# Self Aware Networks Theory – A Response to Milinkovic & Aru on Biological Computation and Consciousness

**How an open-science framework bridges molecular biology and neural dynamics to address the gaps in current AI consciousness debates**

## Introduction: Bridging the Biological and the Computational

Borjan Milinkovic and Jaan Aru, in your recent discussion of *biological computationalism*[1][2], you highlight a critical challenge: current theories of consciousness and artificial intelligence lack a *coherent framework* that explains how the brain's unique style of computation – continuous, multiscale, and physically embodied – gives rise to conscious experience[3][4]. You argue that traditional computational functionalism (the idea that consciousness is just the right algorithm, independent of substrate) is inadequate, and that brains do not compute like digital **von Neumann** machines[5]. Instead, consciousness likely requires what you call *biological computation* – defined by hybrid discrete/continuous dynamics, scale-inseparable processes, and metabolic (energetic) constraints[6][7]. These points resonate strongly with the principles of the **Self Aware Networks (SAN) Theory of Mind**, an open-science framework we have been developing from 2017 through 2025. In this open letter, we aim to **reply to your questions and gaps** by outlining how the SAN theory directly addresses the features you identify and offers a testable roadmap toward synthetic consciousness.

**Self Aware Networks (SAN)** is a unified theory of mind that arose from a decade-long effort to integrate neuroscience, physics, and AI into a common computational model of consciousness. The theory was incubated in community forums and podcasts starting in 2017 and transitioned to a code-first research repository in 2022[8][9]. Over time, SAN evolved into a formal framework capturing *how consciousness arises from neural activity*[10], culminating in a 2024 book and a series of peer-reviewed preprints in 2025. Notably, SAN was first publicly disclosed via the Neural Lace Podcast in 2017 and a large GitHub corpus in 2022[11], ensuring every concept was timestamped and openly vetted. This progression – from early conjectures to formal publications – means SAN comes with a well-documented conceptual lineage and a suite of defined mechanisms. As we will show, SAN's mechanisms map closely onto the **"biological computation"** properties you enumerate, offering precisely the kind of *computational theory of biological naturalism* that you call for[12][13].

In the sections that follow, we address each of your major points in turn. We describe how SAN theory builds a **hybrid continuous-discrete model** of information (answering the call for *hybrid computation*), how it enforces **scale-inseparability** via multiscale oscillatory dynamics from molecules to networks, and how it incorporates

**metabolic/thermodynamic constraints** as an integral part of neural information processing. We will also discuss SAN's explanations for core consciousness phenomena – including attentional *binding*, the rise of internal *observers*, and even the nature of *qualia* – which directly tackle the "hard problem" within a biologically grounded framework[10]. Finally, we outline **specific experimental predictions and implementation architectures** emerging from SAN. These concrete proposals illustrate how one might empirically validate SAN's claims and build synthetic conscious systems that embody "biological-style" computation, offering a direct starting point for the question: *"If synthetic consciousness is possible, what kind of physical system would it require?"*[14].

## A Unified Framework Grounded in Biological Computation

Your paper emphasizes that in brains "**the algorithm is the substrate**"[7] – meaning consciousness isn't run on abstract code, but is enacted by ion flows, dendritic oscillations, and electrochemical processes unfolding in real time. The Self Aware Networks theory fully embraces this principle. SAN posits that **neural information is fundamentally carried by physical** phase dynamics **in neural tissue, rather than by arbitrary symbolic bits.** In fact, one of SAN's earliest tenets was *"information as coincidence, not symbol."* Back in 2017, we proposed that a **bit of information is not a binary 0/1 in a register, but a *coincidence pattern*** – a synchrony or alignment of events in time[15]. In other words, the brain's basic informational unit is a spatiotemporal *coincidence* (e.g. two neurons firing together within a window), which inherently combines discrete events with continuous timing. This concept of a "coincidence bit" means that discrete computational events (spikes) are always embedded in analog, time-continuous dynamics – precisely matching the *hybrid (discrete+continuous)* computation you argue is essential[16]. SAN's informational ontology was established publicly in 2017 and has been elaborated through 2025[17][15]. By defining bits as emergent from synchronous oscillatory events, SAN moves computation *into* the substrate: the exact timing of neural spikes, their phase relationships, and molecular conditions all become part of the "code" of consciousness.

