Behavioral and Brain Sciences

cambridge.org/bbs

Metabolic considerations for cognitive modeling

Philipp Haueis[a,b] and David Colaço[c]

[a]Department of Philosophy, Institute for Studies of Science, Universität Bielefeld, Bielefeld, Germany; [b]Institute of Philosophy, Leibniz Universität Hannover, Hannover, Germany and [c]Munich Center for Mathematical Philosophy, Ludwig-Maximilians-Universität München, München, Germany
philipp.haueis@uni-bielefeld.de; https://philipp-haueis.de/
david.colaco@lmu.de; https://sites.google.com/site/colacodavidj/

Target Article

Cite this article: Haueis P and Colaço D. (2026) Metabolic considerations for cognitive modeling. *Behavioral and Brain Sciences* **49**, e453: 1–78. doi:10.1017/S0140525X25103956

Target Article Accepted: 21 October 2025
Target Article Manuscript Online: 18 November 2025
Commentaries Accepted: 15 May 2026

Keywords:
metabolism; cognition; models; constraints; information processing; neural computation; artificial intelligence; energy use

What is Open Peer Commentary? What follows on these pages is known as a Treatment, in which a significant and controversial Target Article is published along with Commentaries (p. 19) and an Authors' Response (p. 71). See bbsonline.org for more information.

Corresponding author:
Philipp Haueis;
Email: philipp.haueis@uni-bielefeld.de

Short Abstract

The human brain makes up 2% of body mass but consumes closer to 20% of the body's energy. Nonetheless, it is more energy-efficient than most computers. Although these facts are well known, models of cognitive capacities and neural systems rarely address metabolic factors. Here, we argue that metabolic considerations should be integrated into cognitive models. Metabolic considerations can be used to evaluate existing and generate novel models. Our account of these considerations offers fresh insights into debates over cognitive models of mental effort, philosophical views of multiple realizability and medium independence, and the comparison between biological and artificial computational systems.

Long Abstract

The human brain makes up just 2% of body mass but consumes closer to 20% of the body's energy. Nonetheless, it is significantly more energy-efficient than most modern computers. Although these facts are well known, models of cognitive capacities rarely account for metabolic factors. In this paper, we argue that metabolic considerations should be integrated into cognitive models. We distinguish two uses of metabolic considerations in modeling. First, metabolic considerations can be used to *evaluate* models. Evaluative metabolic considerations function as explanatory constraints. Metabolism limits which types of computation are possible in biological brains. Further, it structures and guides the flow of information in neural systems. Second, metabolic considerations can be used to *generate* new models. They provide: a starting point for inquiry into the relation between brain structure and information processing, a proof-of-concept that metabolic knowledge is relevant to cognitive modeling, and potential explanations of how a particular type of computation is implemented. Evaluative metabolic considerations allow researchers to prune and partition the space of possible models for a given cognitive capacity or neural system, while generative considerations populate that space with new models. Our account suggests cognitive models should be consistent with the brain's metabolic limits, and modelers should assess how their models fit within these bounds. Our account offers fresh insights into the role of metabolism for cognitive models of mental effort, philosophical views of multiple realization and medium independence, and the comparison of biological and artificial computational systems.

1. Introduction

There are two oft-cited tidbits of information about the brain. First, the brain is a metabolically hungry organ. While making up 2% of human body mass, it consumes about 20% of the body's calories (Mink et al., 1981; Hofman, 1983). Human brains use this much energy because they have a large number of neurons; metabolism scales linearly with neuron number (Herculano-Houzel, 2011), and scales with brain size by a factor of 5/6 (Karbowski, 2007). Second, the brain is far more energy-efficient than any computer engineered so far. While the human brain uses 10–20 Watts, the supercomputer used by the BlueBrain Project, for instance, would use an estimated 2.7 GigaWatts to realistically simulate 1 second of human brain activity (Stiefel & Coggan, 2023). Even the best systems currently available, which are not on par with human cognition, consume multiple orders of magnitude more energy than human brains.

Though these tidbits are often cited by scientists and philosophers, metabolic considerations are rarely explicitly included in models of cognitive capacities or neural systems, and they are often not addressed at all. This is somewhat surprising, since computation is bounded by resources available to brains (van Rooij, 2008; Colombo & Piccinini, 2023), and quantitative estimates of the cost of neural information processing exist (Howarth, Gleeson, & Attwell, 2012;

© The Author(s), 2025. Published by Cambridge University Press

Levy & Calvert, 2021). This raises the question: what can we learn by including metabolism when we model cognition?

