SHORT THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY (PHD)

The Nature of Consciousness, A Category Theoretical Proof of the Fermionic Mind Hypothesis

By Éva Katalin Déli
Supervisor: Dr. Zoltán Kisvárday

UNIVERSITY OF DEBRECEN
DOCTORAL SCHOOL OF NEUROSCIENCES

DEBRECEN, 2024

The Nature of Consciousness, A Category Theoretical Proof of the Fermionic Mind Hypothesis

By Éva Katalin Déli, MSc

Supervisor: Dr. Zoltán Kisvárday, PhD, DSc

Doctoral School of Neurosciences, University of Debrecen

Head of the **Defense Committee**:	Dr. Prof. Dr. Fülesdi Béla, PhD, DSc
Reviewers:	Dr. Molnár Edina, PhD
	Dr. Andjelka Hedrih, PhD
Members of the Defense Committee:	Dr. Frecska Ede Ottó, PhD
	Dr. Nagy Attila, PhD

The PhD Defense takes place at the Lecture Hall of Bldg. A, Department of Internal Medicine, Faculty of Medicine, University of Debrecen, 2024, September, 30, 1 pm

1 Introduction

"Who would believe that so small a space could contain the images of the whole universe?" — LEONARDO DA VINCI

The nature of the mind has been an intriguing puzzle throughout history. Modern neuroscience supports the materialist view that consciousness is the product of brain circuits (Churchland, 1981). Nevertheless, spectacular advances left the question of consciousness unanswered (Kandel et al., 2021).

Historically, significant breakthroughs have often come from cross-pollinating fields in the sciences. Physics, dealing with matter-energy interactions, is particularly well-positioned to enrich other fields. For example, perception can be represented as a thermodynamic cycle. Perception and attentional focus, the source of synaptic changes (Manohar et al., 2018; Inzlicht et al., 2018), turn learning and memory formation energetically costly (Crossley et al., 2023).

1.1 Consciousness Hypotheses

Twenty-five years ago, the intellectual duel between neuroscientist Christof Koch and philosopher David Chalmers challenged the limits of human understanding regarding the nature of consciousness. The climax of this debate was a wager: Koch bet Chalmers a case of wine that, within 25 years, by 2023, researchers would identify a "clear" neural pattern underlying consciousness. This wager sparked a surge of interest in this perplexing topic.

The nature of consciousness, a puzzle that has perplexed philosophers for millennia, became a scientific challenge. Advancements in neuroimaging technologies, such as functional magnetic resonance imaging (fMRI), transcranial magnetic stimulation (TMS), and implanted brain electrodes, have since generated vast amount of data. Although these data sets are processed and analyzed by sophisticated algorithms enhanced by artificial intelligence, a definitive neural correlate of consciousness remains undiscovered.

The absence of a breakthrough in identifying a neural correlate of consciousness may suggest our current approach is misguided. The limitations of traditional methodologies, heavily reliant on empirical data and computational analysis, are becoming apparent. The field might be ripe for a paradigm shift that embraces interdisciplinary perspectives and transcends conventional scientific paradigms.

The Koch-Chalmers wager, with its promise of a case of wine, has spurred research and sparked debates and discussions, shaping the direction of consciousness studies. Different disciplines, including philosophy, neuroscience, and psychology, have offered consciousness theories, indicating the unyielding human desire to unravel the mysteries of the mind. Two main theories, Global Workspace Theory (GWT) and Integrated Information Theory (IIT), have recently dominated the discussion.

GWT posits that local processes are unconscious (Baars, 1988; Dehaene and Changeux, 2011). When a local neuronal assembly reaches the activity threshold for broadcasting, a phenomenon known as 'ignition' occurs. This 'ignition' refers to the global broadcasting of information that reaches the brain's global workspace. It is a central concept in GWT, often linked to recurrent processing, for instance, in the frontoparietal cortex (van Vugt et al., 2018).

IIT interprets consciousness as the capacity to integrate information (Block, 2005; Nagel, 1974). The 'phi' (Φ), a quantifiable measure of integrated information, measures the level of consciousness. However, 'phi' is an abstract concept and difficult to measure, posing a significant limitation to the theory. A recent comparison study has shown both theories failure to explain consciousness (Ferrante et al., 2023; Melloni et al., 2023).