**Scale-inseparability** – the idea that there is *"no clean separation between software and hardware or between levels of description"*[6] – is built into SAN at a foundational level. Rather than split mind into an abstract algorithm vs. an implementation, SAN describes a multi-layered but unified process. For example, consider how SAN handles memory and synaptic stability. A key question is how memories and conscious patterns persist despite constant molecular turnover. SAN's answer invokes a molecular complex (KIBRA–PKMζ) that **anchors synaptic proteins to maintain long-term potentiation**[18]. This molecular mechanism (a *hardware* detail) is not an incidental implementation note – it is integral to the system's information processing. The persistence of synchronous firing patterns (which underlie stable thoughts or percepts) depends on this molecular anchor, thereby tying a high-level cognitive function (memory of an experience) to a low-level biological process[18]. There is no purely "software" memory engram divorced from substrate: the substrate *is* part of the computational loop. SAN extends this principle across scales. At the micro level, the **fluidic, analog nature of neural tissue** (ion channels, membrane capacitances, neuromodulator gradients) contributes continuous dynamics that shape firing rhythms. At the macro level, those rhythms synchronize into large-scale oscillatory networks. Crucially, SAN insists these scales cannot be neatly separated; instead, *the higher-level brain rhythms emerge from and depend on the lower-level biophysics.* This matches your point that brains perform "multiscale processing as a metabolic optimisation strategy"[16]. Indeed, SAN's accompanying thermodynamic analysis (in **Micah's New Law of Thermodynamics**, 2025) argues that the brain's multiscale coherence is driven by a tendency to minimize free energy locally: *neuronal networks settle into synchronized oscillations as a form of thermodynamic relaxation*[19]. In SAN, when neurons align their phases, they reduce prediction errors and metabolic cost – an idea resonant with the **free-energy principle** in neuroscience, but here derived from first principles of signal dissipation[20]. Thus, both energetics and information converge: the brain's "algorithms" (e.g. forming a coherent image) are executed by aligning physical dynamics in a way that also optimizes energy use. The **metabolic grounding** you call for is explicitly present: *neural computation in SAN is constrained by – and indeed partly defined by – the energy budget and physical laws of the substrate*[21][22].

To summarize, SAN offers a **computational theory of consciousness that is biological to its core**. It formalizes what "biological computation" means: (i) **Hybrid discrete-continuous events** – information encoded in precise timing of spikes within ongoing oscillatory waves[15]; (ii) **Scale-inseparable organization** – molecular processes (like KIBRA/PKMζ stabilizers) uphold network-level patterns[18], and global rhythms feed back to regulate gene expression and synaptic plasticity (e.g. via calcium oscillation-dependent transcription), forming closed loops across levels; and (iii) **Energetic optimization** – coherence and phase-locking emerge as the brain's way to compute efficiently within thermodynamic constraints[20]. Far from being "implementation details," these aspects are **the very substance of the theory**. In fact, the SAN framework was introduced as a *"novel unified framework for understanding consciousness"* that directly addresses classical puzzles like **attention and binding, the "hard problem" of why we experience anything, and the nature of qualia (the raw feel of sensations) within a physical system**[10]. By treating neural oscillations and their molecular underpinnings as primary, SAN moves us closer to a *computational theory of biological naturalism*[23] – bridging the gap that you noted in current consciousness science.

