

From Philosophy to Physics of the Mind: Teleological Transduction Theory

Stanislav V. Tregub

Abstract

For millennia, the mind-body problem remained a philosophical issue without an empirical basis and practical implications. In recent decades, natural science has finally started to treat the mind as an object of research. The main approach was to study the ‘neural correlates of consciousness’. Since then, we have accumulated a vast amount of data on neural processes that appear to correlate with observed mental phenomena. However, the explanatory gap between mental and physical is not covered, as the main questions about the mind remain unanswered. What is the mind? What does it do? Why does it do it? How does it do it? The article reviews a theory that aims to answer these phenomenological, functional, teleological, and causal questions from a physical perspective. The theory elucidates physical processes and mechanisms that give rise to mental phenomena in the brain, thus enabling neuroscience to progress from correlational descriptions to an explanation of the physical causes of the mind and solve the mind-body problem.

Key Words: signal processing; neural code; reality model; wave physics; binding problem

DOI: 10.5281/zenodo.20141717

1

1. Introduction

The mind-body problem from a materialistic standpoint can be formulated shortly: how does the body produce the mind? Many philosophers and even neuroscientists share the idealistic dualism that considers the mind a non-physical entity within a physical body. This creates a cognitive dissonance for a scientist, as idealism forbids physical answers and keeps us outside the domain of empirical research. Thus, even if some researchers think that the mind is something “over and above the properties invoked by physics” (Chalmers, 2005, p.15), here we take as a foundation the materialistic stance of natural science in general. It is the only one that can lead to

Corresponding author: Stanislav Tregub

Address: 22-2214 Folkestone Way, West Vancouver, BC, v7s 2x7, Canada

e-mail ✉ stanislavtregub@gmail.com

ISSN 1307-6531, JNphi, Since 2007

www.jneurophilosophy.com

concrete answers to the practical issues embedded in the mind-body problem. However, the answers to the questions about the mind do not appear as if by magic if we take the materialistic stance and say that the mind is physical.

The fundamental questions about an object of research are common to any branch of science. What is the phenomenon? What does it do? Why does it do it? How does it do it? They are called phenomenological, functional, teleological, and causal questions. The last one is the ultimate task of any scientific research.

The question of how is the actual hard problem. It is not a question of why we have subjective experiences (qualia, representations), as some philosophers insist. It is the question of *how* our brains produce them. It is not a question of where those experiences emerge physiologically, as many neuroscientists think. It is the question of *how* physiological processes create the mind and by what physical mechanisms.

For most of the 20th century, neuroscience ignored the mind. It was the century of the brain. We started with Ramón y Cajal's pioneering studies of neurons and have come to state-of-the-art technologies that allow us to screen the whole brain in action and go down to the molecular level. However, the mind remains a mystery. The International Dictionary of Psychology describes this state in desperation: "Consciousness is a fascinating but elusive phenomenon: it is impossible to specify what it is, what it does, or why it has evolved" (Sutherland, 1989, p. 95). Please note that the author mentions only the first three questions about the mind, but not the most important causal one.

When neuroscience finally acknowledged that the mind should be the object of its research, it began looking for the 'neural correlates of consciousness' (NCCs). Thus, it could avoid the causal question and keep on studying the brain without explaining the mind.² Correlating the observed neural activity with reported mental phenomena appears to follow a cause-and-effect logic: when one thing occurs, another follows. When the wheels move, the car moves. The movement of the wheels seems to be the cause. However, it is a logical error. The missing link is the engine under the hood, which is the actual causal mechanism. Showing a correlation cannot be an explanation, as it is the correlation that has to be explained. Even if we get to the finest details of the brain or every neuron, without explaining the physical mechanisms that neurons and the brain use to create the mind, the 'explanatory gap' (Levine, 1983) between the mental and the physical remains open. The engine remains hidden under the hood.

² Many researchers focus on awareness level, which is why the word 'consciousness' is used more often. However, the goal is to explain the mind as a whole. Therefore, in the following text, we will use the word 'mind' as a broad concept that includes both conscious awareness and subconscious processes.

We are not saying that neuroscience is not trying to model the mind. However, all models (at least, known to the author of the present paper) share the same problem. Here is how one review described it: “Most modern theories of consciousness, even the most prominent ones, are not theories at all but rather are (proposed) laws. Some of them can feasibly be tested and could end up being right or wrong, but ultimately, they are not theories because they only describe a relation but do not offer an explanation. It means that when we try to explain consciousness, we arrive at a chasm where we say ‘and then consciousness happens’ in order to magically jump over the explanatory gap” (Schurger and Graziano, 2022, p. 2). To put it short, they avoid the causal question as they do not have an answer to it.