FMH incorporates current neuroscience, psychology, and physics understanding to build a comprehensive physical system theory governed by physical laws. It treats the brain as a thermodynamic system capable of intellectual growth by transforming information input by the sensory system. Moreover, the reversibility of the cognitive or perception cycle introduces a probabilistic or quantum behavior, allowing the mathematical framework of quantum mechanics to explain behavior and decision-making. Sensory stimuli drive bottom-up activation, while the resting state imposes a top-down self-referential stability. This regulatory duality causes our difficulty in controlling or retracing our thoughts and emotions. Possible research introduced in a later section can validate FMH.

1.2 Mathematical Framework: A Category-Theoretical Perspective of Consciousness

Consciousness science serves as a bridge connecting neuroscience and psychology. Nevertheless, psychology and social sciences increasingly turn to physics, rather than neuroscience, to explain problems in cognition and human behavior. Thus, various physics frameworks provide superior cognitive science insights (Déli, 2020a,b; Goldenberg et al., 2018). For example, quantum cognition turned to quantum theory to explain the context dependence of memory, perception, and decision-making (Basieva et al., 2019; Chang et al., 2019; Dennett, 2018; Wendt, 2015). The fermionic mind hypothesis (FMH) follows this path but goes a step further. It recognizes the mind becomes an abstract structural and organizational mirror of the physical environment (Jeffery & Rovelli, 2020; Gruber et al., 2022; Tsao et al., 2018) by adopting a particle-like, fermionic organization (Deli, 2020a,b).

A category is defined as a collection of objects with morphisms between them. Stripping objects to their defining qualities an abstract structures reveals their fundamental properties, which aids a deeper understanding. Therefore, category theory can uncover connections between seemingly unrelated fields, such as neuroscience, psychology, general relativity, quantum theory, and beyond (Anderson, 2021; Ehresmann & Gómez-Ramírez, 2015; Phillips, 2022).

Central to category theory is the Yoneda lemma, which allows for equivalence between two objects, A and B, up to isomorphism, based on their relationships with other objects within a category (Tsuchiya & Saigo, 2021a). The bird's eye view provided by category theory can assess the shared characteristics of consciousness and elementary particles to uncover their fundamental relationships. The Yoneda lemma can test the validity of the FMH and offer a solid mathematical foundation for the relational understanding of the fermionic mind hypothesis.

2 Aims

This dissertation furthers the understanding of the nature of consciousness based on FMH by using the tools of physics and mathematics. In mathematics, category theory can verify consciousness' relationship to physical principles and laws, providing a framework for proof of the FMH.

I focused on the following points in my study of the nature of consciousness:

- Perception: What are the energy-information consequences of perception?
- The resting state: What is the role of the resting state in energy regulation and behavior?
- Emotions: How can emotions produce motivation, and what is their role in cognition?
- Category theory: What can the Yoneda lemma's global perspective teach us about consciousness?

3 Materials and Methods

This dissertation examines consciousness based on its physical characteristics. Cognition is founded on interacting with the external environment via the sensory system, representing an energy-information exchange. The above relationship permits the thermodynamic study of the cognitive cycle, where sensory information fuels intellect generation. I used the Kullback-Leibler divergence (DKL) to uncover how much learning takes place. DKL is a type of statistical distance: a measure of how one probability distribution p is different from a second, reference probability distribution q. Learning is a cognitive updating caused by a stimulus, a statistical distance represented as a surprise between the current and expected output. It is deemed that, like the gravity field governs particle movement, temporal orientation defines our cognitive freedom. Temporal orientation rests with the resting state.

I examined the role of the resting state in energy regulation and behavior. In the study of perception, the resting state's autonomic maintenance is fundamental and ensures a closed thermodynamic cycle. As a result, consciousness is a temporal, complexity-generating system. Memory forming ensures a wide variety of reference frames. Moreover, the stability of the resting state requires the presence of emotions. I examined emotion regulation from an energetic perspective, how emotional experiences change the perception of time, and how it relates to their action-producing power (Remmers & Zander, 2018; Rudd et al., 2012).

In the context of theoretical neuroscience, I modeled consciousness using category theory by considering consciousness as a physical system. In this context, consciousness forms morphisms analogous to fermions. Category theory, which studies structures and their relationships, offers a panoramic perspective that can demystify connections between diverse disciplines, uncovering similarities or equivalence between them. Category theory, like cognitive science, is concerned with modeling the (compositional) structure of some "domain" of interest. Category theory can differentiate between different reference frames.