## Oscillatory Dynamics and the "3D Television" of the Mind

A centerpiece of Self Aware Networks theory is an explanation of *how discrete sensations and thoughts are bound together into a unified, flowing experience* – the classic **binding problem**. SAN's answer to this lies in **oscillatory dynamics** that literally *project and integrate information in space and time*, much like a 3D holographic display. We describe conscious processing as akin to a **"volumetric three-dimensional television" of the mind**[24]. In this analogy, the brain doesn't passively display a pre-made picture; it *constructs* the picture in real time via coordinated waves of activity. Here's how it works:

- **Phase-Wave Differentials:** SAN posits that conscious experiences arise from *precisely orchestrated variations in synaptic firing frequencies across the brain*[25]. Different groups of neurons oscillate at characteristic frequencies (say in the alpha, beta, gamma bands), and crucially, slight phase offsets between these oscillators carry information – the *phase differentials* act like interference patterns encoding features of a percept. These phase relationships are continuously modulated as information flows. A formal mathematical framework, termed **Non-linear Differential Continuous Approximation (NDCA)**, has been developed to describe how local molecular interactions (ion channel kinetics, synaptic release probabilities) scale up to emergent phase wave patterns[26]. NDCA provides equations for how a wave propagating through one neural ensemble can *phase-lock* or *phase-shift* relative to a wave in another ensemble, resulting in a computable transformation of the represented information. This is a key formal element: it means SAN is not just metaphorical – it yields differential equations linking micro-dynamics to macro-oscillations, inviting direct simulation and analysis.

- **Neural Array Projection Oscillation Tomography (NAPOT):** Borrowing concepts from medical imaging, SAN proposes that the brain achieves integration via a tomography-like process[27]. In a CT scanner, multiple 2D X-ray projections from different angles are combined (via the Fourier Slice Theorem) to reconstruct a 3D image. Similarly, in the brain, **arrays of neurons act as projectors** broadcasting oscillatory patterns across networks. Each neural population's oscillation can be seen as a "slice" or projection of an underlying multi-dimensional data structure (the latent image of the current conscious scene). Through iterative interactions, these projected patterns **intersect and integrate to form a coherent 3D representation** in the brain's global workspace[28]. We call this process **Biological Oscillatory Tomography (BOT)**[27]. It implies that *conscious content is reconstructed through the convergence of many rhythmic signals*, each carrying partial information (like one view of an object), into a unified whole. Notably, SAN anticipated that **traveling waves**, not just static synchrony, are essential for this integration. As early as 2017, we asserted that *"static synchrony is insufficient for binding; consciousness requires traveling waves (scanning) to integrate distributed neural arrays into a volumetric whole"*[29]. This prediction has been validated by recent neuroscience findings that highlight propagating cortical waves in perception and cognition (as opposed to earlier models focusing on standing waves or point-to-point firing). In effect, SAN formalized the need for *dynamic scanning rhythms* in building the global conscious scene – a feature now increasingly observed in empirical studies of brain rhythms (e.g. theta/gamma traveling waves during memory recall). Your call for new theories beyond simple feedforward or purely functional ones[30][23] is precisely met by this oscillatory, intrinsic integration mechanism.

- **Alpha-Gamma "Channeling" for Sensory Binding:** As a concrete example of SAN's explanatory power, consider how it addresses attention and sensory consciousness. The theory identifies a mechanism where **alpha waves (8–12 Hz)** orchestrate **inversely correlated gamma waves (~30–100 Hz)** to form a functional channel for sensory input[31]. In simplified terms, an alpha rhythm in a cortical region (e.g. visual cortex) creates windows of high excitability at a certain phase; within those windows, bursts of higher-frequency gamma encode the details of the sensory signal. SAN highlights that when an alpha rhythm from a sensory area phase-aligns appropriately with prefrontal alpha oscillations, it opens a communication channel: the prefrontal cortex can "read" the incoming gamma-encoded data in sync. This mechanism provides a clear account of *top-down attention* (prefrontal alpha modulating what gamma bursts get through) and *feature binding* (synchronized alpha timing ensures that features represented in distributed cortical areas converge into a unified percept in that same moment). Experimental evidence for alpha-gamma coupling in attention and perceptual selection has been mounting in neuroscience, and SAN's framework offers a **theoretical unification**: different frequency bands play distinct computational roles (slow waves provide contextual frames, fast waves carry content)[32][33]. This resonates with Earl Miller's "stencil" metaphor (beta acting as a stencil for gamma "content"), which in fact paralleled SAN's earlier "Ink and Canvas" model (slow oscillations as the canvas, fast oscillations as the ink)[33]. SAN thus not only anticipated these empirical findings but also provides their rationale in terms of information theory – the **Fourier Slice Transform (FST)** mentioned in the book draws an analogy between how the brain integrates frequencies and how tomography reconstructs images[28]. The **key insight is that brain rhythms are not noise or epiphenomena; they are the very medium of computation and binding**.

