consciousness. Their failure is principled. What about theories that propose consciousness is more fundamental than spacetime, and somehow creates spacetime? Such theories are at the heart of ancient philosophies, idealism, and scientific investigations into anomalous phenomena possibly beyond the scope of standard physics.

To be taken seriously by science, these theories must pass tough tests. Their claim that spacetime and objects emerge from consciousness must be rigorous, with a formal description of consciousness and its dynamics. It must explain how the Minkowski space of special relativity arises, with its slowing clocks and contracting rulers in moving inertial frames. It must explain why the speed of light is the same in all inertial frames. It must show how the elementary particles of the Standard Model arise, how they get their masses, and how mass curves spacetime. It must explain why massive objects cannot travel at the speed of light. These are just starters. Physics is an advanced field, with a wealth of theory and experiment to be explained by any theory that claims consciousness is more fundamental.

Accounts of consciousness have been around for millennia. So how are they doing? How many can derive Minkowski space, explain why there is no preferred inertial frame, and explain why the speed of light is the same in all inertial frames?

Zero.

That's right. None. Theories of consciousness do not boot up spacetime any better than theories of spacetime boot up consciousness. The failure of physicalist theories is principled: spacetime and objects are undefined below the Planck scale, and don't exist when not perceived. Perhaps the failure of consciousness theories is also principled.

But perhaps not. If spacetime and objects are forms of conscious experience, as idealist and spiritual traditions propose, then a theory starting with conscious experiences might boot up spacetime, physics and chemistry. To do this, it would of course need a mathematical model of an observer. Idealist and spiritual traditions so far offer none.

Remarkably, neither do physicalist theories. Scientists are observers. But Newtonian physics assumes that observers don't affect what they observe and can be safely ignored. Einstein's relativity mentions observers, but they are identified with frames of reference, consisting of coordinates and clocks.

Quantum theory is forced to put observers front and center. When physical systems are not observed, they evolve according to the Schrödinger equation, which is deterministic and unitary. But experiments reveal that if a system is observed then its state collapses to a single value, a process that is random and non-unitary. So an unaltered Schrödinger evolution cannot create a theory of observation. Any theory it might propose would be unitary, and observation is not unitary. Replacing observers with physical "measuring devices" won't help. In quantum theory, physical devices have unitary evolution. But measurements are not unitary.

Decoherence doesn't help either. Decoherence can transform a pure quantum state with complex coefficients into a mixed state with real coefficients. But that is not enough. Measurement gives a single outcome, decoherence does not.

So quantum theory highlights the observer, but can't construct a theory of the observer. This is a central and open problem. As the physicist Frank Wilczek put it,"The relevant literature is famously contentious and obscure. I believe it will remain so until someone constructs, within the formalism of quantum mechanics, an "observer," that is, a model entity whose states correspond to a recognizable caricature of conscious awareness . . . That is a formidable project, extending well-beyond what is conventionally considered physics."

**The Trace Logic Of Conscious Observers**

Wilczek poses a worthy challenge: construct a formal model of an observer "whose states correspond to a recognizable caricature of conscious awareness." However, this challenge cannot be met within the formalism of quantum mechanics, which demands unitary evolution. But it might be met by a deeper formalism that does not start with unitary evolution, and allows unitary evolution to somehow arise. That formalism must show precisely, no hand waves, how spacetime and quantum theory arise from a dynamics beyond spacetime, a dynamics "whose states correspond to a recognizable caricature of conscious awareness."

Let's try. And let's start simple.

In fact, let's start where Wilczek suggests: with a conscious observer whose states are its experiences. A simple observer might have just three experiences, say red, green and blue. So it has three states. But, as we noted before, humans taste thousands of flavors, see millions of colors, and smell trillions of odors. So complex observers are possible, with trillions of states.

But why stop at trillions? There are many experiences no human has encountered. Some fish, sharks, rays, and chimaeras have electroreceptors, called ampullae of Lorenzini, which allow them to detect electric fields. Some insects and birds can see the polarization of light. The list goes on. The set of possible experiences is enormous, perhaps infinite.

Experiences change with each shift of attention and movement of eyes or body. Even if I sit still, my experience can change. I may wait, for instance, at a traffic signal, watching a red light. Then it turns green. The transitions of colors at a traffic signal are predictable: red turns to green to yellow and back to red in an endless cycle. Most transitions between experiences in daily life are less predictable. Indeed, we normally need a web of probabilities to describe how experience evolves.

We can write these probabilities in a simple array. For instance, at the traffic signal we see red, green, and yellow. If the light is red, we know the probability that the next color will be red, green, or yellow: 0, 1 and 0. If it is green, we know the probability that the next color will be red, green, or yellow: 0, 0 and 1. If it is yellow: 1, 0 and 0.

We can arrange these probabilities in a square array, or matrix, having three rows and three columns. The first row lists the transition probabilities if our current experience is red: 0, 1, 0. The second row lists the transition probabilities if our current experience is green: 0, 0, 1. The third row lists them if our current experience is yellow: 1, 0, 0.

Now imagine a defective traffic light with different probabilities. For instance, the first row might be .2, .3, .5, the second row .4, .5, .1, and the third row .3, .3, .4. This would be bad for traffic. But notice that each row sums to 1, so each row is a probability measure.