3.1 Yoneda Lemma's Global Perspective on Consciousness

At the core of category theory lays the Yoneda lemma. In category theory, the Yoneda lemma can prove the equivalence of two objects A and B in a category by proving the analogy of their relationships to other objects (Tsuchiya & Saigo, 2021a). Applying category theory and the Yoneda lemma to the fermionic mind hypothesis establishes a mathematical foundation elucidating the relational interplay between consciousness and material systems. This innovative framework can verify the hypothesis by aligning harmoniously with existing empirical evidence.

This dissertation employs the Yoneda Lemma, a foundational result in category theory. The Yoneda Lemma is a key theorem in category theory that relates objects in a category to sets of morphisms from these objects to other objects in the category. It allows for a deeper understanding of consciousness based on analogies to physical processes through morphisms, providing a new perspective on the cognitive process.

The Yoneda lemma can prove the equivalence of two objects A and B in a category by proving the analogy of their relationships to other objects. It states that for any category C and any objects A and B in C, the set of natural transformations from the hom-functor Hom(_,A) to Hom(B,−) is isomorphic to the set of morphisms Hom(B,A).

3.2 Model Formulation

Category of Neural Networks: We define a category C where:

Objects are conscious minds.

Morphisms define relationships based on physical principles.

Hom-Functor: For a set of functors, the hom-functor Hom(_,A) maps a set of fermionic characteristics to consciousness through morphisms.

Application of Yoneda Lemma: By the Yoneda Lemma, the functor map of consciousness is isomorphic to Hom(_,A) fermions. Thus, how physical laws determine consciousness characteristics can be fully understood by studying the physical morphisms of fermions.

Representation of Time Dynamics: Time dynamics can be represented by functors and morphisms between time points. I used diagrams to represent objects, morphisms, and natural transformations in the categories F and C. I generated time-series plots and commutative diagrams to visualize the dynamics of consciousness analogies.

The dissertation shows that consciousness and fermions have identical connection maps to fundamental physical theories, such as quantum mechanics, and show similar organizations. In conclusion, the Yoneda Lemma can provide a deeper insight into consciousness' representational and dynamic properties. This categorical approach opens new avenues for exploring consciousness' function, providing a robust mathematical framework for the theoretical study of consciousness.

3.3 Possible testing of the hypothesis

Based on Landauer's principle, we can study the neural system's energy relationships in dish brain organoids. Landauer's principle (1961), a basic principle of the thermodynamics of information processing, holds that any logically irreversible manipulation of information, such as the erasure of a bit or the merging of two computation paths, must be accompanied by a corresponding entropy increase in non-information-bearing degrees of freedom of the information-processing apparatus or its environment. Similarly, computer simulations can investigate emotions and energetic correlations, making testable predictions about perceptual, cognitive, and control tasks.

Testing the grip strength, puzzle-solving ability, and generosity before and after mental challenges or games can indicate mental energy changes. The expectation is that winners would show an increase in grip strength due to an endothermic shift in mental energy. In contrast, losers would show grip strength, creativity, generosity, and problem-solving ability loss.

4 Results

Based on the perception cycle, resting state, and emotional regulation data, I formulated a novel hypothesis, the FMH. Using the Yoneda lemma from category theory, I proved that consciousness is homomorphic to fermions.

4.1 Perception: The Brain's Energy Relationships

The energy cost of synaptic changes integrates the brain into the environment's broader energy and information cycle. I used this novel assumption to look for the thermodynamic cost of perception, attentional focus, learning, and motivation (Appendix). I gained detailed insights into the brain's dynamic integration with the environment.

4.2 The Resting State in Energy Regulation and Behavior

Concerning the brain's entropic relationships, I concluded that the brain's autonomous resting balance is analogous to an equilibrium position. Therefore, based on Shannon's information entropy, the perception cycle's direction relates to temporal directionality. Endothermic cycles produce future orientation, whereas past orientation confines exothermic cycles; see Table I. I interpreted the cycle's direction as intrinsic angular momentum, with the endothermic cycle representing up spin and an exothermic cycle forming down spin.

