# On motivating irruptions: the need for a multilevel approach at the interface between life and mind

**Ignacio Cea** - Center for Research, Innovation and Creation; and Faculty of Religious Sciences and Philosophy, Temuco Catholic University – Chile; email: igneocj@gmail.com

(This is the manuscript version. The final, published version appeared in ***Adaptive Behavior***, June 2023**:** https://doi.org/10.1177/10597123231184651**)**

**Abstract**

In a recent remarkable paper, Froese (2023) presents his *Irruption Theory* to explain how motivations can make a behavioral difference in motivated activity. In this opinion article, we review the main tenets of Froese's theory, and highlight its difficulty in overcoming the randomness challenge it supposedly solves, i.e. the issue of how adaptive behavior can arise in the face of material underdetermination. To advance our understanding of motivated behavior in line with Froese's approach, we recommend that future work should endorse a multilevel pluralistic approach to causation and explanation in which motivations could genuinely play an irreducible role. Additionally, in line with the life-mind continuity thesis, we suggest that the best place to look for the interplay between motivations and nonmotivational physical, biological, and dynamical factors, may be at the level of the continuous feeling of being an embodied, living organism.

**Keywords:** motivation; motivated behavior; enactivism; mental causation; life-mind continuity

## Irruption Theory and the efficacy of motivations

Froese's *Irruption Theory* (2023) aims to elucidate how motivations, particularly those imbued with value, can make a behavioral difference in *motivated activity*. The latter is defined as "a particular kind of end-directed activity that is driven by the intrinsic preferences of the system in question" (Barrett, 2020b, p. 90). Froese' approach is centered in offering a novel, enactive conceptualization of motivated activity, such that the role of motivations is considered, and they are not left to the status of mere epiphenomena. Froese (2023) refers to this approach as a "motivation-involving" (p. 17) account of motivated activity, which is proposed as advancing an enactive account in line with the libertarian philosophy of freedom and agency. According to it, the motivated activity of an agent is not completely determined by her material embodiment but also irreducibly caused by her motivations. Importantly, although Froese doesn't explicitly identify himself as either a substance or property dualist, he endorses a 'relaxed naturalism' (Hutto & Satne, 2018) where motivations are inescapably mental and irreducible to their material counterparts: "an agent's motivations are taken to constitute a distinctive and intrinsic part of nature in addition to its materiality" (Froese 2023, pp. 15-6). With these commitments in place, Froese's *Irruption Theory* is intended to overcome current limitations of enactive proposals

that also regard aspects of the subjective activity of an agent as irreducibly efficacious (Di Paolo et al., 2010; Fuchs, 2018, 2021; Thompson, 2007; Thompson & Varela, 2001; Varela & Thompson, 2003; see also Meling & Scheidegger, 2023), but which doesn't provide a satisfactory account of how this could be the case.

Froese's main target of criticism is the common enactive strategy of attempting to explain the effect of conscious activity on neural dynamics as a form of downward causation or global-to-local constraint (Thompson & Varela 2001; Thompson 2007). In our opinion, he correctly points out that picturing this causal relationship in terms of global properties of brain activity (e.g. global order parameters, collective variables, etc.) having a determinative influence on the dynamics of the local components of the neural substrate is unsatisfactory. We fully agree that "there is simply no conceptual room for a motivation-involving account if all efficacy is offloaded to a nonmotivational organizing property of the material body" (Froese 2023, p. 8). Although this common enactive view endorses a multilevel view of causal interactions that surely enriches the standard reductionistic picture, as long as the causally determinative global property remains deprived of any intelligible mental attribute, it seems to be an instance of a non-mental global property of a physical substrate having an effect on its local properties. A complementary critique is offered by Cea (2023), who states that the causal relationship between experience and the local activity of its neural substrate arguably does not correspond to a form of global-to-local or downward determinative influence, because these interlevel forms of causation presuppose a whole-parts relation that does not exist between conscious activity and local neural processes.