In this paper, we argue that *metabolic considerations* can and should be used to evaluate and generate models of cognitive capacities and neural systems. We provide examples of both evaluative and generative insights in this paper, which we use to lay out how appeals to metabolic considerations can be productive in cognitive science and neuroscience. To do this, we first provide a conceptual review of the neuroscientific and philosophical literature, highlighting both (a) why knowledge of metabolism is crucial for explaining cognition, and (b) that empirical and theoretical work is often *too generic* for modeling specific cognitive capacities. Section 3 fills this lacuna by distinguishing between *evaluative* and *generative metabolic considerations* in cognitive modeling. Our account provides concrete examples and general guidelines for including metabolism in model evaluation and generation. Cognitive modelers should use metabolic knowledge to limit their search to biologically plausible candidates and to generate novel hypotheses about how brains process information. Section 4 draws consequences of our view for debates in cognitive science (mental effort and energy use), philosophy (multiple realizability/medium independence), and artificial intelligence (AI) (energy use in biological and artificial systems).

We conclude that cognitive models should be consistent with the metabolic resources that brains have at their disposal. Current metabolic considerations offer some bound on the activities of the brain and mind. Thus, modelers should determine how their models fit into this bound or at least consider whether their models can be understood in terms of this bound. If our arguments are correct, then metabolism provides a powerful and currently underestimated tool for guiding our search for scientific models of the mind.

2. Previous neuroscientific and philosophical work on metabolism

2.1. Neuroscientific background

Neuroscience has amassed a wealth of knowledge about metabolism that is relevant to cognitive modeling. This section provides an overview of this knowledge and assesses its specificity for cognitive modeling. While being the crucial starting point, metabolic knowledge often remains too *generic* for modeling projects. This sets the stage for an account of metabolic considerations that helps researchers go beyond this generic knowledge. It shows how metabolism can constrain and generate models of *specific* cognitive capacities or neural systems.

Philipp Haueis is assistant professor of philosophy and member of the Institute for Studies of Science (ISoS) at Bielefeld University, Germany. He has published numerous articles in general philosophy of science on scientific concepts, context of discovery and pursuit, concept formation and tool development, and in philosophy of cognitive science and neuroscience on cognitive and noncognitive brain functions, neural circuits, cortical hierarchy, exploratory experiments, mechanism discovery, and multiscale modeling. In 2018, he was awarded the price for best dissertation by the faculty of human sciences at Otto-von-Guericke University Magdeburg, Germany.

David Colaço is a postdoctoral researcher in the Munich Center for Mathematical Philosophy at the Ludwig Maximillian University of Munich. He received a DFG individual grant on "three methodological problems in memory science." He has published 20 articles, book chapters, and book reviews, including general philosophy of science works on phenomena, definitions, conceptualization, replication, and artifacts, and philosophy of cognitive science and neuroscience works on integration, engrams, and memory systems. He won the 2024 Philosophy of Memory Essay Prize Award from the Centre for Philosophy of Memory and the 2020 Griffith Award in philosophy from the Southern Society for Philosophy and Psychology.

2.1.1. Metabolic considerations: The very idea

To understand metabolic considerations, let's start with an example. How does the visual system code for decreases in local contrast? One option is to use decreasing firing rates from a high baseline. Maintaining that baseline, however, would use a lot of energy. By contrast, having dedicated "off" cells is metabolically feasible. Koch (2004, p. 477) uses this example to illustrate how "metabolic considerations" can constrain the style of computation. On our account, Koch's example uses metabolism to *evaluate* which models of a computational task are biologically plausible (section 3.1).

The example illustrates that *metabolic considerations* involve reasoning from brain metabolism to models of cognitive capacities or neural systems. *Brain metabolism* includes processes that convert nutrients (e.g., glucose) into energy (e.g., ATP) and processes that remove neurotoxic byproducts of such energy conversion (Dienel, 2014). *Cognitive capacities* are activities or dispositions that enable organisms to process information about bodily states and the sensory environment that is relevant to behavior (Bechtel, 2008). Paradigmatic cognitive capacities include perception, memory, attention, and decision-making.

Models of cognitive capacities are idealized, abstract systems whose elements stand in a relation to assumed elements of the target capacity. Many models are mathematical and posit computational operations to explain cognitive capacities (Moustafa, 2017). There are also dynamical models that use mathematical equations but eschew attributing computations to the brain. Mathematically, dynamical models are often compatible with the assumptions of information theory, as long as neither is interpreted as getting at the fundamental nature of the cognitive system (Beer & Williams, 2015). Similarly, the account developed here focuses less on model ontology but instead highlights whether these models include metabolic considerations *at all*. Do modelers explicitly represent brain metabolism, constraining known models or generating new ones? Metabolic considerations also apply to models of neural systems, that is, systems of functionally and structurally connected neurons and other tissue, which are not of any one cognitive capacity.