These square arrays, in which each row is a probability measure, are called Markov matrices, in honor of the mathematician Andrey Markov. Markov matrices describe probabilistic transitions that depend only on the current state. A dynamics whose transitions are described by a Markov matrix, such as the dynamics of a traffic light, is called a Markov chain.

Markov matrices will be our formal model of conscious observers. To model a decent range of human experiences we need matrices with trillions of states, not just three. But we can imagine that there are conscious observers corresponding to each matrix of any size, from 1 to infinity.

This is a bold proposal: beyond spacetime there are countless conscious observers. They can be modeled with Markov matrices. They are not products of physical processes in spacetime. Instead, they create spacetime and all physical processes within spacetime. Spacetime is just one of innumerable headsets that conscious observers construct to navigate their interactions with other conscious observers.

Can we build spacetime from Markov chains? That seems a tall order. These chains have been around since Markov's work in 1906. If they could build spacetime, that would surely have happened already. So what is new here?

What's new is a "trace logic" that unites all Markov chains. It is the key to booting up spacetime, whether flat or curved, from the dynamics of conscious observers.

To explore the trace logic, we must first understand the trace of a Markov matrix. Unfortunately, there are two distinct meanings of trace, so we must take care to avoid confusion. It can refer to summing the diagonal elements of a matrix. That's not the meaning here.

To understand the meaning here, consider again the Markov matrix for the defective traffic light. The first row is .2, .3, .5, the second row .4, .5, .1, and the third row .3, .3, .4. Suppose you watch that defective light. But suppose you can't see green; you can only see red and yellow. So as you watch, you see red transition to red or to yellow. You see yellow transition to red or yellow. But you see no transitions involving green.

If you watch awhile, you'll see a pattern of transitions between red and yellow: It turns out that the chance is 44 percent that red stays red, and 56 percent that it turns yellow; the chance is 54 percent that yellow stays yellow, and 46 percent that it turns red. This describes a Markov matrix. The first row is .44, .56; the second is .54, .46. This red-yellow matrix is what we are calling a <u>trace</u> of the red-green-yellow matrix.

Here is the key point. If I start with a Markov dynamics, but ignore some states, then the new Markov dynamics I see is the trace. So the trace relates the dynamics seen by different conscious observers. Indeed, it induces a logic on all conscious observers. Here I assume that some observers have experiences that are subsets of the experiences of other observers.

This may sound strange. What could it mean to have a logic on all conscious observers? After all, logic is about relationships of entailment, conjunction, disjunction and negation among propositions, among statements that can be true or false. For instance, the proposition *Chris is a bachelor* entails the proposition *Chris is a man*; to be a bachelor you must be a man. From the proposition *Chris is a man* and the proposition *Chris is unmarried* we can construct a conjunction: the proposition *Chris is a man and Chris is unmarried*. We can also construct the disjunction: *Chris is a man or Chris is unmarried*. From *Chris is a man* we can construct the negation *Chris is not a man.* So, in everyday usage, a logic refers to propositions related by entailment, conjunction, disjunction and negation. How does this translate to conscious observers and their matrices?

It's easy to see that each matrix indeed represents a proposition. For instance, to describe the matrix for transitions of red and yellow lights, I could say "If I see red now, then there's a 44 percent chance I'll see red next, and a 56 percent chance that I'll see yellow next." That is a proposition. Indeed it is a conjunction of two propositions. It can be true or false.

Okay, but these are complex propositions. A matrix on a thousand states is a proposition. But this proposition is the conjunction of literally a million smaller propositions. How can we possibly work out notions of entailment, conjunction, disjunction and negation? Surely we'll get lost in the woods of complexity.

Fortunately there's a way to think about logics that eases their generalization to Markov chains. The key is to understand entailment.

First note that, trivially, *Chris is a man* entails *Chris is a man*. No surprise. Each proposition entails itself. We say that entailment is "reflexive."

Second, suppose I have two propositions, and each entails the other. Then, as far as entailment is concerned, they are equivalent. We say that entailment is "antisymmetric."

Finally, note as before that *Chris is a bachelor* entails that *Chris is a man.* Now also note that *Chris is a man* entails that *Chris is a mammal*. From these two entailments we can conclude that *Chris is a bachelor* entails that *Chris is a mammal*. We say that entailment is "transitive."

So entailment is reflexive, antisymmetric, and transitive. If we find a relationship among Markov matrices with these properties, then we might find a complete logic as well.

I tipped my hand. There is a relationship: the trace. That it's reflexive and antisymmetric is fairly obvious. That it's transitive is not so obvious, but was proved by the mathematician Chetan Prakash.

So there is a logic on the set of all Markov matrices, and thus on the set of all conscious observers. It's a complex logic that mathematicians call a "partial order." For partial orders the convention is to say "meet" instead of conjunction and "join" instead of disjunction. I will observe this convention.

The trace logic has no maximal matrix. This means there is no special matrix whose traces create all possible matrices. Instead there an infinite number of directions in which we can enlarge the state space for Markov matrices to approach an infinite number of states, and for each set of states there are infinitely many distinct Markov matrices.