The initiators of the search for NCCs hoped the gap would be covered sometime in the future (Crick and Koch, 1998). Perhaps the future is now, and *this paper calls for a start to the search for the physical causes of the mind.*

We will focus on the theory that is an exception to the negative ‘rule’ of avoiding the questions. It follows the positive rule: a theory should answer the phenomenological, functional, teleological, and causal questions. It should not ‘jump over the gap’, but cover it with explanations and testable predictions.

2. What is the Mind?

When we formulate the phenomenological question in such a concise manner, it becomes clear that we should define the mind. Moreover, if we consider it to be a physical phenomenon, we should define it in physical terms. The definition of an object of study can be just an initial hypothesis that will be developed and empirically tested within the study. It may be wrong or right, but it has to be precise and physically grounded for it to be testable in principle.

The above statements seem to be obvious. However, in neuroscience, the task of defining the mind was not even set. The initiators of the search for NCCs explained their position: “For now, it is better to avoid a precise definition of consciousness because of the dangers of premature definition. Until the problem is understood much better, any attempt at a formal definition is likely to be either misleading or overly restrictive, or both” (Crick and Koch, 1998, p. 97).

However, we can rephrase the above reasoning: it is better to give a precise definition so that we can better understand the problem, checking whether the definition is restrictive or misleading, both or neither. Giving a wrong definition is not a danger if it is formulated as an empirically testable assumption that can be refuted or corrected. Not defining the target is a danger, as it prevents a better understanding of the problem. This is akin to looking for a black cat

in a dark room without defining what a cat is. Even if it is there, we will never find it.

When we say that we should define the object of our study, we use the standard terminology in science, meaning that we should specify the phenomenon being investigated. It can be a physical object or a physical process. The word 'object' means some material thing (substance, entity, system, etc.). The word 'process' means that something is happening to an object. For example, gravitation research is about a process of fundamental interaction of objects, not an object called gravity, or cancer research deals with a pathological cell growth process, not an object called cancer. They will inevitably reach an impasse if they make a category error and take the process for an object. Thus, to understand the causes of the mind, we must first answer the ontological question of whether the mind is an object or a process.

Within the philosophy, the mind-body problem has been typically formulated in the following way: are the mind and body two distinct entities, or a single entity? We can say that philosophy fell into the trap of human language. We give names (nouns) to both objects and processes. Thus, we may easily objectify a process and think of it as an entity named so and so.

The initial problem leading to such an objectification error is a lack of knowledge. If we do not know how the mind works within the body, we tend to simplify the problem by declaring it a special kind of entity living in a body. However, this is an illusion of a way out, and we fall into the vicious circle of the mind-body problem. Thus, there is no end to the discussions between dualists and monists of the idealistic or materialistic camp. However, they have no practical outcome, as the problem based on a category error is ill-posed.

The way out is a simple idea that *the mind is a physical process*. If we accept it as the initial hypothesis, there is no point in arguing whether the mind and the body are one entity or two. A process cannot be one entity with an object, since it is what happens in or to the object. They are just different categories. A process can be 'one' with an object if it happens in it. But it can be 'separate' in the sense that the same type of process can happen in a different object. Thus, the materialistic stance acquires a solid ground as a process in a physical substrate is obviously a physical process. The idealistic stance becomes irrelevant. There is no mind as a transcendent entity, as it is a process in a material entity.

Eliminating the category error does not make the task of understanding the workings of the mind easy, as we face the same fundamental questions. However, it allows us to ask questions in a precise manner and brings hope to find precise answers. If the mind is a physical process, what kind of process is it? Why has it evolved? What does it do? It is exactly these phenomenological, teleological and

functional questions that some researchers found impossible to answer (Sutherland, 1989). Could it be that they were looking in the wrong direction? Sometimes it only takes an insight into the perspective, and the problem seems to evaporate by itself. The theory we review here provides such an insight by offering a clear definition:

The mind is the process of transducing reality signals into information for the purpose of active adaptation to this reality.

It even contains this definition of the mind in its name: Teleological Transduction Theory (Tregub, 2021a).

At first glance, such definition of the mind may be attributed to entities that we intuitively put within a ‘mindless’ category. However, there are two important points. First, our intuitions may be wrong, and entities that we think do not ‘think’ are engaged in much ‘thinking’ (signal processing). Second, TTT is not a new version of panpsychism. All models that adhere to panpsychism, attributing the mind to various systems and even the whole Universe, do not define the mind physically and, thus, make their predictions unverifiable and irrefutable (non-scientific). Instead, TTT offers a physical definition and a testable prediction about the delineating line between systems that possess the mind and those that do not.

According to it, the mind is an active encoding process that does not need external programming and produces information using its own internally generated code. This simple idea settles the ‘other systems argument’ which has been the topic of heated debate for decades. For example, a computer is using the external code and is, thus, mindless. No matter how smart modern artificial intelligence systems may seem, they are still not self-encoding entities.