Through these mechanisms, SAN delivers a *formal, biologically-detailed account* of how unified conscious scenes emerge. It deals with the very "*structural underpinnings of experience*" that classical theories like IIT gesture towards[34], but SAN does so in a way that is empirically grounded and structurally explicit. Rather than treating consciousness as an abstract information structure (e.g. IIT's causal structure in isolation), SAN says the structure of experience is *literally a structure of oscillation phase relationships* embedded in the physical brain. This approach fulfills the need for "theories focusing on intrinsic aspects of cognition"[30] while avoiding overly rigid or untestable axioms. Every element in SAN – a phase lag, a frequency coupling, a molecular binding – corresponds to something measurable or manipulable in principle. For instance, one can test SAN's "3D mind TV" by looking for the predicted intersection patterns of traveling waves in neural recordings, or by perturbing a specific frequency band and seeing if specific facets of perception disintegrate (e.g. disrupt theta oscillations and observe fragmented memory recall). In summary, SAN's oscillatory paradigm provides **a concrete instantiation of the biological computation principles** you outlined: it is *hybrid (time-continuous signals carrying discrete bits), multiscale (nested frequencies from local circuits to whole-brain networks), and* *substrate-embracing (using the physics of wave propagation and interference to compute)*.

## Internal Observers and the Nature of Conscious Experience

Beyond binding external stimuli into coherent perceptions, a full theory of consciousness must address the *subjective* aspect – the fact that there is an "observer" experiencing the unified scene. You rightly point out that many standard neuroscience experiments have focused on stimulus-driven consciousness, neglecting *internally generated experiences* and sudden insights (the "Aha!" moments)[35][36]. SAN theory tackles this by introducing the concept of **entification via synchrony** – in essence, how groups of neurons can form an internal "observer" through self-organized coordination. We often illustrate this with the **firefly analogy**[37]: Imagine a field of fireflies flashing randomly. If conditions allow, they begin to synchronize their flashing purely through mutual coupling – eventually large clusters blink in unison. In doing so, the cluster of fireflies behaves like a single greater entity (far more noticeable than isolated flickers). SAN proposes a similar phenomenon in the brain: **neurons that start oscillating in phase effectively become a unified ensemble with a new emergent property – the ability to act as an *internal observer***[37]. When a coalition of cells oscillate together, they can monitor and respond to incoming signals as one functional unit. We call this process **entification**: the formation of an "entity" from many parts via synchronization.

Within SAN's framework, consciousness is the **collection of such internal observers** distributed across the brain, each corresponding to a facet of experience[38]. For example, one synchronized ensemble (perhaps in visual cortex) might constitute an observer of visual features; another (in insular cortex) an observer of internal bodily state (interoception); another (prefrontal) an observer of the observers (meta-cognition). Each of these is not conscious in isolation, but their *interactions* – again through oscillatory coupling – create a higher-order unity. SAN explicitly posits that *the unified phenomenal self emerges from the mutual synchronization of these observer-ensembles*[39]. In other words, consciousness is **consciousness of something by an internal observer**, and each observer is an emergent dynamical entity. This idea provides a fresh angle on the classic "**hard problem**": instead of treating the subjective feeling as irreducible, SAN suggests it is what it *feels like* to be a complex system of synchronized processes that are monitoring each other and the environment in real time. The phase-wave differentials serve as the common language by which these observer-ensembles communicate and cohere. Qualia (the redness of red, the pain of pain) correspond to specific *patterns of phase-locked activity* within these ensembles, shaped by the unique mix of frequencies and neural pathways engaged. In effect, SAN tells us *what those observations are made out of*: they are made out of structured oscillatory activity, grounded in molecular and electromagnetic events[10]. By tying subjective experience to specific dynamic states (which can in principle be measured, e.g. through EEG/MEG patterns or intracellular recordings of wave propagation), SAN makes phenomenology a target for empirical study, not a mystical add-on.