The main components of Froese's theory are its three axioms, working hypothesis, and three main theses. The three axioms are: 1) *motivational efficacy*: motivations are irreducibly efficacious; 2) *incomplete materiality*: the impossibility of measuring how motivations make a causal difference; and 3) *underdetermined materiality:* the physical underdetermination of motivated behavior. The theory's main working hypothesis, in turn, states that "*The more an agent's embodied activity is motivated, the less that activity is determined by its material basis"* (p. 10, italics in the original). Importantly, this entails that irruptions are associated with increased levels of unpredictability as measured e.g., by the informational property of entropy. Finally, the three theses are: 1) *irruption thesis*: the living body is an incomplete system in the sense that the more motivations are involved in behavior, the more materially underdetermined the latter will be; 2) *scalability thesis:* microscopic material underdetermination is amplified in the living body to a macroscopic level in which behavior can be affected; and 3) *attunement thesis:* the living body responds to amplified irruptions (i.e. material underdetermination due to the involvement of motivations) in a "context-sensitive and adaptive manner" (p. 11), such that the resulting behavior is neither random nor rigid, but appropriate to the circumstances and preferences of the organism.

Very importantly, the third thesis of attunement carries the heavier burden of explaining how this approach to motivated behavior overcomes the well-known "luck objection" to libertarian accounts of free will (McKenna & Pereboom, 2016). In our context of discussion, it is the problem of explaining how the material underdetermination of an agent's actions does not entail that her purported autonomously motivated behavior ends-up

being really a random o lucky outcome of indeterministic processes. In other words, it is the double challenge of explaining how motivated behavior could be both materially underdetermined and non-random. Crucially, Froese (2023) asserts that his irruption theory "solves this double challenge" (p. 9). In the next section, we argue that irruption theory falls short of overcoming this problem.

**The problem of the attunement of scaled up irruptions**

As mentioned, the main task of explaining how motivated behavior is non-random even if it is materially underdetermined is placed in the third *thesis of attunement*. Although Froese (2023) mentions various promising models that could ground and further develop the attunement thesis, such as the embodied approach to habits (Ramírez-Vizcaya & Froese, 2020), implicit body memory (Froese & Izquierdo, 2018), and meta-stable attunement (Bruineberg et al., 2021), he favors the *self-optimization model* of adaptivity and regulation (Froese et al., 2023). The latter proposes that a system, such as a neural network, can spontaneously optimize its internal organization to better satisfy its constraints over time (i.e. solve a "constraint satisfaction" problem). This is achieved through a process of learning and adaptation, involving reversible dynamics, irreversible dynamics, and historicity. Reversible dynamics allow the system to converge into and maintain a state of equilibrium, while irreversible dynamics involve the system transiently entering a far-from-equilibrium state, increasing the chances of converging into a new, different state of equilibrium. Historicity involves the accumulation of correlations of previous states, increasing the chances of those correlations reoccurring. The model also incorporates the concept of 'resets', where a certain number of states in the system are changed or 'reset' to a new state, simulating the system opening up to its environment.

Through this process, the system can find 'deeper' equilibria, or states that satisfy more constraints, enhancing its adaptivity and regulatory capabilities. Froese attributes such importance to this model for the underpinnings of the attunement thesis, that he states that

> Irruption theory in combination with the self-optimization model, therefore, holds potential to provide a novel systems theoretic answer to the main concern raised against libertarian interpretations of motivated activity, i.e., regarding how an agent's behavior can be effective even when it is underdetermined. (Froese 2023, p. 15)