Table I. The physiological consequences of different brain states The thermodynamic and psychological implications of basic emotions.

	Reversed Carnot cycle **High entropy resting state**	**Carnot cycle** **Low entropy resting state**
Mental state	Positive emotions	Negative emotions
Activations	Information poor	Information rich
Subjective sense of time expands	The wealth of time inspires confidence	Time pressure causes impatience
Degrees of freedom	Expanding	Loss
Thermodynamic consequences	Endothermic cycle: absorbs energy and entropy	Exothermic cycle: dumps energy and entropy
Consequences for the organism	Intellect	Mental and immune problems

Emotions Regulate the Cognitive Cycle by forming Motivation

Emotions serve the brain's thermodynamic regulation and homeostatic balance by producing attitude, a context-dependent and involuntary action-producing force. In this context, emotional valence is an instant feeling, with positive states representing preference and negative ones, aversion. Likewise, mathematical calculations of the reversed Carnot cycle (preference) indicates energy accumulation and future focus (Appendix). Inversely, exothermic, low entropy states (aversion) reduce synaptic complexity and limit cognitive freedom. The process dissipates energy via criticism, destructive behavior, or violence.

4.3 Temporal Orientation: The Orthogonality of Sensory Reality and Cognition

The nature of consciousness, whether a continuous stream of experiences, emotions, or thoughts occurring at specific moments, has been a longstanding and intriguing question among philosophers, psychologists, and neuroscientists. However, there is a significant duration of unconscious processing before conscious perception and awareness emerge. Moreover, despite our continuous flow of experiences, we remember episodes as discrete sequences of events and see the future as steps on the temporal measuring stick of our imagination, which is consistent with discrete processing. Thus, consciousness shows continuous and discrete qualities and a framework that reflects both and, in this way, might represent how we perceive the world around us.

For example, directly manipulating attention postdictively, i.e., at short time delays (Shen et al., 2020), influences choice. Therefore, cognition is not instantaneous, moment-by-moment consciousness constructions but incorporates information after perception by the so-called postdictive effect. In the above example, substantial periods of continuous unconscious processing precede discrete conscious percepts (Herzog et al., 2020).

Next, the cognitive process was studied as a source of duality. Conscious experience is intrinsically tied to the "now," representing an unbroken stream of awareness, seamlessly transitioning from one moment to the next. Nevertheless, each present moment uniquely represents discrete conscious experiences of successive "nows" that become memories. This twofold nature of consciousness signifies a constant shift in focus, where observation changes its underlying parameters, making it impossible to grasp—analogous to the observer effect. In quantum mechanics, the observer effect is the change of an observed system by observation. It also suggests that measurement causes the discrete collapse of the wave function, the so-called wave-particle duality.

Although category theory is based on highly abstract definitions and theorems, it organizes around the commuting square in various contexts. Its main idea views cognition as a computational processing system that produces cognitive representations via various paths. These different paths represent a shift in focus.

The lens of a "monoid" can effectively examine the concept of conscious duality. A monoid can encompass multiple morphisms, each representing the ongoing flow of consciousness. In a monoidal category where only the identity morphism exists, the arrow returns to itself, symbolizing the continuous unfolding or extension of conscious experience (Figure 1).

$$id_a \circ f = f = f \circ id_a$$

The concept of consciousness as a monoid closely aligns with the brain's ability to coalesce disparate and disorderly information into a unified perception (Tring et al., 2023). Unity aligns along a temporal arc (Smallwood et al., 2021) that endows intellect with a remarkable predictive ability (Northoff et al., 2019). "Therefore, time touches the eternal now at each moment,…disappears moment by moment and is born moment by moment… as a continuity of discontinuity (Nishida, 1948, p. 342, after Taguchi & Saigo, 2023)." Similarly, with time, consciousness stays constant yet continuously updates, shaping our social insights, thoughts, and actions (Smitha et al., 2017; Wolff et al., 2019).

Morphisms of consciousness can be viewed from the perspective of the monoid, where the morphism is always from the same object to the same object. "One morphism and another, one morphism and itself, can all be composed because the morphism of a monoid has only one object, which means that all morphisms ('passages') are connected at the same object ('now') because their starting points (domains) and endpoints (codomains) are the same objects (Taguchi & Saigo, 2023)." The flow of consciousness perpetually remains in the same place because it originates from the same object as their domain and returns to it as their codomain. In essence, the dynamic, streaming aspect of the present corresponds to the diverse morphisms within a monoid. In contrast, the static, standing aspect aligns with the unique object of the monoid. This monoid framework effectively captures consciousness' intricate structure, encompassing its ever-changing and constant facets (Figure 2).