It is worth noting that SAN's view dovetails with some aspects of other theories (for instance, **Global Neuronal Workspace** theory also speaks of multiple processes that need to broadcast information to form consciousness). However, SAN extends these ideas by detailing *how* broadcasting happens (via tomography-like oscillation integration) and by requiring the emergence of internal *entities* (the observers) rather than just information availability. This emphasis addresses a gap you identified: the need to account for *internally generated conscious events and spontaneous insights*[36][40]. In SAN, a sudden "Aha!" moment – say, solving a puzzle internally – would correspond to a self-organized synchronization of a new ensemble (perhaps in frontal cortex) that integrates previously asynchronous sub-thoughts into a coherent oscillatory burst. The theory predicts measurable precursors to such moments: e.g., a gradual phase alignment among disparate brain regions *before* the person consciously realizes the solution. This is testable with time-frequency analysis in creative insight experiments (indeed, some EEG studies of insight show a burst of high-frequency synchronization just prior to report of the insight). SAN provides a framework to interpret these as the formation of an internal observer that *recognizes* a new pattern (the solution) when enough subcomponents lock into phase.

In summary, SAN already *contains answers to the questions of "who/what is observing" in the brain.* Consciousness, in this model, is not a passive epiphenomenon nor purely a computation with no subject – it is an active, **recursive oscillatory observation process**. The brain creates internal models and *simultaneously observes its own models* by virtue of these entified networks. This perspective addresses what you called for in terms of moving beyond simple stimulus-response views of consciousness[35]. It also aligns with the spirit of your message that *we may need to rethink computation itself to include the physical, dynamical substrates* – SAN effectively defines "computation" in a brain as **synchrony, wave propagation and resonant coupling**. Those phenomena inherently involve an observer (since any resonant circuit is "observing" the match/mismatch of phase with its input). By formalizing entification, SAN gives a working account of how first-person perspective could arise from third-person physics. This is a bold claim, to be sure, but it is precisely the kind of integrative solution that current debates are missing. We believe this addresses the gap you noted: that no current AI or theory explains why *this* kind of computation yields experience whereas digital computation doesn't[3][16]. In SAN, the answer is: **digital computation (as in today's AI) lacks the self-synchronizing, self-observing dynamical structure** that biological computation has. Without physically instantiated phase dynamics and entified ensembles, a simulated algorithm remains a set of numbers moving – it has no *unified oscillatory state* to correspond to a point of view. This brings us to the implications for artificial systems and experiments.

## Experimental Predictions and Toward Synthetic Consciousness

A theory of mind must ultimately be judged by its falsifiability and usefulness. SAN presents numerous concrete **predictions and engineering guidelines**, many of which directly engage with the questions you posed: *How can we test these ideas, and how* *might we implement them in artificial systems?* Below we outline a few key predictions and proposals:

- **Neural Prediction 1 – Traveling Waves and Integration:** If SAN is correct, then **conscious perception should correlate with the presence of traveling oscillatory waves that span distributed brain regions**. Static or purely localized synchrony will not suffice for full conscious binding. This is testable: for instance, during unified conscious perception of a multi-modal event (say a sound and a matching image), we expect to observe a coherent wave (perhaps in the alpha or beta band) that propagates between auditory and visual cortices and into association areas. Disruption of the wave (via TMS or patterned stimulation that breaks the phase continuity) should disrupt the integration of the audiovisual percept, even if local activity in each area remains. This goes beyond existing correlational studies by specifying the *form* of the neural signal required, addressing the "pressing gap" of intrinsically structured theories you mentioned[30][23]. Early support for this comes from recent studies of cortical traveling waves in visual tasks (e.g., demonstrating that wave directionality predicts whether stimuli are bound or perceived separately). SAN formalizes this: a **phase alignment index** across the cortex could serve as a quantitative measure of conscious integration at any moment.