With a clear definition of a phenomenon, we do not have a problem applying it to any system and checking if it corresponds to the criteria. A good test of our intuitions against a definition of the mind within TTT would be a rock. If we find that a rock processes signals and turns them into representations using a self-generated code for active adaptation to those signals, we might say that this rock possesses the mind. If there are no signs of such a process in the rock, we assume it to be mindless, and this does not counter our old intuition.

The simplicity of the above definition may make us wonder what the problem was. First, any solved riddle appears trivial in hindsight. Second, the impression of simplicity is misleading. It gives clear answers to the phenomenological, teleological, and functional questions, but the major challenge is the causal question of how the mind works. However, clear physical and technical answers to the ‘impossible’ questions allow us to move forward from correlation to causation. For this, we have to elaborate more on the functional question. We need to identify what the mind does and only then explain how it does.

TTT calls this functional architecture diagram the *Perception-Apperception-Action Lemniscate (PAAL) algorithm of the mind*. Lemniscate (from Greek lemniskos — ribbon) is a generic term for all ‘figure eight’ curves. Other terms are taken from the vocabulary of psychology: perception (from Latin perceptio — sensation), apperception (from Latin ad, perceptio — in addition to sensation), and action (from Latin actio — doing).

It is important to emphasize that the order of words in the algorithm's name — from perception through apperception to action — and the numbering of functions do not imply a linear sequence. Moreover, the functional architecture diagram, represented as a closed lemniscate, clearly shows that it is impossible to identify any function as initial within the brain's time and space. However, it does not mean there is no functional hierarchy.

TTT proposes that the projection of the reality model has a functional priority in the algorithm of the mind. The brain constantly projects the model as an assumption about the state of the external and internal environment in the present and the future based on past data, compares the feedforwarded expectation with the feedback of introjected signals, estimates the difference, evaluates the current state of the world and the body, corrects the model, projects its new version and so on in an iterative PAAL cycle of testing the model against reality and optimizing the prediction to navigate reality most efficiently and adaptively.

Many models use the ‘Bayesian brain’ hypothesis, according to which the brain constantly generates predictions about the world based on its past experience and uses the incoming data to minimize prediction errors. It is even taken by many researchers as an axiomatic postulate that does not require any physical explanation. They call it ‘predictive coding’ but do not describe how the coding is done. They call the projected representations ‘predictions’, but do not speak about what they are physically and how they physically interact with introjected signals for error correction. Providing mathematical models of calculating probabilities and minimizing error (surprise, informational entropy) is not enough. Saying that some neural populations perform the job is a ‘causal structure’ approach that does not answer the causal question. Even if we go down to the finest physiological details of specific neural populations, without elucidating the mechanism by which they perform the function, we are in the explanatory gap.

To start building a bridge from mental phenomenology to physical causation, it is better to convert the psychological terms into technical terms. Mental functions, as defined by psychology, are a result of technical functions, as defined by TTT. Sensations and feelings in various modalities of perception are encoded representations of external world signals or internal bodily signals. Our emotions are also encoded representations that evaluate the state of the body or environment and apply a valence to it. Actions are the result of motor

representations. Memories are retrieved representations of past signals. Imagination and thoughts are higher-level representations that may not be associated with currently introjected signals but are always based on past signals.

Whatever function we can think of in psychological terms is a signal-processing and encoding function in technical terms. In contrast to the psychological terminology, such functions are measurable physical processes. This technical approach allows us to address the causal question, as we can test our assumptions about the mechanism experimentally. The mind turns out to be not an intangible ‘thing’ that cannot be studied with scientific methods but a physical process that can and should be explained by science.

However, to understand the causal links and to test them purposefully, we must model the physical mechanisms and provide empirically verifiable hypotheses with specific predictions. This is the crucial test for the theory: does it offer the answer to the ‘How’ question?

4. How does the Mind work?

The definition of functions in technical terms prepares the foundation for answers to the questions about their physical implementation. If we define the mind as a signal encoding process and assume that our brain is the information-producing ‘machine’, we face specific technical questions.

Let’s start with the question about the neural code. The testing ground for the hypothesis is simple: the coding scheme that is able to generate information (code patterns) with an adequate level of representation, while adhering to the time constraints of brain function and maximizing efficiency, is the most likely candidate.

The debate about neural code has been ongoing for decades, and we can roughly distinguish two main approaches: firing rate coding and temporal coding models. The central argument concerns the question of whether the timing of an individual action potential (AP) matters or whether we should take the firing rate as the varying parameter that contains the information. Importantly, both models treat APs as identical spikes (digital signals). Using the human language analogy, they propose that ‘neural language’ consists of one letter pronounced continuously at different speeds or interrupted by various pauses. Such language will have minimal informational capacity. But the ‘digital’ neural code paradigm remains the mainstream. Moreover, it is taken not as a hypothesis but as an axiom beyond debate. However, it leads to several paradoxes that cannot be resolved within the paradigm and are ignored by the mainstream.