Metabolism is traditionally backgrounded in computational modeling. From pioneering texts to many canonical approaches today, computational modelers proceed from the assumption that the brain is energy-efficient. They acknowledge that the brain is much more energy-efficient than supercomputers (Churchland & Seijnowski, 1992; Stiefel & Coggan, 2023). Yet, when it comes to knowledge relevant for modeling cognitive capacities, most textbooks and overview articles background or ignore metabolism (Churchland & Seijnowski 1992, pp. 50–59; Rieke et al., 1999; Dayan & Abbott, 2005; Anderson, 2014; Kriegeskorte & Douglas, 2018). The result is a modeling practice that "neglect[s] completely the energetic aspect of brain functioning, as if neural information processing were for free and performed in some abstract 'mathematical' hyperspace" (Karbowski, 2024, p. 1).[1]

Against this widespread tendency, the concept of metabolic considerations foregrounds knowledge about metabolism when

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

evaluating and generating models of cognitive capacities or neural systems. Modelers should not stop at qualitative descriptions of information processing as "expensive" or "cost-efficient." As the next sections show, neuroscience now provides detailed quantitative knowledge that can be incorporated into cognitive modeling.

2.1.2. Efficiency of brain metabolism: No reason to background

The most generic way of including metabolism in cognitive modeling is assuming that the modeled system is efficient. "Efficiency" can be understood as a maximization of useful output units per unit of input. Yet, that brain metabolism is efficient can mean different things (Theriault et al., 2023, Table 1). Processes in the brain can maximize (1) ATP output per input unit (*energy production efficiency*), (2) energy consumption per second (*energy consumption efficiency*), (3) speed of signal transmission in meter/second (*signal transmission efficiency*), (4) information units sent per time unit (*information rate efficiency*), or (5) information units sent per ATP molecule (*information-energy efficiency*). When discussing how metabolic considerations matter to cognitive modeling, we must distinguish these different senses of "efficiency."

Consider, for instance, (1) energy production efficiency, which arises from evolutionary pressures on brain organization. Organisms with less production-efficient brains must spend more time finding food, leaving them at a reproductive disadvantage (Pontzer, 2015). If "efficiency" just meant *global* energy production efficiency via natural selection, then cognitive models could justifiably background metabolism. Brains would use energy efficiently because the organism as a whole has evolved to produce energy efficiently.

However, there is more to say about efficiency (Padamsey & Rochefort, 2023). To begin, global energy production is just one application of production efficiency (1). While global glucose brain metabolism produces energy close to the theoretical maximum, its efficiency drops when organisms execute tasks (Dienel, 2019, p. 950). On the local level, however, brain metabolism is much less production-efficient. Alongside oxidative metabolism (~30 ATP per glucose molecule) the brain continually uses glycolysis, which typically is thought of as a process which only occurs when oxygen supply is deficient. Yet, in the oxygen-rich brain this process occurs as *aerobic* glycolysis, which produces 2 ATP per glucose (section 3.2.2). Furthermore, action potentials are less (2) consumption-efficient but more (3) transmission-efficient than chemical diffusion (Sengupta Laughlin & Niven, 2014). Finally, thick axons are more (4) rate-efficient than thin axons, whereas for (5) information-energy efficiency, thin axons outperform thick axons. This quick survey reveals that backgrounding the brain's metabolism "because its efficient" hides potentially valuable information for cognitive modeling.

When models make efficiency considerations explicit, they often remain generic. Two prominent examples illustrate this point. One of the first computational models to incorporate metabolism found that model neurons do not maximize the number of distinct patterns they represent (Levy & Baxter, 1996). Rather, they maximize the amount of represented information per energy unit – they are information-energy-efficient. It does not matter for this model *what* the neuronal firing patterns represent. The finding is a generic statement about information-energy-efficient coding, not a specific constraint or guidance to model cognitive capacities.