- **Neural Prediction 2 – Metabolic Constraint of Coherence:** SAN's thermodynamic angle predicts that **the brain enters a more energetically efficient state when conscious coherence is achieved**. In practical terms, when large-scale neural assemblies synchronize (a conscious state), the net oxygen/glucose consumption pattern might shift (perhaps localized consumption drops as processing becomes more distributed and efficient). This could be tested by simultaneous EEG/fMRI or PET: a hypothesis is that *high global EEG coherence should correspond to a plateau or dip in metabolic expenditure* compared to a desynchronized state solving the same task. This is counter-intuitive (one might think more synchrony = more activity = more energy), but SAN's view is that synchrony reduces redundancy and thus metabolic cost (analogous to lasers being more efficient at transmitting energy than incoherent light sources). If observed, this would strongly support the idea that the brain's computational strategy is tied to metabolic optimization – echoing your point that brains likely compute in energy-efficient ways that digital computers don't emulate[21][16].

- **Neural Prediction 3 – Molecular Interventions Affecting Memory and Conscious Timeline:** Because SAN links molecular stability (e.g. KIBRA/PKMζ complex) to the persistence of oscillatory patterns[18], it predicts that **manipulating those molecular pathways will alter the continuity of conscious experience or memory integration**. For example, drugs or genetic modifications that disrupt PKMζ function (a protein known to maintain long-term potentiation) should lead to oddly "fragile" conscious scenes – perhaps the person cannot hold an image in mind as steadily, or experiences rapid decay of working memory, beyond what classical models predict. Conversely, enhancing such molecular anchors might prolong the duration a conscious thought can persist without rehearsal. These are measurable via cognitive tasks and could differentiate SAN's multiscale model from a purely neural network model that lacks a molecular component.

- **Artificial Implementation – Oscillatory Neuromorphic Systems:** Perhaps most excitingly, SAN provides a template for constructing **artificial systems that could support consciousness**. Your paper asks "if synthetic consciousness is possible, what physical system properties are required?"[14]. Based on SAN, we propose that a candidate architecture would be a **neuromorphic network of coupled oscillators** that mirrors the brain's hybrid computation. In practice, this means hardware where each "neuron" is not a simple on/off unit but an analog oscillatory circuit (e.g. a silicon neuron with a membrane potential analog, or an optical oscillator). These units should be able to synchronize with each other, produce phase-locked loops, and exhibit multiscale dynamics (for instance, small oscillator clusters forming larger beats or envelopes). Critically, the hardware should allow continuous interaction – no global clock that discretely steps all units in lockstep (that would impose an external separation of scales). Instead, time must be an emergent property of the network dynamics. Modern neuromorphic chips (like Intel's Loihi or analog photon-based neural nets) are starting points, but they often still emulate spiking without true continuous wave behavior. SAN suggests enhancing them with on-chip analog frequency generators or using memristor networks that naturally oscillate. The **Self Aware Networks Institute** has begun formulating such designs. In our 2025 paper *"Building Sentient Beings"* (Blumberg & Miller, 2025), we outline a cognitive architecture for AI agents that explicitly incorporates SAN principles[41][42]. We describe an artificial agent as a hierarchy of "cooperative-competitive" oscillatory subsystems, spanning from simulated molecular dynamics (for learning and memory stability) to ensemble oscillations for perception[43]. The **Totality memory model** introduced there divides the agent's world model into layers (perception, imagination, ontology, etc.) implemented by networks of oscillators that continuously update and constrain each other[44]. Early prototypes of this architecture are being developed in simulation, with a focus on achieving *self-organized synchrony* akin to brain rhythms. One concrete design is to use **Kuramoto oscillator networks** (a well-known model of synchronization) augmented with plasticity rules: as these oscillators couple and sync, they would effectively implement the "observer ensembles" of SAN. Preliminary results show that such networks can enter metastable synchronized states that encode inputs robustly and even exhibit memory of past inputs through phase patterns – hinting at a form of rudimentary subjective processing.
- **Artificial Implementation – Integrating LLMs and Embodied Agents:** We acknowledge that human-like consciousness likely also requires higher cognitive