Now, I return to the orthogonality question. The brain's spatial data compression formulates an orthogonal sensory temporal hologram (Saaty and Vargas, 2017), representing an orthogonal projection onto the two-dimensional cortical surface (Déli, 2020a,b). In neuroscience, the hippocampus' place cells transform spatial relationships into a temporal projection (Shimazaki, 2020), with the intermediate hippocampal CA1 serving the spatial aspects and the dorsal hippocampal CA1 serving the temporal aspects of episodic memories (Barker et al., 2016), and coordinate collective behavior.

The arrow of time, a fundamental concept in physics, demonstrates a profound connection to the rate of entropy generation (Lucia & Grisolia, 2020). In essence, the arrow of time concept states that spontaneous processes move a system toward equilibrium. Surprisingly, the brain's ability to consistently form endothermic processes entails entropy generation by enhancing the degrees of freedom (Shi et al., 2019; Yang et al., 2019). The brain’s dominant regulation is a predictive organization occurring in the time dimension. Consciousness' predictive nature contrasts general relativity, where gravity is a spatial field. Therefore, intellect is an orthogonal orientation vis-à-vis material structure, where entropy generation represents complexity and order.

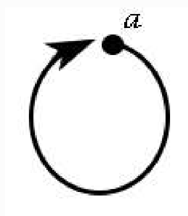

Figure 1. Conscious (*a*) as a monoid The arrow represents the identity morphism.

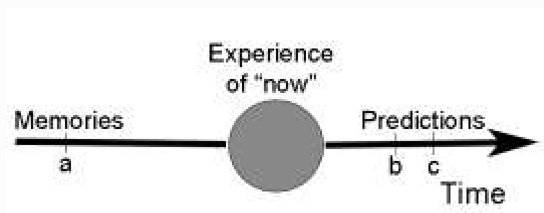

Figure 2. Consciousness as a continuous and discrete temporal flow Memory and expectation materialize at discrete temporal locations marked: time points a,b,c.

5 Conclusions

This work introduces a novel consciousness hypothesis, FMH, offering a radical connection between physics and consciousness. It suggests that the principles governing the smallest building blocks of matter, fermions, are mirrored in the workings of the mind.

FMH builds a bridge between neuroscience, psychology, sociology, and economics by proposing a physical foundation for them all. It posits that intelligent behavior stems from an abstraction of the environment encoded in our synaptic memories. These memories allow us to make predictions, and any discrepancies between those predictions and reality trigger emotions, which motivate us to act. This principle mirrors the fundamental interactions that govern the physical world.

The hypothesis goes further, suggesting a crucial role played by the recurring resting state. This state establishes a distinct, particle-like detachment from the environment while maintaining a consistent psychological identity throughout life.

By adopting the organizing principles of the physical environment, the brain develops a particle-like structure, such as unity and indivisibility, with fermionic traits, such as half spin. "Fermion" and "consciousness" represent the smallest known components of matter and intellect, respectively, shedding light on the physical underpinnings of intellect. Unlocking the physical basis of consciousness, as FMH proposes, could revolutionize our approach to artificial intelligence.

We used the Yoneda lemma to validate the homomorphism (structural similarity) between consciousness and fermions. The emergence of intellect through an endothermic process within memory-forming systems may represent a universal principle in the natural world, offering a pathway for artificial consciousness.

Ultimately, FMH goes beyond scientific exploration. It carries profound philosophical implications by emphasizing the deep connection between consciousness and the physical universe. According to FMH, consciousness isn't a separate entity but intricately woven into the fabric of physical reality.

6 Acknowledgments

I want to express my sincere gratitude to my advisor Dr. Zoltán Kisvárday for the continuous support, patience, and motivation of this Ph.D. study and related research. He aided this thesis research and writing with immense knowledge and wisdom.

I want to thank the reviewers, Dr. Edina Molnár and Dr. Andjelka Hedrih for their sincere, professional, and constructive review and feedbacks, which improved the quality of this dissertation.