This means that no single matrix corresponds to an omniscient observer. It also means that the meet and join of pairs of conscious observers does not, in general, exist. Nor does the negation of a single conscious observer.

These are remarkable conclusions to draw about conscious observers. Each follows from the mathematics of partial orders applied to the trace logic on Markov matrices. Mathematics can reveal what intuition might not discern.

The trace logic has another striking feature. Pick any matrix, and consider its traces. Together they form a simple and well-behaved logic, called a Boolean logic. Meet, join and negation are all well-defined. So the trace logic is not Boolean, but each matrix has a Boolean sub-logic beneath it. Simple assumptions about the nature of observation entail a remarkably rich logic of observation.

The trace logic offers new insights on old problems. For instance, an open problem in the science of consciousness is the "combination problem": How can different phenomenal states combine? This problem arises, for instance, in panpsychist theories claiming that elementary particles of physics, such as photons and electrons, have conscious experiences.

One aspect of the combination problem is this question: When and how can different subjects of experience combine? William James famously raised it in The Principles of Psychology (1895). "Take a sentence of a dozen words, and take twelve men and tell to each one word. Then stand the men in a row or jam them in a bunch, and let each think of his word as intently as he will; nowhere will there be a consciousness of the whole sentence. We talk of the 'spirit of the age,' and the 'sentiment of the people,' and in various ways we hypostatize 'public opinion.' But we know this to be symbolic speech, and never dream that the spirit, opinion, sentiment, etc., constitute a consciousness other than, and additional to, that of the several individuals whom the words 'age,' 'people,' or 'public' denote. The private minds do not agglomerate into a higher compound mind."

Contemplation of this combination problem led to the discovery of the trace logic. It offers a precise solution: Two conscious observers can combine if their join in the trace logic exists. Their combination is precisely their join.

**The Trace Logic Of Intelligence**

The trace logic on conscious observers concerns conscious experiences. Conscious observers can be generalized to conscious agents that have conscious experiences and a repertoire of actions. The trace logic extends to conscious agents, and formalizes an important aspect of intelligence.

That aspect is surprise. Suppose that each time I act, I'm surprised by what happens. I reach for a glass, and it crashes on the floor. I take a step and fall on my face. I put on a shirt and rip a hole in it. Well then, I'm not too bright. The more I'm surprised by the outcomes of my actions the less intelligent I am. There is more to intelligence than not being surprised. But there is not less.

As it happens, trace minimizes surprise. Given an environment governed by some Markov chain, the trace of that chain on a subset of states gives a new chain that minimizes surprise for that subset and environment. We can see this in two ways.

The first is easy. The trace gives you transition probabilities that are empirically correct: they describe exactly the pattern of transitions you see long term. No matrix does better.

The second is simple and fascinating, but a tad longer to explain. Every finite Markov matrix has a "stationary measure," which gives the long-term probabilities of being in each state. For instance, it turns out that the defective traffic light spends about 30.682 percent of its time red, 37.5 percent green, and 31.818 percent yellow. Its trace onto red and yellow spends 49.091 percent of its time red and 50.909 percent yellow.

Here's the remarkable thing. Let's look at the ratio of red time to yellow time. For the traffic light it's 30.682/31.818, which is .964. For the trace it's 49.091/50.909, which is .964. They're identical. The trace preserves the relative times for red and yellow. This is true in general, no matter how many states are in the trace. This means that the stationary measure for a trace is identical to the stationary measure of the original on those states, up to a scale factor. We say that the stationary measure of the trace is a "normalized restriction" of the stationary measure of the original. This is a second way that the trace minimizes surprise: It predicts precisely the same pattern of long-term behavior as the original.

Another perspective on this is worth noting. There is, as we've seen, a logic on all Markov matrices: the trace logic. But there's also a logic on all probability measures: the Lebesgue logic. In this logic, one measure entails another if it's a normalized restriction. Like the trace logic, the Lebesgue logic is non-Boolean, but has Boolean sublogics associated with each probability measure. The map from a matrix to its stationary measure is a so-called homomorphism from the trace logic of observers to the Lebesgue logic of probabilistic beliefs. This just means that the logic of observation meshes properly with the logic of belief.

The trace logic, by minimizing surprise, is a logic of intelligence, of sentient behavior. Minimizing surprise is central to active inference theories of sentient behavior, and is a promising approach for building artificial intelligence.

**Building Spacetime Physics From The Trace Logic**

Let's now sketch how spacetime may emerge from the trace logic. For this we need one enhancement. To each matrix we attach a counter that increments with each update of state. That's all.

To motivate his first theory of space and time, the special theory of relativity, Einstein proposed two assumptions: (1) the laws of physics are identical for all observers in uniform motion and (2) the speed of light is the same for all such observers. He showed that these assumptions lead to counterintuitive predictions about what observers see.

First, if you and I each carry a clock, and you move past me at constant velocity, I see your clock ticking more slowly than mine. Second, if you and I each carry a meter stick, and you move past me at constant velocity in the direction your stick points, then I see your stick as shorter than mine.