Unfortunately, we think that the self-optimization model as a way to unpack the attunement of motivated behavior as scaled-up irruptions suffers from the same problem that Froese attributes to the common enactive strategy of invoking global order parameters to explain the efficacy of mental processes on neural dynamics and behavior. Trying to understand the efficacy of motivations either as global-to-local determinative influences or as self-optimization *are both ways of conceiving the effect of motivations in nonmotivational terms.* More specifically, in the case of self-optimization, the properties that drive the behavior of the system towards deeper states of equilibria are the system's ability to iterate between equilibrium and far-from-equilibrium dynamics, combined with a form of historicity. However, these capacities follow from the structural, dynamical and functional properties of the system, none of which allude to anything near "motivations as such" to work as they do. In other words, although these systems are remarkable in their capacities

to self-organize towards global coordinated constraint satisfaction, and hence, towards forms of adaptive, context-dependent behavior, they do it entirely in virtue of nonmotivational properties, i.e. their behavior has a full causal explanation in terms of nonmotivational factors. Therefore, without substantive further development, in which an intelligible causal and/or constitutive link between motivations as such and the nonmotivational properties that drive self-optimization is established, it is far from clear how the latter combined with irruption theory can solve the problem of randomness or luck.

A similar problem arises for a view in the theoretical vicinity of Froese's: Barrett's *thermogenic* approach to motivated behavior (Barrett, 2020a, 2020b). Barrett explores the possibility that motivated behavior may be a particular form of a more general physical phenomenon. He considers the thermodynamically-driven end-directed activity common to all living systems and physical dissipative structures, in which "extremal properties" such as the Principle of Maximum Entropy Production (PMEP) (Swenson, 2020; Swenson & Turvey, 1991) purportedly offer "thermodynamic reasons" for the emergence of order at macroscopic levels. Barrett proposes that extremal properties such as PMEP "might be co-opted and directed by organisms as the driving force of motivated behavior" (Barrett 2020b, p. 3). In other words, the motivated behavior of living organisms may be a particular case of a very general physical principle that guides the achievement of macroscopic order in far-from-equilibrium thermodynamical systems.

However, if we are seeking a motivation-involving account of motivated behavior in which motivations as such act as causal influences in end-directed behaviors, we still need an intelligible link between "thermodynamic reasons" and "motivating reasons" such that we could understand how motivations influence behavior based on principles such as PMEP. Moreover, such an account would need to respect the distinction between an inherently psychological and experiential, end-directed motivation like a desire; and a mindless thermodynamic extremizing tendency such as PMEP. Additionally, it would need to show how the influence of the psychological motivation is not causally excluded by the physical extremizing property; nothing less than the causal exclusion problem (Kim, 1998, 2005) applied to motivations under a thermogenic approach.

## Conclusions and recommendations for future work

We conclude that Froese' Irruption Theory, as well as Barrett's related thermogenic approach, are very promising but still incomplete frameworks to think about the interplay between mental and nonmental (e.g. physical and dynamical) properties in motivated activity. More specifically, they seem to excessively downplay the explanatory and causal role of the phenomenological and psychological aspects of motivated activity, privileging, instead, its nonmental underpinnings. We suggest that future motivation-involving accounts of motivated activity, especially from an enactive perspective, should avoid explaining the role of motivations in entirely nonmotivational terms. Instead, we recommend endorsing a multilevel, pluralistic approach to causation and explanation, such as those being recently discussed in theoretical assessments of psychiatry and psychedelic medicine (Aday et al., 2023; Carhart-Harris, 2018; Gauld et al., 2022; Girn et al., 2023; van Elk & Yaden, 2022; Yaden & Griffiths, 2020). Concerning the explanatory and causal role of the phenomenological dimension in psychedelic medicine, for instance, it has been recently

declared that "the case for subjective effects playing a major role in enduring beneficial effects is compelling… underlying neurobiological-based mechanisms are undoubtedly necessary but likely not sufficient to confer full beneficial effects" (Yaden and Griffiths 2022, p. 570). The researchers argue that the subjective, lived character of e.g., a mystical-type experience, most probably plays an irreducible role in determining the long-term effects of psychedelic interventions, in addition to mechanisms at the neurobiological level. More generally, the multifaceted nature of mental disorders and the interventions needed to address them are calling for a pluralistic approach in which "cognitive and phenomenological processes… may be more explanatorily powerful… [we need] research frameworks that productively integrate constructs and examine pharmacological interactions at multiple levels of explanation" (Aftab and Stein, 2022, p. 2). We think that the same is urgently needed for our understanding of motivated behavior. We need to develop a multi-level pluralistic theory in which motivations as such figure in our accounts, without being replaced by nonmotivational factors. In other words, the challenge is to understand how phenomena at multiple levels of description interact, integrate and fit with each other non-reductively, that is, without necessarily trying to explain, e.g. psychological phenomena in terms of non-psychological processes.