The same is true for the role of efficiency in *design principles* of neural systems, such as "*send only what is needed; send at the lowest acceptable rate; minimize wire, that is, length of all neural processes*" (Sterling and Laughlin 2015, p. 54f., emphasis in original). These principles are derived from empirical regularities, such as the law of diminishing returns for the information rate in axons (Figure 1).[2]

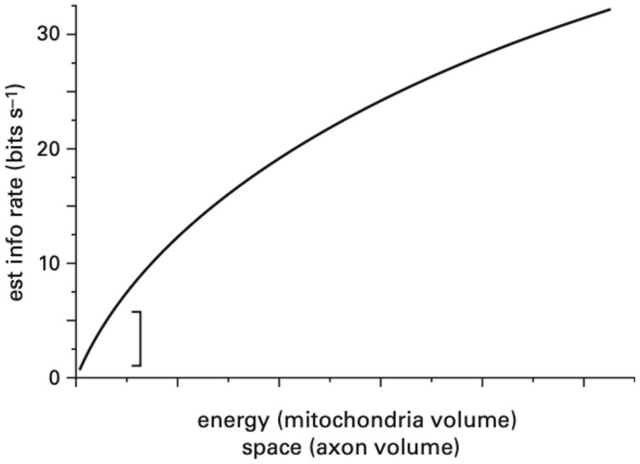


Figure 1. The relation of the information rate in myelinated axons (y-axis) to their ability of producing energy in mitochondria (x-axis). If the axons get thicker than 0.7 µm, the information rate rises slower than the energy production rate. Reproduced from Sterling and Laughlin (2015), Figure 3.6. The diagram is based on empirical data from Perge, Koch, Miller, Sterling, and Balasubramanian (2009).

Design principles generically relate information rate efficiency and energy consumption. However, they cannot be directly applied to models of neural systems with specific information rates underlying cognitive capacities. To do so, modelers must specify what information is needed to execute that particular capacity, or what rate is acceptable for that specific system. Design principles provide a useful starting point but do not themselves constitute metabolic considerations needed for cognitive modeling.

2.1.3. Global energy budgets: Energy for cognition is limited

Global energy budgets are more specific than efficiency considerations, as they distinguish whether energy is spent for signaling versus non-signaling or "housekeeping functions" (Atwell & Laughlin, 2001). This distinction puts an *upper bound* on the energy that brains can allocate to cognitive capacities. Despite a lack of a mark of the cognitive (Allen, 2017), neuroscientists and philosophers commonly characterize cognition as information processing relevant to the behavior of the organism (de Charms & Zador, 2000; Bechtel, 2008; Haueis, 2018). Since only signaling functions are related to information processing, only that part of the energy budget can be used to fuel cognitive processes.[3] Figure 2 shows that recently, more direct estimates of non-signaling functions reduced the estimated energy available for information processing from 75% (Figure 2A) to 60% (Figure 2B). Energy budgets thus provide global metabolic constraints for cognitive modeling, since energy availability limits the amount of information processing brains can perform (section 3.1.1).

Being a global upper bound, however, energy budgets provide no quantitative metabolic considerations to model specific cognitive capacities. Unless one is evaluating the energy cost of whole-brain models (Ritter et al., 2013; Meier et al., 2022), these budgets provide limited guidance by estimating relative costs of neural activities. The budget in Figure 2A, for example, estimates that among signaling functions, generating action potentials is cheap