Why didn't we notice this before? Because normally your velocity relative to me is small, nowhere near the speed of light. So I see your clock and meter stick as nearly identical to mine. But there are fast-moving objects in the universe, and when we check we find that these counterintuitive predictions are correct. This is the way spacetime works.

So already we have an important hurdle for the theory of conscious observers. If, as we claim, conscious observers are more fundamental than spacetime, and somehow create spacetime, then we must show how this odd behavior of clocks and rulers naturally arises from conscious observers and their trace logic.

We can, and it's surprisingly simple. Let's start with clocks. Recall the traffic-light example. Suppose Fran sees red, green, and yellow, and Sid just sees the trace on red and yellow. Fran's counter increments each time red, green, or yellow appears. Sid's counter increments only if red or yellow appears. So Fran's counter runs faster than Sid's. That's the key insight.

But we need one more step. In Einstein's spacetime, if Fran and Sid are in relative motion, then Fran sees Sid's clock advance more slowly. But, equally, Sid sees Fran's clock advance more slowly. Our first simple example doesn't capture this symmetry.

No problem. This is easy to fix and leads to an insight. Consider a traffic light with four colors, red, green, blue, and yellow. Fran sees just the trace on red, green, and yellow, while Sid sees just the trace on red, yellow, and blue. Fran cannot see blue, and Sid cannot see green. So Fran cannot see Sid completely and Sid cannot see Fran completely. All Fran can see of Sid is Fran's trace of Sid, and all Sid can see of Fran is Sid's trace of Fran. So Fran sees Sid's counter run slowly, and Sid sees Fran's counter run slowly. We get the symmetry of Einstein's spacetime.

We must take the limitations of observers seriously. I do not see unadorned truth. I see just a trace. From my trace your counter looks slow. But from your trace my counter looks slow. Who is right? You're never absolutely right, because you only see a trace. Who is wrong? You're never wrong, so long as you recognize that you only see a trace. Spacetime is not the structure of observer-independent reality. It is the structure of observer-dependent traces. It is a VR headset.

We have glanced at clocks, now let's look at rulers. The same simple ideas apply, but with an interesting twist. Rulers measure distance. We need a notion of distance for conscious observers. So we must define a natural notion of distance between states of a Markov matrix.

This was done in 2017 by mathematicians Peter Doyle and Jean Steiner, using the commute time between states. The commute time between states *a* and *b* of a Markov chain is the expected number of steps to go from *a* to *b* and back. For example, with the normal traffic light, in which red always goes to green, green to yellow, and yellow to red, it's easy to see that the commute time between any two colors is 3. But for our example of a defective light, computing the commute time is more complex, because there are countless paths to consider. A key intuition is that larger commute times between states means it's harder to travel between states, so they're farther apart. Doyle and Steiner discovered that the square root of the commute time between *a* and *b* has the properties of Euclidean distance in *n*-dimensional space. So distances in this space are the *square root* of the commute times between states of conscious observers beyond spacetime.

This means we must revisit counters and clocks. We saw that the counters of conscious observers beyond spacetime slow down in a trace. We proposed that this explains why clocks slow down in spacetime. But exactly how are clocks and counters related? Given the result of Doyle and Steiner, we propose that clock time in spacetime is the square root of counter time for conscious observers beyond spacetime.

Can this proposal work? In Einstein's special theory of relativity, the relative distortions of time and length are given by identical expressions. In the trace logic, trace-induced distortions of counter time and commute time are not generally identical.

But they are identical for one class of matrices. Recall the normal traffic light, where red goes only to green, green to yellow, and yellow to red. Its matrix is called a 3-cycle, because it cycles endlessly in a fixed sequence through three states. A cycle on *n* states is called an *n*-cycle.

Our proposal works for *n*-cycles. Projecting the square root of counter time to clock time, and the square root of commute time to distance, can precisely match the distortions predicted by special relativity.

Should we worry that it works just for *n*-cycles? No. To the contrary, Einstein used just light, and no other particles, to build special relativity. We use just *n*-cycles, and no other matrices, to link trace logic and special relativity.

Then are *n*-cycles the trace-logic analog of photons? Photons are massless and always travel at the speed of light, the fastest possible speed. No massive particle moves at the speed of light. These are precise properties. Can *n*-cycles project to these properties of photons?

They can. The average commute time between states is fastest for *n*-cycles. Any other matrix has a slower average commute time. In this sense, *n*-cycles are the fastest matrices.

What about mass? Photons are massless. In what way are *n*-cycles massless? What is the natural analog of mass in matrices?

Intuitively, mass is related to influence and connectedness. A massive object strongly affects other objects. A light object has less influence. An analog of this for matrices is the interconnection between states. If a state connects directly to many others, then it's quite influential. If it connects to just one, it's less influential.

So an analog of mass for a Markov matrix is some measure of the average connectedness of its states. A natural candidate is the entropy rate, a central concept in information theory. A matrix with higher entropy rate has more connectedness of states and more complex behavior. So let's propose that entropy rate projects to mass in spacetime. Conscious observers whose matrices have higher entropy rate appear in spacetime as more massive.