Given the centrality of the life-mind continuity thesis for the enactive approach (Thompson 2007) and other influential views (Kirchhoff & Froese, 2017; A. Seth, 2021; A. K. Seth & Tsakiris, 2018; Wiese & Friston, 2021), we further recommend placing special attention to the inherently motivational dimension of the ubiquitous feeling of being alive, recently conceptualized as the *continuous organismic sentience* ("COS") that pervades all experience (Cea & Martínez-Pernía, 2023). This COS would be a pre-reflective, ongoing, background bodily feeling exerting a very basic but allo/homeostatically significant influence on the organism's behavior. Importantly, it would provide the most basic and ubiquitous *motivational direction* regarding the agent's behavior along a two-dimensional space: i) action vs. inaction; ii) perpetuate current state vs. change current state. So, for instance, a pleasant, low-aroused feeling of relaxation would motivate the subject to remain inactive to perpetuate her current homeostatically beneficial state; while an unpleasant, but highly aroused feeling of anxiety would dispose the subject to act to change her current noxious state. Importantly, this basic, but ubiquitous *motivational directionality* might be the most intimate link between life and mind, at least in human consciousness, and hence, the best place to start exploring the interconnections and reciprocal, multilevel influences between nonmotivational physical, biological and dynamical properties, and motivations as such.

## Declaration of Conflicting Interests

The Author declares that there is no conflict of interest

## Funding

***The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article:*** *This work was supported by the ANID-Fondecyt Postdoctoral grant [3210707]; and the Center for Research, Innovation and Creation, Temuco Catholic University, Chile.*

## References


Aday, J. S., Carhart-Harris, R. L., & Woolley, J. D. (2023). Emerging Challenges for Psychedelic Therapy. *JAMA Psychiatry*. https://doi.org/10.1001/jamapsychiatry.2023.0549

Barrett, N. (2020a). Extremal properties and self-preserving behavior. *Adaptive Behavior*, *28*(2), 113–118.

Barrett, N. (2020b). On the nature and origins of cognition as a form of motivated activity. *Adaptive Behavior*, *28*(2), 89–103.

Bruineberg, J., Seifert, L., Rietveld, E., & Kiverstein, J. (2021). Metastable attunement and real-life skilled behavior. *Synthese*, *199*(5–6), 12819–12842.

Carhart-Harris, R. L. (2018). Serotonin, psychedelics and psychiatry. *World Psychiatry*, *17*(3), 358. https://doi.org/10.1002/WPS.20555

Cea, I. (2023). Commentary: Not in the drug, not in the brain: Causality in psychedelic experiences from an enactive perspective. *Under Review*.

Cea, I., & Martínez-Pernía, D. (2023). Continuous Organismic Sentience as the integration of Core Affect and Vitality. *Journal of Consciousness Studies*, *30*(3–4), 7–33. https://doi.org/https://doi.org/10.53765/20512201.30.3.007

Di Paolo, E., Rohde, M., & De Jaegher, H. (2010). Horizons for the enactive mind: Values, social interaction, and play. In *Enaction: Towards a new paradigm for cognitive science*.

Froese, T. (2023). Irruption Theory: A Novel Conceptualization of the Enactive Account of Motivated Activity. *Entropy*, *25*(5), 748.

Froese, T., & Izquierdo, E. J. (2018). A dynamical approach to the phenomenology of body memory: Past interactions can shape present capacities without neuroplasticity. *Journal of Consciousness Studies*, *25*(7–8), 20–46.

Froese, T., Weber, N., Shpurov, I., & Ikegami, T. (2023). From autopoiesis to self-optimization: Toward an enactive model of biological regulation. *BioRxiv*, 2022–2023.