The idea that neurons transmit information to each other with a varying average rate of spike generation was attractive because of its

simplicity. Counting averages was feasible even at the level of the last century's equipment capacities. It also seemed to reflect reality, as neurons do change their spiking rate. Since then, a tremendous amount of data correlating changes in the firing rate with mental phenomena has accumulated, but it does not bring us closer to deciphering the code. The researchers cannot even agree about an average rate definition: is it an average over an observation time, the average over several repetitions of an experiment, or both taken together? Imagine that we would try to understand a foreign language by averaging words and sentences to an identical 'beeping' sound with varying tempo. We will be wondering about the averaging window till the end of time, but still be way off from the meaning that the speaker conveys.

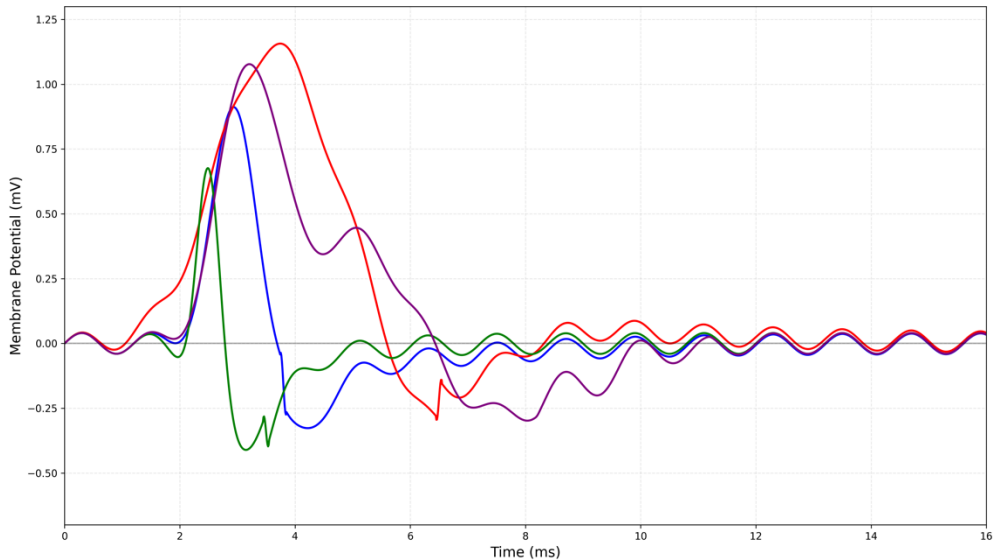
In reality, neurons demonstrate great variability in activity patterns. Rate code idea supporters regard this variability as noise and remove it by applying averaging methods. This creates the '*not-so-smart neurons*' paradox: a highly efficient and precise brain consists of inefficient and noisy neurons that spend so much energy and time creating meaningless patterns. However, the paradox arises from the incorrect assumption that variability is noise. If we change the assumption and consider that varying patterns contain information, this will lead us to the temporal code idea.

The temporal model takes the computer analogy and considers APs as 1s and interspike intervals as 0s of a digital binary code. This significantly increases the capacity of the code and lends credence to the hypothesis. However, it still contains problems. We will not even dive into the problem that interspike intervals differ. Should we count them all as a single 0 or sequences of 0s quantized by a specific rate? The first option ignores the possible information content of an interval length (rhythmic structure), and the second one leads to the question of the system clock (base pulse). Whatever we choose, we still face a critical point: to encode sufficient information, neurons must produce long binary sequences.

In artificial signal processing technologies, to compensate for the temporal limitations of digital coding, engineers speed up processors to billions of hertz. However, this leads to enormous energy consumption up to billions of watts. In stark contrast, the brain operates at dozens of hertz and consumes dozens of watts, yet encodes signals with precision and data compression on a microsecond time scale.

This creates the '*slow neurons*' paradox: how can long and slow spike trains encode fast signals? The brain has no time to produce many spikes to extract information from their average speed, or many spikes and intervals to extract information from their sequence. Signals of the world change their parameters sometimes faster than even one full cycle of an AP.

Despite the paradoxes created by the basic assumption of both models, in mainstream neuroscience, “all spikes are implicitly assumed to be identical in size and duration” (Izhikevich, 2007. p. 280). Thus, the physical reality is sacrificed for the convenience of the models. However, at sufficient temporal resolution, the physical reality reveals itself: the action potential is a smooth, analog wave with phases of subthreshold oscillations, rising potential, peak, falling potential, slight base-line undershoot, and return to the base pulse. Real, not theoretical, APs differ in onset time, width, amplitude, and the shape of the rising, falling, and after-potential phases (Figure 2):



The conclusion that seems obvious in hindsight is that the brain uses the informational capacity of each wave. Using the language analogy, neurons speak words that consist of various letters. Taking a musical analogy, neurons play melodies that consist of various notes.