Does this work for light? Yes. An *n*-cycle has zero entropy rate, so its projected mass is zero. In fact, any Markov matrix with a 1 somewhere in each row has zero entropy rate. Any other matrix has an entropy rate greater than zero, so its projected mass is nonzero. It also has slower average commute times than *n*-cycles. So I propose that *n*-cycles of enhanced Markov chains project to light, and that they can be used to construct spacetime. Specifically, as *n* grows to infinity they converge to flat spacetime.

Let's summarize. Why does nothing travel faster than light? Because no matrix proceeds through states faster than an *n*-cycle. Why is there no preferred inertial frame? Because there is no greatest matrix in the trace logic. Why do massive particles never travel at the speed of light? Because they are not projections of *n*-cycles. The theory of conscious observers explains what special relativity assumes.

If we just need *n*-cycles to construct flat spacetime, what about the other matrices? Most are not *n*-cycles. Should we ignore them?

No. I propose that some matrices, with unequal counter and commute times, generate a spacetime that converges to curved spacetime. I also propose that most matrices don't generate spacetimes that converge to any conventional spacetime structures, flat or curved. These matrices are outside the purview of our spacetime headset. No surprise. Headsets are convenient fictions, not windows on reality. It will be fascinating to investigate the infinite realm beyond our headset.

## Quantum Theory From The Trace Logic

We have focused so far on spacetime. But what about quantum theory? What is its relationship to conscious observers?

It describes their long-term dynamics. We have focused on the step-by-step dynamics of conscious observers, which is like studying each frame of a movie. But we can step back and see their long-term dynamics, which is like watching the movie at normal speed. For this, we study the "harmonic functions" of enhanced chains. Chetan Prakash showed that some harmonic functions are identical to quantum wave functions of free particles.

A Markov matrix has real probabilities. But its long-term behavior is described by the complex wavefunctions of quantum free particles. So I propose that quantum wavefunctions describe the long-term dynamics of conscious observers.

If this is so, quantum theory is not a fundamental theory of reality. It just sees the asymptotic behavior of a deeper reality: a Markov dynamics of conscious observers. This deeper level has variables that are hidden from quantum theory.

Hidden variables? Didn't John Bell prove a theorem that rules out theories with hidden variables? And haven't experiments confirmed that hidden variables cannot exist?

Not at all. Bell proved that any theory with hidden variables must be nonlocal, or measurement dependent, or both. Conscious observers are both, and more.

The theory of conscious observers uses Markov chains and their states. These are prior to spacetime and locality, and in some cases may be used to construct spacetime. Thus they are nonlocal. Moreover, the trace logic is a logic of measurement. Each observer is a different measurement context. So it is measurement dependent. And finally, Bell's theorem targets hidden variables that are deterministic. Conscious observers are probabilistic. So Bell's theorem does not preclude the hidden-variables theory of conscious observers.

It's no surprise that conscious observers are nonlocal. Locality is just a feature of our spacetime headset, not a deep insight into reality. It's also no surprise that the dynamics of conscious observers is measurement dependent. They are, after all, observers.

If, as we propose, quantum wavefunctions arise from the long-term dynamics of conscious observers, then that dynamics should explain the many weird properties of quantum theory. So there are many hard-nosed tests for this proposal.

One weirdness is Heisenberg's uncertainty principle. You can't know the exact position and momentum of a particle at once. The better you know position, the less you know momentum, and vice versa. Similarly for time and energy.

The trace logic explains this with "sampled" traces. Recall our defective traffic light, and the trace we got by attending to red and yellow. I noted that if you watch long enough, you see a pattern of transitions: the chance is 44 percent that red stays red, 56 percent that it turns yellow, and so on. Now watching is sampling. You can take small samples or big. If you want to know the current state, you want a small sample: just one state. To know the entropy rate, you need a large sample. You can't sample small and large at once. This limitation underlies the uncertainty principle.

## Transcending The Headset

Galileo argued that the book of nature "is written in the language of mathematics, and its characters are triangles, circles, and other geometric figures." Mathematics permits a precision of prediction that invites precise experimental tests. This fuels a virtuous circle of theory and experiment, which in the last four centuries has transformed our worldview.

Science repeatedly confirms that spacetime and its contents can be accurately described by mathematics. This seems almost magical. As the physicist Eugene Wigner remarked, "The miracle of the appropriateness of the language of mathematics for the formulation of the laws of physics is a wonderful gift which we neither understand nor deserve."

Paradoxically, the magic of mathematics has served, in the eyes of many, to demystify nature. For them, nature is a machine and, as the philosopher Gilbert Ryle put it, there's no reason to believe that there is a "ghost in the machine." They may remind us that the church, which forced Galileo to recant and finish his days in house arrest, claimed that its authority was backed by miracles; but today, the weight of miracles favors science, whose theories of the machine routinely spin off technologies that, to Galileo's peers, would be indistinguishable from magic.

So it's no surprise that almost all scientific theories of consciousness start with a neural or computational machine, and expect consciousness, or the illusion of consciousness, to somehow emerge from unconscious ingredients. They argue that the mechanistic approach has time and again proved remarkably adept and powerful. True, it hasn't cracked consciousness yet. But be patient. With the discovery of DNA it demystified the problem of life and it will, in due course, unravel the problem of consciousness.