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

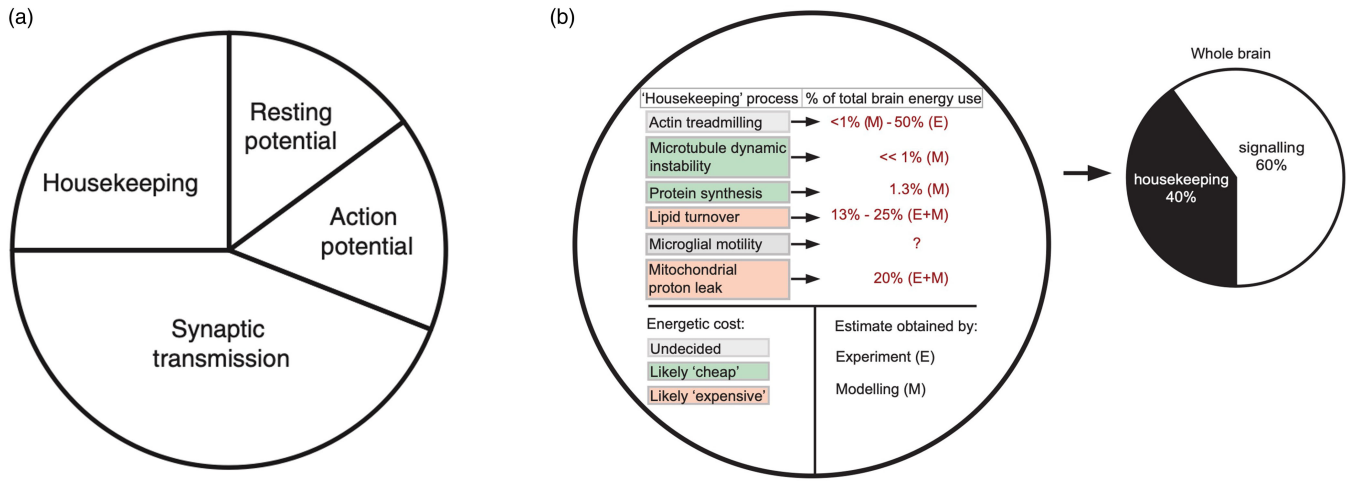


Figure 2. Two recent energy budgets of signaling and non-signaling ("housekeeping") functions in the brain (adapted from Howarth et al. 2012, Figure. 1B and Engl and Atwell 2015, Figure 5).

(16%) compared to synaptic transmission (44%). So, in models of a specific cognitive capacity that include both processes, their energy costs should differ by a factor of 3. Nonetheless, global energy budgets are generic: they do not provide concrete energy costs for neural activities underlying specific cognitive capacities.

2.1.4. Numerical values for metabolic cost of information processing

Even more specific than energy budgets are numerical values of how much energy information processing tasks cost. These numerical values are already used to evaluate and generate models of cognitive capacities, such as memory, or specific computations, such as prediction error (sections 3.1.2 and 3.2.2). Such numbers have been calculated, for example, for the cost of *transmitting information* across synapses or axons (Harris, Jolivet, & Attwell, 2012; Sengupta et al., 2013), and for *storing memories* in synapses (Karbowski, 2019; Karbowski & Urban, 2024). These values are relevant for any model that includes neural information transmission and storage as a basis for executing cognitive capacities.

To estimate numerical values, researchers (1) obtain empirical or simulation data of the metabolic cost of a neural activity, (2) choose a theoretical measure of information processing, and (3) estimate the cost of 1 bit of information being processed by the activity (in ATP/bit). Consider first the cost of information transmission at the synapse (Harris et al., 2012, p. 762). It combines the (1) value of 164000 ATP molecules per synaptic vesicle release (Laughlin & Atwell, 2001) with the measure of (2) mutual information, which specifies how much the patterns of postsynaptic currents tell us about presynaptic action potential inputs.[4] The result is that (3) transmitting information across synapses costs 24000 ATP/bit.

Consider next the *cost of storing memories*, which is used to evaluate synaptic memory models (section 3.1.2). In synaptic models of memory, information putatively is stored in synapses, which are shaped by experience-dependent plasticity processes. Yet, how can synapses act as long-term memory storage sites if they constantly rebuild their structure? The molecular process of *autophosphorylation* provides a solution because it allows a system to store information despite molecular turnover (Lisman, 1985; Koch, 2004, p. 455). Because synapses contain CaMKII molecules, which can autophosphorylate, they can act as long-term memory storage sites. In the model discussed in section 3.1.2, the metabolic cost of this process and another protein (PSD-95) during synaptic long-term potentiation was calculated at 31000–91000 ATP/min (Karbowski, 2019). This amounts to step (1) of estimating the cost of information processing, but (2) information theoretic measures and (3) values for ATP/bit are missing in this study. In follow-up work using the same nonequilibrium thermodynamics approach, Karbowski and Urban (2024) specify (2) *information gain* as the theoretical measure.[5] As a result, they estimate (3) the cost of storing information for 10 synapses at 5 million ATP/bit.[6]

Finally, there is the *cost of information transmission in axons*, which is used to generate new models of neural systems underlying specific computations in the brain (section 3.2.2). To estimate energy costs in model axons, Sengupta et al. (2013) calculated trial-averaged ATP molecules/second for fueling the Na^+/K^+ pump that must be restored to transmit action potentials along the axon. To (2) measure information, they subtracted the noise entropy (average of probability distribution for each word occurring at any time) from the total entropy (information capacity of the spike train). To calculate the cost of information transmission, they (3) plotted the ATP/bit against axon diameter, which is shown in Figure 3.

In Figure 3, axon model compartments on the blue line spend 71500 ATP/bit if they are 0.56 μm in diameter. This cost rises tenfold for 5.6 μm thick axons, showing that thin axons are more information energy-efficient (Theriault et al., 2023, Table 1).