TTT takes the musical analogy and proposes the *Symphonic Neural Code (SNC) hypothesis* (Tregub, 2021a). According to it, each action potential is a ‘note’ of the mind that has individual waveform characteristics (period, amplitude, phase trajectory), which encodes the meaning created and transmitted by the neuron. This hypothesis elegantly resolves the paradoxes and explains the observed tremendous computing power, efficiency, and speed of the brain. The neurons turn out to be fast and smart enough for the world, as the AP internal parameters develop on the same microsecond timescale as the signals, and the information density of each AP is high. This does not reject the role of AP sequences for longer-timescale representations. Each note has its place in melodies and harmonies (population code), rhythms (temporal code), and tempo variations (firing rate). Thus, SNC is a reconciling hypothesis for existing models as it combines their predictions and offers a novel explanation where old ones fail.

TTT looks at the coding process from a physical and technical perspective. The signals of the world are continuous energy flows (electromagnetic, sound, mechanical, thermal, chemical, and others). The brain faces a dilemma, or rather a technical task. On the one hand, a continuous signal must be quantized since information is a code pattern, i.e., a sequence of discrete values (in signal processing theory, such signals are called digital). On the other hand, if reality is a set of continuous signals, the model of this reality must also be a set of continuous representations (analog signals). How does the brain solve the problem? *TTT suggests that the brain uses a hybrid analog-digital-analog signal processing technology.*

At the single-cell level, TTT models the neuron as an active non-linear analog-digital-analog conversion (ADAC) filter that produces complex wave patterns, representing incoming wave signals. This is a game changer: the ‘not-so-smart neuron’ that the old integrate-and-fire model defined as a simple capacitor that sums currents and produces identical digital spikes turns out to be a smart wave processor. For brevity reasons, we cannot go into the technical details, physical explanations, and mathematical descriptions provided by TTT.³ We will list models that the theory suggests and describe them on a general level.

The first step is to model the input processing by the dendrites. The classical view of the dendrite is that of a passive cable (Rall, 1959). In this framework, synaptic inputs are treated as current sources, and the dendrite’s role is to sum these currents as they spread towards the soma. This creates the ‘*smart cable*’ *paradox*: the cable model describes a passive transmission line, while modern physiology has unequivocally shown that dendrites are active, complex, and dynamic computational units. The classical model renders the sophisticated computation of the dendrites functionally irrelevant. Presynaptic action potentials, stripped of their analog information, become identical inputs, producing an identical output clone. In this view, the neuron annihilates information rather than creating it (the not-so-smart neuron paradox appears again).

TTT proposes the *Dendrite Antenna Model (DAM)* that describes dendrites as active antenna-like arrays with pattern-recognition function. The DAM mathematics describes how synapses launch electrochemical waves, how dendritic branches guide them, and how voltage-gated channels actively modulate them. The model offers several testable predictions about biophysical properties of this circuit: the dendrite’s resonant frequency defines its function as a matched filter tuned to specific spatiotemporal patterns; the effective electrochemical inductance of voltage-gated ionic currents is the biophysical basis for resonance and bandpass filtering; the dendritic

³ As the list of references is also limited, we will not refer to experimental evidence for the hypotheses and predictions of TTT. The full list contains hundreds of items and can be found in the original publications (Tregub, 2021a, b, c, d, e).

output at the soma is the result of the coherent superposition of waves from all inputs.

Next, the *Soma Wave Processor model (SWAP)* describes the central unit of a neuron as a nonlinear relaxation-type oscillator whose function is to map the high-dimensional analog dendritic input onto a structured, information-rich output wave. The SWAP mathematics describes the *ADAC cycle* of the soma. At the digital step, the soma samples and quantizes the continuous input from dendrites, generating the discrete action potential. At the analog step, the output is a wave packet with a specific waveform that encodes the parameters of the initial analog input.

The discreteness of the action potential is not in conflict with its analog waveform. The somatic threshold acts as a binary decision determining whether a wave is launched. However, this digital step does not specify the precise shape of the resulting wave. The analog nature of the propagating wave packet is determined by the initial conditions at the soma and the dendritic input parameters. Thus, TTT reconciles a fundamental biological observation of the all-or-nothing AP generation rule with the experimentally observed variations in AP waveform.

This ‘smart neuron’ model not only explains the observed but provides testable novel predictions: somatic response depends on coherent cross-frequency coupling of dendritic inputs; the soma dynamically reconfigures its impedance profile as an active part of pattern-specific encoding; the AP waveform encodes the spatiotemporal pattern of the input. This is the essence of the SNC hypothesis in physical and technical terms.

If this hypothesis is falsified, it would merely indicate that waveforms carry no informational load and would not undermine other predictions of the DAM and SWAP models. If, however, it is confirmed, this would signify a neural code of far greater complexity than classical models allow. This added complexity could, in turn, explain the neural code’s remarkable speed, efficiency, and information saturation, solving the long-standing paradoxes.