Fuchs, T. (2018). *Ecology of the Brain: The Phenomenology and Biology of the Embodied Mind*. Oxford University Press.

Fuchs, T. (2021). *In defence of the human being: foundational questions of an embodied anthropology*. Oxford University Press.

Gauld, C., Nielsen, K., Job, M., Bottemanne, H., & Dumas, G. (2022). From analytic to synthetic-organizational pluralisms: A pluralistic enactive psychiatry. *Frontiers in Psychiatry*, *13*. https://doi.org/10.3389/FPSYT.2022.981787

Girn, M., Rosas, F. E., Daws, R. E., Gallen, C. L., Gazzaley, A., & Carhart-Harris, R. L. (2023). A complex systems perspective on psychedelic brain action. *Trends in Cognitive Sciences*, *27*(5), 433–445. https://doi.org/10.1016/j.tics.2023.01.003

Hutto, D. D., & Satne, G. (2018). Naturalism in the Goldilocks zone: Wittgenstein's delicate balancing act. In *Wittgenstein and naturalism* (pp. 56–76). Routledge.

Kim, J. (1998). *Mind in a physical world: An essay on the mind-body problem and mental causation*. MIT press.

Kim, J. (2005). *Physicalism, or Something Near Enough*. Princeton University Press. https://doi.org/10.1215/00318108-2007-027

Kirchhoff, M. D., & Froese, T. (2017). Where there is life there is mind: In support of a strong life-mind continuity thesis. *Entropy*, *19*(4), 169.

McKenna, M., & Pereboom, D. (2016). *Free will: A contemporary introduction*. Routledge.

Meling, D., & Scheidegger, M. (2023). Not in the drug, not in the brain: Causality in psychedelic experiences from an enactive perspective. *Frontiers in Psychology*, *14*. https://doi.org/10.3389/fpsyg.2023.1100058

Ramírez-Vizcaya, S., & Froese, T. (2020). Agents of habit: refining the artificial life route to artificial intelligence. *Artificial Life Conference Proceedings 32*, 78–86.

Seth, A. (2021). *Being you: A new science of consciousness*. Penguin.

Seth, A. K., & Tsakiris, M. (2018). Being a Beast Machine: The Somatic Basis of Selfhood. *Trends in Cognitive Sciences*, *22*(11), 969–981. https://doi.org/10.1016/j.tics.2018.08.008

Swenson, R. (2020). The fourth law of thermodynamics (LMEP) and cognition from first principles: commentary on Barrett's "On the nature and origins of cognition as a form of motivated activity." *Adaptive Behavior*, *28*(2), 105–107. https://doi.org/10.1177/1059712319856536

Swenson, R., & Turvey, M. T. (1991). Thermodynamic reasons for perception--action cycles. *Ecological Psychology*, *3*(4), 317–348.

Thompson, E. (2007). *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*. Harvard University Press.

Thompson, E., & Varela, F. (2001). Radical embodiment: Neural dynamics and consciousness. In *Trends in Cognitive Sciences* (Vol. 5, Issue 10, pp. 418–425). https://doi.org/10.1016/S1364-6613(00)01750-2

van Elk, M., & Yaden, D. B. (2022). Pharmacological, neural, and psychological mechanisms underlying psychedelics: A critical review. *Neuroscience & Biobehavioral Reviews*, *140*, 104793. https://doi.org/10.1016/J.NEUBIOREV.2022.104793

Varela, F., & Thompson, E. (2003). Neural synchrony and the unity of mind: A neurophenomenological perspective. In A. Cleeremans (Ed.), *The Unity of Consciousness*. Oxford University Press.

Wiese, W., & Friston, K. J. (2021). Examining the continuity between life and mind: Is there a continuity between autopoietic intentionality and representationality? *Philosophies*, *6*(1), 18.

Yaden, D. B., & Griffiths, R. R. (2020). The subjective effects of psychedelics are necessary for their enduring therapeutic effects. *ACS Pharmacology & Translational Science*, *4*(2), 568–572.