2.1.5. Beyond generic uses of metabolism in cognitive modeling

The previous sections sort metabolic knowledge along a continuum of generic to specific when it comes to modeling cognitive capacities and neural systems. Backgrounding metabolism "because it's efficient" is the most generic. The various senses of "efficiency" challenge that this attitude is appropriate. Even when specific senses of efficiency are addressed, the models usually are generic. They only specify certain kinds of codes as efficient or certain design principles that equally apply to many brain activities or systems. Energy budgets are more specific because they distinguish signaling and housekeeping functions, but they only put an upper energy limit on the sum of cognitive capacities. The most specific are numerical costs of specific information processing

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

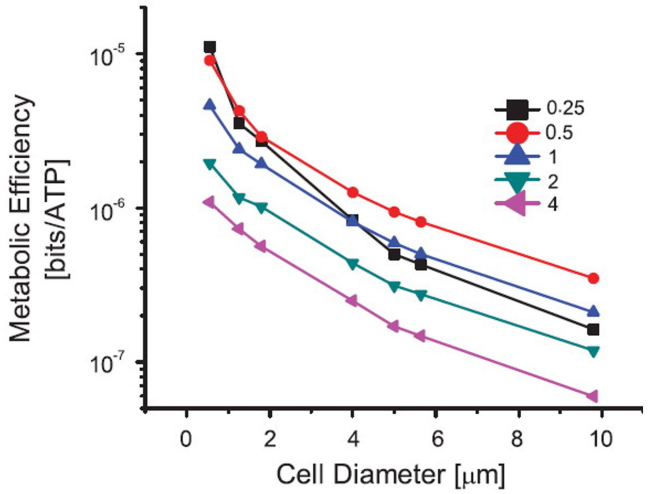


Figure 3. The relation between information-energy efficiency (bits/ATP) and axon diameter in single-compartment models with different ion channel densities (where 1 represents the channel densities of Hodgkin and Huxley's original model of the giant squid axon). Adapted from Sengupta et al. (2013), Figure 3B.

activities by particular neural entities. Yet, these numbers often are extrapolated from animal studies and are not specific to a neural system of cognitive significance. To wit: it is one thing to know the cost of sending a bit of information across the synapse. It is another to know the cost of computing the Laplacian convolved with the Gaussian in the early visual system (to use the classic example of Marr, 1982, p. 337).

The account of metabolic considerations in section 3 shows how these limitations can be overcome once researchers use metabolic knowledge to constrain and generate models of cognitive capacities or neural systems. Rather than treating it as generic or mere background knowledge, metabolic considerations can and should play an active role in cognitive modeling. Due to disciplinary compartmentalization, however, metabolic knowledge from neurochemistry is often not related to information processing or computation. Thus, cognitive and computational neuroscientists may perceive this knowledge as "merely" metabolic or "noncognitive" and thus irrelevant (see section 2.2). Metabolic knowledge, even if relevant and quantitative, may be *formatted in the wrong way* to be useable in cognitive modeling projects. For example: while ATP/bit can be used to evaluate existing or generate new cognitive models, ATP/second cannot, as it is not directly related to information processing in cognitive models.

These issues highlight three desiderata for metabolic knowledge to act as full-blown considerations in cognitive models. First, this knowledge should be *specific* to a particular computational operation (e.g., contrast detection, computing prediction error), a cognitive capacity (e.g., early visual processing, memory), or neural system (e.g., hippocampus, dorsal visual stream). The *relevance* of the metabolic process to the computational operations underlying a cognitive capacity should also be made explicit (e.g., the costs of storing or transmitting information). Finally, metabolic knowledge should be *legible*, that is, its format needs to be compatible with how the cognitive model is formulated (Colaço, 2025).

2.2. Philosophical work on metabolism

While most philosophers of mind and cognition ignore metabolism, a few recognize its importance. This section summarizes this work and reveals a tendency that parallels the science above. While philosophers provide *general* arguments that scientists should consider metabolism (sections 2.2.1 and 2.2.2), they do not address how metabolic considerations can be concretely included in cognitive modeling (section 2.2.3). This provides further motivation that a detailed, philosophically informed account of metabolic considerations in cognitive modeling is needed.

2.2.1. General philosophical considerations of metabolism

There are agreed-upon points among the few philosophers who discuss metabolism and its relation to cognition. At its broadest, they address that, in our terms, the energy intensive but consumption-efficient nature of brains raises questions of how the brain's energy use relates to the activities it performs. For example, in an introduction to the computational theory of mind, Piccinini and Colombo declare:

> Since computing systems consume energy and operate under thermodynamic and metabolic constraints, they must be efficient if they are to stick around and generate offspring (Colombo & Piccinini, 2023, p. 32).

Likewise, philosophers emphasize the two tidbits mentioned in section 1. Take Eliasmith:

> A typical brain weighs between 1 and 2 kg and comprises only 2% of our body weight. Nevertheless, brains account for about 25% of the energy used by the body. This is especially surprising when you consider the serious energy demands of muscles, which must do actual physical work (Eliasmith, 2013, p. 34).