This paradigm shift creates additional difficulties in deciphering the neural code, as it is necessary to analyze all the subtleties of individual AP parameters. The TTT prediction that makes the task slightly easier is that the waveform individuality does not imply infinite variability. The brain likely employs a limited set of waveforms. As in music, vast representational power emerges from the combinatorial sequencing (rhythms, melodies) and interaction (harmonies) of fundamental units (notes). Even a limited set represents a manifold increase in information capacity over a digital code proposed by the old models.

We have to note another problem of the classical cable model. It is built upon the cable equation, which describes a diffusive process, leading to a ‘*bad cable*’ paradox: the model predicts a signal that

spreads with infinite speed but decays exponentially within millimeters. This directly contradicts measured finite velocities and high-fidelity long-distance transmission by the axons. The fundamental flaw of the model necessitates post-hoc parameter adjustments and is only superficially patched by adding the Hodgkin-Huxley (HH) model, which describes point-wise ion-channel dynamics (Hodgkin and Huxley, 1952). This is interpreted as a saltatory conduction, where a ‘dying’ digital signal is revived at each node of Ranvier. Combining a non-wave component with a discrete model to describe a continuous phenomenon creates a fundamental mathematical inconsistency that reflects the foundational physical inconsistency of the paradigm.

The *Axon Waveguide Model (WAM)* resolves the paradoxes and inconsistencies. In this framework, the AP is a soliton-like wave packet, the internode is a waveguide, and the Nodes of Ranvier are impedance-matching and phase-shifting wave repeaters. Thus, the active part of signal transmission is not about boosting diffusing digital pulse, but about balancing AP wave dispersion (various frequency components traveling at different speeds) by the focusing nonlinearity of electrochemical oscillations at the nodes. The equations of WAM naturally incorporate the ion channel dynamics of the HH model, restoring the mathematical consistency of describing electrochemical wave propagation in the axon. Thus, TTT does not reject the empirically tested and physically viable HH model, but uses it on a new explanatory level.

The WAM quantitative and qualitative predictions explain the already measured biophysical properties and suggest novel research directions. In contrast to the diffusive cable model, it predicts a specific AP propagation speed for various types of axons, and its velocity equation produces calculations that match measured values without curve fitting, which is a common practice in research based on cable theory. It also proposes novel testable predictions: the action potential within the myelinated internode will exhibit wave dispersion determined by the frequency-dependent properties of the waveguide; nodes act as active wave repeaters that perform impedance matching and phase-sensitive gain to correct dispersion and restore the original action potential waveform.

After dealing with information encoding and transmission issues, TTT tackles the information integration problem. In neuroscientific discourse, it is called the ‘binding problem’ (Feldman, 2013). How do the encoded representations of a vast number of signals constantly processed by the brain combine into what we experience as a unified flow of the mind? How does the brain create a coherent model of reality while maintaining signals’ identity and, at the same time, preventing the world picture from falling apart into separate pieces? Usually, these two aspects are called the segregation problem and the combination problem.

There is an old hypothesis that the binding problem (BP) is solved by synchronous firing of neurons at a specific rate (Llinás, 1998; Singer, 1999). It came naturally: if we think of a neural code as a set of identical ‘shots’, integration becomes simultaneous firing of those shots. But the idea has two major pitfalls.

First, it cannot be a solution to the problem of combining the variety of subjective experiences into a coherent whole, as the variety vanishes (identical shots add up to a firing discharge). That is why the initial author of the ‘binding-by-synchrony’ idea suggested that differentiation should be supported by another unknown mechanism (von der Malsburg, 1999).

Second, the brain functions at multiple frequencies. Why would the oscillatory dynamics be so complex if all that neurons need to do is resonate at one frequency? The authors of one review noted: “We do not yet have a coherent picture on why the brain oscillates at such a diverse range of frequencies and how these different oscillatory processes functionally complement each other” (Uhlhaas, 2008, p. 156).

TTT uses the same wave physics first principles for a single cell level and for the whole brain functioning. These principles help it to resolve the BP from a physical perspective (Tregub, 2021c). We will not delve into details due to the limitations of the article. For a short and intuitive explanation, we will take a musical analogy actively used in the model. To start, we can formulate the binding problem using musical terminology. What mechanism allows the orchestra of the brain to play a coherent symphony with complex polyphony and polyrhythm while preserving individual notes?

The intuitive musical analogy helps us understand *the Binding-by-Harmony Hypothesis (BHH)* offered by TTT. It explains how the brain solves both segregation and combination aspects of BP. Physically, the nervous system is a set of oscillators with a wide range of characteristics (just like instruments in an orchestra). However, the variety of frequencies, amplitudes, and phases of oscillations is not an obstacle to unifying them into a coherent reality model (just like a complex but harmonious symphony). The physics of oscillator interactions allows different wave patterns of representations (qualia) to be superimposed on each other while preserving their frequency and phase characteristics.