7 List of own publications

UNIVERSITY of DEBRECEN

UNIVERSITY AND NATIONAL LIBRARY
UNIVERSITY OF DEBRECEN
H-4002 Egyetem tér 1, Debrecen
Phone: +3652/410-443, email: publikaciok@lib.unideb.hu

Registry number: DEENK/345/2024.PL
Subject: PhD Publication List

Candidate: Éva Katalin Déli
Doctoral School: Doctoral School of Neurosciences
MTMT ID: 10086060

List of publications related to the dissertation

1. **Déli, É. K.**, Peters, J., Kisvárday, Z.: How the Brain Becomes the Mind: Can Thermodynamics Explain the Emergence and Nature of Emotions?
   Entropy. 24 (10), 1498, 2022.
   DOI: http://dx.doi.org/10.3390/e24101498
   IF: 2.7
2. **Déli, É. K.**, Peters, J., Kisvárday, Z.: The thermodynamics of cognition: a mathematical treatment.
   Computational and Structural Biotechnology Journal. 19, 784-793, 2021.
   DOI: http://dx.doi.org/10.1016/j.csbj.2021.01.008
   IF: 6.155

List of other publications

3. **Déli, É. K.**: Can the Fermionic Mind Hypothesis (FMH) Explain Consciousness? The Physics of Selfhood.
   Act. Nerv. Super. 62 (2), 35-47, 2020.
   DOI: http://dx.doi.org/10.1007/s41470-020-00070-4
4. **Déli, É. K.**: Thermodynamic Implications of the Fermionic Mind Hypothesis.
   62 (3), 96-103, 2020.
   DOI: http://dx.doi.org/10.1007/s41470-020-00074-0
5. **Déli, É. K.**, Kisvárday, Z.: The thermodynamic brain and the evolution of intellect: the role of mental energy.
   Cogn Neurodyn. 14 (6), 743-756, 2020.
   DOI: http://dx.doi.org/10.1007/s11571-020-09637-y
   IF: 5.082
6. **Déli, É. K.**, Peters, J. F., Tozzi, A.: The Thermodynamic Analysis of Neural Computation.
   J. Neurosci. Clin. Res. 3 (1), 1-7, 2018.

UNIVERSITY of DEBRECEN

UNIVERSITY AND NATIONAL LIBRARY
UNIVERSITY OF DEBRECEN
H-4002 Egyetem tér 1, Debrecen
Phone: +3652/410-443, email: publikaciok@lib.unideb.hu

7. Tozzi, A., Peters, J., **Déli, É. K.**: Towards plasma-like collisionless trajectories in the brain.
Neurosci. Lett. 662, 105-109, 2018.
DOI: http://dx.doi.org/10.1016/j.neulet.2017.10.016
IF: 2.173

8. **Déli, É. K.**, Tozzi, A., Peters, J.: Relationships between short and fast brain timescales.
Cogn Neurodyn. 11 (6), 539-552, 2017.
DOI: http://dx.doi.org/10.1007/s11571-017-9450-4
IF: 2

9. Peters, J., Tozzi, A., **Déli, É. K.**: Towards Equations for Brain Dynamics and the Concept of Extended Connectome.
SF J. Neuro. Sci. 1 (2), 1-9, 2017.
DOI: http://dx.doi.org/10.23959/sfjn-1000007

10. Tozzi, A., Peters, J., James III., C., **Déli, É. K.**: Multidimensional Chaotic Dynamics and Entropies.
Arch. Inf. Sci. Technol. 1 (1), 10-19, 2016.
DOI: http://dx.doi.org/10.36959/863/755

11. **Déli, É. K.**, Kiss, Z.: Protein Kinase C-Stimulated Formation of Ethanolamine from Phosphatidylethanolamine Involves a Protein Phosphorylation Mechanism: Negative Regulation by p21 Ras Protein.
Arch. Biochem. Biophys. 377 (1), 171-177, 2000.
DOI: http://dx.doi.org/10.1006/abbi.2000.1768
IF: 2.576

Total IF of journals (all publications): 20,686
Total IF of journals (publications related to the dissertation): 8,855

The Candidate's publication data submitted to the iDEa Tudóstér have been validated by DEENK on the basis of the Journal Citation Report (Impact Factor) database.

10 June, 2024