However, as I mentioned earlier, high-energy theoretical physics has found a canary in the coal mine: the Planck scale. Spacetime is the foundation of nature's machine. But that foundation is riddled with holes at the Planck scale. Our best theories of spacetime entail, without equivocation, that spacetime cannot be fundamental. As Nima Arkani-Hamed and other physicists put it: "Spacetime is doomed." If spacetime is doomed, so are its particles and other physical objects.

Also in jeopardy are theories of consciousness that start with causal or functional properties of physical systems such as brains or computers. Physics tells us that elementary particles are symmetries of spacetime; they are "irreducible representations of the Poincare group." If spacetime is not fundamental then neither are its particles nor its objects, such as brains and computers, which are composed of particles. If a theory of consciousness nevertheless asserts that such objects have causal or functional properties, then it must justify that assertion.

Correlation is inadequate justification. Specific neural activity may indeed systematically precede specific conscious experiences. Crowd gatherings may indeed systematically precede train arrivals. In neither case does correlation imply causation.

Spacetime is doomed. It is not the fundamental fabric of reality. It's the format of a headset. Science cut its teeth studying our headset. It sharpened an impressive array of theoretical and experimental tools, which appeared to limn a theory of everything. Instead it limned the inside of a headset, one of countless possible headsets.

This discovery is a breakthrough for science. We can begin to explore outside the tiny bubble of our headset. How vast is the realm outside? Infinite. How many new concepts will be needed? Infinity. But our hard-earned tools are ready to launch the adventure and evolve in the process.

Theoretical physicists and mathematicians have recently started to peek beyond spacetime. What they find so far are positive geometries, whose structures encode likelihoods of particle interactions. They are static objects outside spacetime that somehow describe dynamical systems inside spacetime. Like the obelisk in 2001: A Space Odyssey, these positive geometries are pregnant with meaning. But like Kubrick's clueless primates, howling and pounding on the silent monolith, we are for now confounded about their origin.

But no worries. These are first steps of science beyond the headset. We should expect a labyrinth of surprises. The trace logic and theory of conscious agents may, at best, be useful tools to begin our exploration. They will surely be replaced in due time, as will each successive theory. There is infinite job security in science.

Planck's scale spells doom for physicalist theories of consciousness. So the science of consciousness is destined to progress beyond physicalism. Unfortunately, Plancks' principle explains how it's likely to progress: one funeral at a time. It's hard work to find plausible assumptions, explore their implications, and test their predictions with well-designed experiments. This requires heavy personal investment by scientists. It's understandable that decades of investment in a theory might leave one attached to its foundational assumptions. So it's often a new generation of scientists that picks up where the last left off, and entertains new foundations.

As we've discussed, there is in principle no end to the search for deeper assumptions. In practice, however, our species may at some point simply fail to have the concepts needed to continue this search. We presume that there are hard limits to the conceptual repertoires of ants, ravens and macaques. That may also be in the cards for homo sapiens.

But for our next step, I propose the assumption of conscious agents with their trace and Lebesgue logics as a framework for theories of consciousness. For this framework to be taken seriously, it must pass hard-nosed tests. I propose six in the epilogue.

**Implications Of Transcending The Headset**

If spacetime is a headset, and reality beyond can be described by conscious agents and the trace logic, then there are fascinating social, personal, scientific and philosophical implications.

First, what is life? This question may seem trivial. Point to any object, and even a child can say if it's alive. Rock? No. Ant. Sure.

Life is easy to see, but remarkably difficult to define. We might suggest that something is alive if it can eat, excrete, metabolize, grow, move, breathe, react to stimuli, and so on. Or we might define that "life is a self-sustaining chemical system capable of Darwinian evolution" as proposed by some exobiologists. But for any list or definition we make, we find exceptions and counterexamples. Some are canvassed by Sara Imari Walker in her book Life as no one knows it.

Sara presents an "assembly theory" of life, developed with Lee Cronin and others. According to assembly theory "Life is the only thing in the universe that can make objects that are composed of many unique, recursively constructed parts." Some objects are easy to assemble and others more difficult, requiring more steps. The assembly index of an object is the smallest number of steps needed to assemble it. The larger the index, the less likely the assembly is an accident. If you find multiple copies of a high-index object, this indicates the action of sophisticated information processes unique to life. All objects with high index and multiple copies constitute "Life". Sara suggests that assembly theory expands our conception of life and thus enhances our chance to detect it in forms alien to us.

Near the end of her book, she observes that “What we are after in this book—to understand what we are—may already be right in front of us. We cannot see ourselves clearly because we have not built a theory of physics yet that treats observers as inside the universe they are describing…”.

I agree. Physics needs a theory that treats observers as inside the universe they are describing. That is precisely what the trace logic delivers. Each observer is a trace of a larger dynamics. Being a trace, each observer is necessarily inside the universe of dynamics that it describes. The observer is not objective. It participates in the creation of the very dynamics it observes.

This creative participation is the key point of John Wheeler’s famous it-from-bit proposal: “Directly opposite to the concept of universe as machine built on law is the vision of a world self-synthesized. On this view, the notes struck out on a piano by the observer-participants of all places and all times, bits though they are, in and by themselves constitute the great wide world of space and time and things.” Wheeler was prescient. We propose that observers “by themselves constitute the great wide world of space and time and things” via the trace logic, which induces the time dilations and length contractions at the foundation of relativistic spacetime.