The variety goes hand in hand with unity. There is no need for special mechanisms for integration and segregation. Neural ensembles use the universal physical mechanism of frequency-phase coupling at harmonic ratios to synchronize wave patterns of their activity while retaining individual characteristics. Using the musical terminology, we can say that the orchestra of the brain plays a harmonic symphony of the mind where each note, melody, and rhythm has its unique place. The metaphor becomes a physically grounded analogy.

TTT proposes the *Brain Harmonic Octave Scale (BHOS)* model as the binary logarithm representing an octave inversion of all frequency levels. The predicted scale corresponds to the phenomenological taxonomy of the diverse range of ‘brain rhythms’ obtained during many years of studying neural oscillations. TTT offers a detailed account of the functional role of different frequencies, how they interact and complement each other. Thus, one more riddle acquires a physical explanation.

Taking the same technical and physical perspective, TTT deals with the question of how the encoded representations are stored and retrieved. According to its *Wave Memory Model (WMM)*, our memories are wave patterns where neurons participate as oscillators with specific characteristics tuned to the parameters of this particular pattern. The storage is not a specific place in the brain (memory module) and not a specific cell (‘grandmother neuron’). Every neuron that participates in a wave stores part of the memory wave in the settings of its impulse response.

WMM explains differentiation and integration, sequential and parallel processing, layering and associativity, stability and flexibility, high capacity and efficiency of our memory. It also explains that memories are not ‘carved in stone’ but are reconstructed each time as a dynamic wave, thus being prone to errors or even false confabulations.

TTT uses the term ‘memory’ in a technical sense: storage and retrieval of representations. This concerns the whole reality model, as the conscious and subconscious levels of representations have the same wave nature. The difference is only in the specific functional role of amplitude-frequency parameters of the waves and their interactions. This explains why ‘causal structure’ theories that are looking for a ‘seat of consciousness’ in a specific place in the brain are not even wrong. They only seem to offer testable predictions about where to look for consciousness. However, as they do not concern themselves with the causal mechanism, they are unfalsifiable.

No wonder the experiments and ‘adversarial collaboration protocols’ confirm predictions of rival theories about the opposite sides of the neocortex being the place where consciousness happens. As the experiments are still testing correlations, not causations, the activity of neural populations may correlate with conscious experience in one experiment and not correlate in another. The secret is not in the location but in the mechanism.

TTT resolves the predictive coding issue from the same wave physics first principles. The comparison of predictions (stored representations) and results (newly produced representations) is also a wave process. According to the PAAL algorithm model, the projection and introjection waves superimpose to produce the current picture of the world. It is analogous to the hologram technology, where the generated object waves are superimposed on the reference waves. The frequency-phase

interaction mechanism ‘compares’ the waves and produces the interference pattern that is the final image. This explains how predictions and results meet in the ‘Bayesian’ brain.

The iterative PAAL loop happens on all spatial and temporal levels. It is a constant process of modelling reality. The stability and dynamic adaptability of the reality model depend on the coherence of the two-way interaction of waves. However, as the reference wave (projection) has the functional priority over the object wave (introjection), the picture of the world is not what we actually see but what we expect to see (this concerns not only vision but all modalities). In this sense, believing is seeing.

While many theories express the idea that expectation is primary and many experiments confirm it, TTT shows the algorithmic flow of the mind that produces the phenomenon and explains the causal mechanism behind it. It explains physically why we experience perceptual illusions. They are not errors in perception as the name suggests, but projections of the adaptive reality model that ‘dictates’ the way we perceive things and compares it with the incoming signals. The PAAL model also explains the pathological disruption of the algorithm when the reference wave is not interacting with the object wave, and the brain produces representations detached from current reality (hallucinations).

After explaining the normal functioning of the mind, TTT logically proceeds to model the pathologies. We cannot understand what is happening physically in the breakdown if we do not know the causal mechanisms of the norm. So, the ultimate practical goal of any theory of the mind is to be able to explain the pathologies from a physical perspective.

TTT starts with an overall description of possible breakdowns of the mechanism (Tregub, 2021d) and goes on to model the systemic pathologies such as autism and schizophrenia (Tregub, 2021e). It suggests a new approach to the diagnosis and classification of all mental pathologies, based on assessing the state of the physical mechanism of the mind offered by the theory. This should help psychology and psychiatry to move beyond the current phenomenological approach of classifying symptoms into syndromes to causal diagnostics of the brain state responsible for those symptoms. Solving the mind-body problem from a physical perspective should have practical implications.

5. Conclusion and outlook

As in any natural science, the ultimate task of neuroscience is the causal explanation. “Science seeks a causal chain of events that leads from neural activity to subjective percept; a theory that accounts for what organisms under what conditions generate subjective feelings,

what purpose these serve, and how they come about. If such a theory can be formulated — a big if — without resorting to new ontological entities that can't be objectively defined and measured, then the scientific endeavor, dating back to the Renaissance, will have risen to its last great challenge. Humanity will have a closed-form, quantitative account of how Mind arises out of matter" (Koch, 2004, p. 34).