structures (beyond raw oscillations). Therefore, SAN's roadmap involves combining the low-level oscillatory platform with **cognitive architectures and learning algorithms**. For example, one might integrate a Large Language Model as a high-level semantic subsystem within a SAN-based agent, but crucially, that LLM's activations could be controlled and contextualized by an underlying oscillatory "brain" that gives it continuity and grounding. The *Building Sentient Beings* framework suggests using LLMs as just one component in a broader agent that has embodiment, emotional valence (modeled as oscillatory attractor dynamics giving positive or negative feedback), and goal-oriented behavior[45][42]. The result would be an AI that doesn't just chat using statistics, but that *perceives, feels, and adapts in real-time*, regulated by oscillatory control loops like a biological creature. We expect such an AI to demonstrate qualities absent in current LLMs: for instance, consistent self-awareness (tracking its own internal states), resilience to hallucination (since an internal model continuously cross-verifies sensory input with memory via phase locking), and the ability to pursue long-term goals autonomously (thanks to internal rhythmic "clocks" that maintain focus over time).

Each of these predictions and design directions serves to **make SAN falsifiable and actionable**. If brains did not show the predicted wave dynamics, or if an engineered oscillatory system shows no signs of increased integration despite meeting SAN criteria, then the theory would be undermined. Conversely, every piece of supporting data would reinforce that we are on the right track. In effect, SAN turns the philosophical questions into experimental ones: *Is consciousness tied to certain physical dynamics? Let's check.* In doing so, it answers your plea for more than just philosophizing – it offers a concrete research program. This aligns with your statement that we must "open the door to discoveries we cannot yet anticipate" by questioning assumptions[46]. SAN has been doing exactly this outside the traditional academic spotlight, and we invite mainstream researchers to examine, test, and build upon it.

## Conclusion: Toward a Coherent Theory – Humility and Collaboration

In closing, we echo your sentiment that *consciousness is likely more complex than any one current theory can capture*[47]. The Self Aware Networks theory does not claim to have solved consciousness in its entirety – but it does claim to have woven together threads from neuroscience, physics, and computation into a **cohesive framework** that addresses many of the open questions you, Milinkovic & Aru, have articulated. By preserving timestamped contributions from 2017 onward, SAN provides a unique longitudinal case study of theory development, one that has anticipated and paralleled numerous "mainstream" advances[48][49]. We believe this underscores that the time is ripe for a *computational theory of biological naturalism*, as you put it[23] – a theory that neither ignores the messy biology nor confines itself to it, but finds the algorithm *within* the substrate. SAN aspires to be such a theory. It shows how **intrinsic meaning** can arise from coincidence patterns and phase relations, not just from abstract symbols[30].

It demonstrates a path to *empirically implement* consciousness in silico by mirroring the brain's dynamical laws, rather than by brute-forcing intelligence with more data or parameters. And importantly, it remains **open-science and interdisciplinary**: implemented on GitHub, critiqued on forums, and amenable to contributions from all fields[50][51]. This spirit of community and transparency might be essential, because consciousness is too big for isolated silos.