He also notes:

> Brains are absolutely fantastic devices. For one, they are incredibly efficient. Brains consume only about 20 W of power – the equivalent of a compact fluorescent light bulb (ibid.).

The metabolism-friendly philosophers highlight that energy use is a biological basis for cognition as "without energy, there is no thinking and no computation" (Thagard, 2022, p. 70). Likewise, they take brain metabolism and information processing to be "tightly intertwined," as "in the brains of creatures like us, you cannot have one without the other" (Cao, 2022, p. 506). These philosophers take metabolism to be fundamentally important for cognition and its study.

2.2.2. Metabolism: Tightly connected or enabling cognition?

Philosophers who defend the importance of metabolism debate *how* it relates to cognition. Some accounts challenge a sharp distinction between the two. Like Cao (2022), Godfrey-Smith argues that cognition and metabolism "are tightly connected in organisms like us," and the "line between the "information processing" side of human brain activity and the metabolic side is porous" (2016, p. 493). On Godfrey-Smith's account, all metabolic systems connected to the environment exhibit cognitive, or at least "proto-cognitive," behaviors. For instance, they sense and respond to environmental stimuli in variable and plastic ways. This connection between metabolism and (proto-)cognition complicates any demarcation between the two. The connection also seems to challenge the feasibility of comparing biological and artificial systems, given their match in information processing but their grave mismatch in energy use (section 4.3).

Processes that seem "merely" metabolic might serve computational roles. For instance, functional hyperemia, that is, excessive blood flow from a metabolic standpoint, may have activity-dependent neuro-modulatory function and hence be computationally relevant (Moore & Cao, 2008). This example suggests that

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

we might reserve judgment about whether processes are "merely" metabolic or "properly" cognitive. Likewise, metabolism can help evaluate the functions ascribed to biological systems. Metabolism might "anchor" the information processing activities of the brain (Cao, 2022, p. 506). Even if metabolism and cognition can be distinguished, their intertwined nature entails that the former constrains the latter (see section 3.1.1).

Metabolism might be necessary for cognitive phenomena, but unlike the mechanistic components often included in explanatory models. For instance, Klein (2018) identifies *resources* which are *transformed* during the modeled process, *not individually important* to the system in which they are embedded, and *causally promiscuous* in their interaction with many different components. Resources also cannot be easily substituted; their *realization* is relevant to the system. *Mechanical parts*, by contrast, persist over the timescale of a process, are individually important to the system, and are causally conservative in their interaction, and mechanistic models are indifferent to their composition (Klein, 2018, Table 1). A toy example of this distinction is a car's engine: a resource is gasoline, while a mechanical part is a check valve. Metabolism would equally count as a resource on this account (Klein, 2022).

There are also etiological dimensions of metabolism. For instance, Godfrey-Smith's account stems from his idea that metabolism, by achieving internal maintenance and coordination despite environmental variability, might be an evolutionary precursor of cognition. Likewise, metabolism arguably plays a role in major transitions in cognitive evolution, where resource efficiency seems to act as a main driver: "the lineage may not have gained any new cognitive functions at all. Rather it was able to do the same thing more efficiently" (Barron, Halina, & Klein, 2023, p. 6). On this account, the evolution of brain structure and function is greatly informed by the use of metabolic resources.

2.2.3. From philosophical interests to cognitive modeling

The preceding shows that a growing minority of philosophers have taken a "metabolic turn," arguing that metabolism is crucial for an adequate understanding of cognition.[7] The focus of these debates is whether "mere" metabolic and "proper" cognitive processes can be distinguished. Those who answer "no" argue that metabolism fundamentally shapes the evolutionary and developmental trajectory of the brain, and they appeal to neural design principles to show that computational processes and energy consumption are tightly related. These considerations speak against the tendency of most philosophers to treat metabolic considerations as mere support for cognition (Chirimuuta, 2024, p. 92), subordinate "implementational-level" details, or background conditions. This tendency is motivated by multiple realizability of mental function (Polger & Shapiro, 2016) and medium independence of computation (Piccinini, 2020), even though these notions *per se* do not speak against including metabolism in cognitive models (section 4.2). By contrast, the claims in this section suggest that we cannot make sense of cognition in the absence of metabolism or confidently judge that a given process is "merely" metabolic. Furthermore, comparing behavior of biological and artificial systems without considering energy use can be misleading (see section 4.3).