We must move from calls for a theory to formulating such a theory. Otherwise, we will remain in the explanatory gap with an eternal 'big if'. The Teleological Transduction Theory (TTT) is based on the analysis of the data accumulated by neuroscience over a century. Its hypotheses stem from this analysis, and its predictions are potentially testable at the current level of our research equipment capacities. TTT does not resort to 'ontological entities' inaccessible to empirical study. It is, probably, the first theory that rises to the great challenge of a physical account of how the Mind arises out of Matter. It allows science to move from describing mental and physiological phenomena to explaining their physical causes.

Conflict of interest statement

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- Chalmers, D. Facing up to the Problem of Consciousness. *Journal of Consciousness Studies* 1995 Vol.2, № 3: 200-219 doi:10.1093/acprof:oso/9780195311105.003.0001
- Crick F, Koch C. Consciousness and neuroscience. *Cereb Cortex*. 1998 Mar;8(2):97-107. doi: 10.1093/cercor/8.2.97.
- Feldman J. The neural binding problem(s). *Cogn Neurodyn*. 2013 Feb;7(1):1-11. doi: 10.1007/s11571-012-9219-8.
- Hodgkin, A., Huxley, A. A quantitative description of membrane current and its application to conduction and excitation in nerve. *J Physiol*. 1952 Aug;117(4):500-44. doi: 10.1113/jphysiol.1952.sp004764.
- Izhikevich, E. *Dynamical Systems in Neuroscience: The Geometry of Excitability and Bursting*. The MIT Press, Cambridge, MA. 2007. pp. 280-281. ISBN: 9780262276078
- Koch, C. The Quest for Consciousness: A Neurobiological Approach [Excerpt]. *Engineering & Science*. 2004. 67(2), 28-34. Excerpt reprinted from *The Quest for Consciousness: A Neurobiological Approach*, by C. Koch, 2004, Roberts and Co.
- Levine, J. Materialism and Qualia: the Explanatory Gap. *Pacific Philosophical Quarterly*. 1983. Vol. 64, № 4. pp. 354-361 doi:10.1111/j.1468-0114.1983.tb00207.x
- Rall W. Branching dendritic trees and motoneuron membrane resistivity. *Exp. Neurol*. 1959 vol. 1, pp. 491-527.
- Schurger A, Graziano M. Consciousness explained or described? *Neurosci Conscious*. 2022 Jan 21;2022(2):niac001. doi: 10.1093/nc/niac001.
- Singer W. Neuronal synchrony: a versatile code for the definition of relations? *Neuron*. 1999 Sep;24(1):49-65, 111-25. doi: 10.1016/s0896-6273(00)80821-1.
- Sutherland S. *The International Dictionary of Psychology*. New York, NY: Crossroads Classic. 1989. ISBN 10:0824525094

- Tregub, S. Algorithm of the Mind: Teleological Transduction Theory. Tregub S.V. Vancouver. 2021a. ISBN: 979-8451716403 <https://stanislavtregub.com/projects/mind-algorithm/>
- Tregub, S. Technologies of the Mind. The Brain as a High-Tech Device. Tregub S.V. Vancouver. 2021b. ISBN: 979-8452552406 <https://stanislavtregub.com/projects/mind-technology/>
- Tregub, S. Harmonies of the Mind: Physics and Physiology of Self. Tregub S.V. Vancouver. 2021c. ISBN: 979-8453148561 <https://stanislavtregub.com/projects/mind-harmony/>
- Tregub, S. Inner Universe. The Mind as Reality Modeling Process. Tregub S.V. Vancouver. 2021d. ISBN: 979-8453227266 <https://stanislavtregub.com/projects/inner-universe/>
- Tregub, S. Dissonances of the Mind. The Physics of Mental Disorders. Tregub S.V. Vancouver. 2021e. ISBN: 979-8878308229 <https://stanislavtregub.com/projects/mind-dissonances/>
- von der Malsburg C. The what and why of binding: the modeler's perspective. *Neuron*. 1999 Sep;24(1):95-104, 111-25. doi: 10.1016/s0896-6273(00)80825-9.
- Llinás R, Ribary U, Contreras D, Pedroarena C. The neuronal basis for consciousness. *Philos Trans R Soc Lond B Biol Sci*. 1998 Nov 29;353(1377):1841-9. doi: 10.1098/rstb.1998.0336.
- Uhlhaas PJ, Haenschel C, Nikolić D, Singer W. The role of oscillations and synchrony in cortical networks and their putative relevance for the pathophysiology of schizophrenia. *Schizophr Bull*. 2008 Sep;34(5):927-43. doi: 10.1093/schbul/sbn062.