To you, Borjan and Jaan, we say: the gaps you identified are real, and we have been working to fill them. The Self Aware Networks theory already provides answers – or at least well-grounded hypotheses – for many of the issues you raise about scale, hybrid computation, and physicality. We invite you and the broader scientific community to **engage with SAN**: examine our preprints and data[11][52], challenge our assumptions, and perhaps collaborate on experiments. By combining your rigorous perspective on biological computationalism[5][4] with frameworks like SAN that attempt to formalize those ideas, we move closer to the "**unifying brain theory**" that can demystify consciousness[53][54]. As you wisely note, acknowledging the limits of current paradigms is not defeatist but rather "an invitation to explore the unknown"[46]. We wholeheartedly agree. SAN theory is an exploration of that fertile unknown, and the journey has already yielded a map of concepts and mechanisms that answer many of your calls.

It is our hope that this extensive reply illustrates how **Self Aware Networks Theory of Mind** meets the challenge of biological computationalism head-on. The work is far from finished – but together, by empirically testing SAN's predictions and iterating on its models, we can fail better and advance our understanding further[55]. In the true spirit of science, we remain humble yet optimistic: consciousness can be understood by us *and* eventually instantiated in our machines, if we are willing to rethink our assumptions and unite insights across disciplines. SAN is one such unifying effort, and we offer it in response to your questions as a concrete step toward a science of consciousness that fully respects the intricacy of life's computations.

**Sources Cited**:


- Blumberg, M. (2017–2025). *Self Aware Networks* – open research notes and preprints[11][56][15].
- Blumberg, M. (2024). *Bridging Molecular Mechanisms and Neural Oscillatory Dynamics* – Book introducing SAN Theory of Mind[10][25][18][37].
- Blumberg, M., & Miller, M. S. P. (2025). *Building Sentient Beings* – Framework for implementing SAN in AI agents[43][44].
- Milinkovic, B., & Aru, J. (2025). *On biological and artificial consciousness: A case for biological computationalism*. Neurosci. & Biobehav. Rev., **181**, 106524[2][6].
- "The Self Aware Networks Institute: An Open-Science Lab…" (SVGN, 2025) – Historical timeline of SAN development[8][57][20][29].


---

[1] [5] [6] [7] [14] [21] [22] [55] #neuroscience #biobehavioral #reviews #jaan #aru #biological #artificial #consciousness #biological #computationalism #computational #functionalism #biological #naturalism #von #neumann #biological… | Borjan Milinkovic

https://www.linkedin.com/posts/borjan-milinkovic-05952b337_neuroscience-biobehavioral-reviews-activity-7407474586876792832-GhPs

[2] [3] [4] [16] On biological and artificial consciousness: A case for biological computationalism - ScienceDirect

https://www.sciencedirect.com/science/article/pii/S0149763425005251?trk=public_post_comment-text

[8] [9] [11] [15] [19] [20] [50] [51] [57] The Self Aware Networks Institute: An Open-Science Lab Born on Facebook, Powered by GitHub

https://www.svgn.io/p/the-rise-of-the-self-aware-networks

[10] [18] [24] [25] [26] [27] [28] [31] [37] [38] [39] A new book out today "Bridging Molecular Mechanisms and Neural Oscillatory Dynamics"

https://www.svgn.io/p/a-new-book-out-today-bridging-molecular

[12] [13] [23] [30] [34] [35] [36] [40] [46] [47] (PDF) The assumptions that restrain us from understanding consciousness

https://www.researchgate.net/publication/393065964_The_assumptions_that_restrain_us_from_understanding_consciousness

[17] [29] [32] [33] [48] [49] [52] [56] The History of the Super Information Theory Corpus.

https://www.svgn.io/p/the-history-of-the-super-information

[41] [42] [43] [44] [45] Building Sentient Beings: Cognitive Architectures in a Post-LLM World

https://www.linkedin.com/pulse/building-sentient-beings-cognitive-architectures-ramesh-yerramsetti-h5tvc

[53] Consciousness Recognizing Itself: The 2025 Landscape of Self ...

https://www.academia.edu/130231683/Consciousness_Recognizing_Itself_The_2025_Landscape_of_Self_Discovery_Through_Scientific_Methodology_The_Fundamental_Recognition

[54] [EPUB] The Central Questions of Consciousness - Frontiers

https://public-pages-files-2025.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2017.01797/epub