This radically reframes the question of life. The universe is self-synthesized by observer-participants. It is itself alive, not a “machine built on law” that sometimes spawns life. Each observer-participant views that self-synthesizing life through a trace—an interface or headset that filters out the vast majority of that life.

But if life, in its infinite variety of interacting forms, is the very fabric of reality, then why have we seen no evidence of alien life? Why the Fermi paradox? As Enrico Fermi famously asked, “So, where is everybody?”

They’re all around us. Our headsets screen out most of life, and desiccate much of what remains into inanimate matter. The life we see in our headset is a thimbleful drawn from an infinite ocean.

So our distinction between living and nonliving is an artifact of the limitations of our headset. That’s why we can’t articulate a principled distinction. An analogy may help. Suppose I talk with a friend via Zoom. I see a desk, chair, plant, walls, and my friend’s face. I start to wonder what’s special about the living pixels of the plant and face. How are they different from the inanimate pixels of the desk, chair, and walls? Well, I’ve just set myself a quixotic task. There is no principled distinction between living and nonliving pixels. Pixels are just pixels. They are the format of a computer interface.

There is a deep unity in diversity of the trace logic. Diversity has been our focus: the infinite variety of matrices and traces. But unity appears as follows. To write a matrix I must first specify the set of all its states, with a certain structure called its measurable space. Suppose a matrix has three states. Then its measurable space has three states. But a matrix in the trace logic cannot be considered in isolation. It is a trace of many bigger matrices, such as one with a trillion states. So its measurable space can't have just three states. It needs at least a trillion. But this bigger matrix is itself a trace of even bigger matrices, ad infinitum. So all matrices in the trace logic have this in common: Each is founded on the same infinite measurable space.

What is the meaning of this infinite shared space? I don't know. Now we probe the limits of the theory of conscious agents. But here are some thoughts, which may lead to a deeper theory.

All experience springs from this infinite shared space. But it itself is not an experience. We could call it the Source of experience. This Source is timeless, and prior to physical notions of space and time. It is the unity behind the diversity of life and experiences.

To better describe the Source, let's again discuss states. They seem to have a clear meaning. For our traffic light, the three states are color experiences: red, green and blue. No problem.

Well, there is a small problem. Is my experience of red the same as yours? What if you're red-green color blind and I'm not? Then the experiences we call red are different. Indeed, for all I know, all of my experiences differ from all of yours.

This problem arises because I can't experience your experiences. I can only feel your pain metaphorically. I may note parallels in how we act and talk, and conclude that our experiences are similar. But this conclusion is fallible.

We learn to talk about our experiences by a process of ostensive definition. Suppose, for instance, that you've never tasted rhubarb. So I give you some, and you take a bite. Then I say, "*That's* the taste of rhubarb." You now know, by ostensive definition, how to talk about the taste of rhubarb. You had an experience. I just told you what to call it. But I still don't know if your rhubarb experience is like mine.

I can't do this with Source. I can't give some Source, have you taste it for yourself, and then say, "*That's* the experience of Source." The Source of experience is not itself an experience. But the Source is not separate from you. It is the infinite shared space that is you at the deepest level, prior to any experience.

There is no final theory of everything in science and, a fortiori, no final theory of you. You transcend all theories, stories and experiences. But since you are you, you can be aware of yourself directly. Ask yourself, for instance, "I wonder what my next thought will be?" Then wait with attention until you have a thought.

If seconds passed before you had a thought, then for those seconds you were thoughtless awareness. *That* is you. That is an ostensive definition of you, of Source. The exploration of you, of awareness beyond thought, of Source, is endless.

If you are Source, so is your neighbor. Source peers at Source through a trace, a headset. Looking past your headset view of your neighbor, with its race and gender, you find yourself. Your neighbor is yourself. Looking past genus and species, animate and inanimate, you find yourself. Separation is an illusion of a trace. Source views itself through countless headsets, and wakes to its nature beyond headsets, thoughts and theory.

This waking can be gradual: expanding a trace to more states. In this case, the headset gets an upgrade, but is not removed.

Expanding a trace means joining with another trace. To consummate this join in the trace logic, the matrix of each trace must change in the process. Old transition probabilities must be replaced with new. Death of the familiar is integral to birth of the fresh. The trace logic minimizes surprise. So this death, though unavoidable, is as gentle as possible. And the fresh, though alien, is as familiar as possible. Surprises that are accepted forge the new join.

If no theory is ultimate truth, are all theories equal? Is earth, air, fire, and water as good as the Standard Model of particle physics? No. Theories are not reality, but they can be more or less useful descriptions of a perspective on reality, of a trace of Source. Experiments can reveal which theories better track the dynamics of our spacetime trace. These are the better descriptions of our perspective.

But science can do more. It can create theories beyond the spacetime headset. Positive geometries in high-energy physics are one example. The trace logic on conscious agents is another. As such theories mature, they will open new technologies that are indistinguishable from magic.

An analogy can help illustrate the possibilities. Suppose you're a wizard at Grand Theft Auto. You know all the tricks and deploy them expertly. Impressive. But now suppose you're the geek who wrote the code for GTA. You're not stuck in its rules. You made them. You can exploit them and change them. You can give the wizard a flat tire, an empty gas tank, a sudden change of location. From inside the game you look like a magician, because you're outside.

Science has explored inside our spacetime headset for centuries, and discovered its limits. Now it ventures outside to realms beyond. This is a watershed for science. Spiritual traditions have long beckoned from beyond that watershed. But science brings powerful new tools to the exploration. A promising rapprochement drops the baggage of each and synergistically combines the best tools of both.

What a delicious croissant! And what an extended reverie. Now for a second bite.

**Epilogue: Conjectures**

If spacetime physics arises from the trace logic of conscious observers, then certain mathematical conjectures must be true. I briefly state some conjectures. Proofs or disproofs of these conjectures are critical tests of the ideas sketched in this essay.

**1**. The trace logic on enhanced $n$-cycles converges to continuum Minkowski space as $n \to \infty$.

Motivation: In our interpretation, $n$-cycles represent zero-mass particles such as light. The constant speed of light in all inertial frames is the foundation of Minkowski space. So we expect to see $n$-cycles converge to Minkowski space in some rigorous sense of convergence, such as Lorentz Gromov-Hausdorff. There may be connections between the partial order induced by traces and partial orders in causal set theory.

**2**. A small class of enhanced Markov chains has a discrete dynamics that converges in the continuum limit to the dynamics of general relativity.

Motivation: Cyclic chains behave like flat spacetime: Traces of cyclic chains transform commute times and counters identically; boosts in special relativity transform space and time identically. But traces of noncyclic chains transform commute times and counters unequally. These unequal transformations may, for a small class of Markov chains, give rise to the curved spacetimes of general relativity, possibly under Lorentzian Gromov-Hausdorff scaling. But we expect that curved spacetimes cannot arise from most Markov chains, such as chains with expander graphs, nearly disconnected components, or Finsler-like limits.

**3**. The Born rule arises from the trace logic applied to a class of enhanced Markov chains.

Motivation: We propose that quantum wavefunctions of free particles describe the asymptotic behavior of enhanced Markov chains. Wavefunctions dictate probabilities of experimental outcomes through the Born rule. So we must show how the Born rule arises from enhanced Markov chains. This may involve Berezin integration, connecting the hiding of states by traces with complex amplitudes.

**4**. Analogs of inflation, Big Bang and an expansion of the universe driven by “dark energy” arise from trace-logic sampling of a small class of enhanced Markov chains.

Motivation: We propose that spacetime is a headset used by some conscious agents. So we must show how spacetime and its history arises from conscious agents and the trace logic. This will involve sampled trace chains. As the sample grows, the estimated Markov matrices evolve, as do the commute times between states. These evolving commute times correspond to evolving distances. We look for a class of enhanced Markov chains for which this evolution corresponds to the evolution of spacetime.

**5**. Elementary particles of the Standard Model of particle physics are projections of recurrent communicating classes (RCCs), or communities of interacting states, of certain enhanced Markov chains.

Motivation: Elementary particles are irreducible representations of the Poincare group of spacetime symmetries. If spacetime arises from conscious agents, then so must elementary particles. We propose that $n$-cycles, which are RCCs, project to photons. Other RCCs may project to other particles. But RCCs may also have internal community structure that projects to interacting bound or confined particles. Decorated permutations classify certain classes of particle interactions, and they also classify RCCs and community structures of Markov chains.

**6**. Scattering amplitudes arise from certain sampled enhanced Markov chains.

Motivation: We propose that a small class of Markov chains projects to spacetime and particles. So we must explain how such chains project to scattering processes and predict scattering amplitudes. Scattering processes evolve over time, suggesting that sampling of Markov chains is central to an explanation. Given a master Markov matrix, one can generate sampled sequences of states of varying lengths and likelihoods. For each sample one can estimate a Markov chain and interpret its RCCs using the particle assignments found in Conjecture 5. There may be connections between positive geometries and collections of estimated Markov chains weighted by likelihoods.

**7**. Planck-scale failures of spacetime arise from tracing on too many states within a small region of a master Markov matrix.

Motivation: Probing ever smaller regions of spacetime requires ever smaller wavelengths of radiation, which in due course pack enough energy in a small space to form a black hole. Higher energies correspond to traces on more states; packing more energy into a small space corresponds to tracing on more states within a small region of the master Markov matrix. At some point the trace sees the discrete nature of the master matrix and the construction of smooth spacetime fails.

**8**. Quantum entanglement arises when an RCC is traced onto two or more disjoint subsets of its states.

Motivation: If $M$ is an RCC, interactions among its states are made explicit by the transition probabilities of $M$. Tracing $M$ onto two disjoint subsets of states, $A$ and $B$, gives two RCCs, $M_A$ and $M_B$. These two matrices give no hint of interaction between them, because they are each an RCC. But the fact that they are traces over hidden states of $M$, entails that they have hidden interactions between them. These hidden interactions may be the source of quantum entanglement.