Nonetheless, even philosophers who are interested in metabolic considerations do not detail *how* these considerations can be incorporated into cognitive modeling. These accounts justify why the mind sciences should care about metabolism, but they provide little concrete guidance for the practicing scientist. One potential reason for this lack is that philosophers largely address the topic of metabolism at a different level of analysis. At most, they generally consider theoretical *lower bounds* on energy requirements for physical computation, such as Anderson and Piccinini's brief discussion of Landauer's principle ((2024) Anderson & Piccinini, 2024, p. 107). These generalist appeals are a first step, but they are insufficient for incorporating metabolic considerations into models.

Our account of metabolic considerations builds upon these philosophical insights by showing how metabolism can inform modeling practices. Because this account treats incorporation as a modeling problem, it is compatible with metabolic processes being an enabling condition for cognition or being proper cognitive factors. The ontological character of metabolic considerations is secondary to their role in evaluating and generating models of cognitive capacities or neural systems.

3. Using metabolism to evaluate and generate cognitive models

The previous section showed that the current science and philosophy of metabolism is often too generic to guide modeling specific cognitive capacities or neural systems. Our account of *evaluative* and *generative* metabolic considerations rectifies this issue by drawing on concepts from philosophy of science (sections 3.1.1 and 3.2.1). Examples from neuroscientific practice show how metabolism can be included when evaluating or generating model structure (sections 3.1.2 and 3.2.2), dynamics (section 3.1.3), and computational rules of cognitive models (section 3.2.2). Each of the sections closes with general guidelines that point out how metabolism prunes, partitions, or populates the space of possible models (sections 3.1.4 and 3.2.4).

3.1. Evaluative metabolic considerations

3.1.1. General characterization

Evaluative metabolic considerations address whether a model of a cognitive capacity or neural system is metabolically viable. They can favor a certain model while deeming other models metabolically implausible or impossible. Incorporating metabolic considerations thus can reveal computational deficiencies in the model.

Metabolism can function as evidence to prune or partition the space of possible models, for example, mechanistic models (Craver & Darden, 2013, chs. 7–8). Models of cognitive mechanisms specify parts and operations that manipulate, store, or transmit information (Bechtel, 2008; Kriegeskorte & Douglas, 2018) or the computations that a neurocognitive mechanism performs (Koch, 2004; Piccinini, 2020). If a model represents a mechanism whose operations are inconsistent with evidence about brain metabolism or energy use, researchers should eliminate it from the space of possible models. As resources, metabolic considerations can likewise "serve to invalidate (or at least cast suspicion upon) certain types of mechanistic explanation" (Klein, 2018, p. 8–9). This idea generalizes, as metabolism can be used to evaluate the plausibility of any model of cognitive capacities, mechanistic or otherwise (section 3.1.4).

Evaluative metabolic considerations also can provide *explanatory constraints* (Ross, 2023). Metabolism constrains cognitive modeling because it contrasts possible with impossible computational operations that underlie cognitive capacities. Metabolism limits at which rates neural systems can process information or which computations are possible to implement (Sengupta & Stemmler, 2014). Explanatory constrains are

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

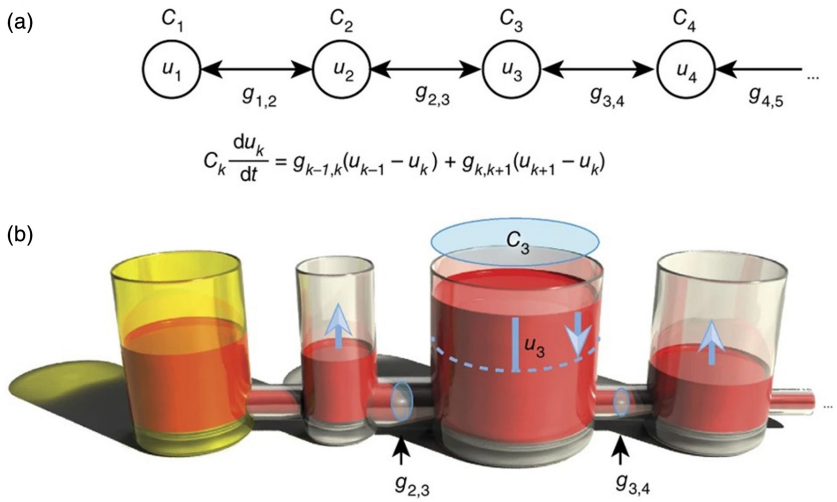


Figure 4. Linear cascade model of synaptic memory consolidation. Adapted from Benna and Fusi (2016) Figure 1.

construed as *external* to the process they limit (Ross, 2023). Information processing and computation in brains require energy, but the metabolic processes producing this constraining resource are generally not viewed as operative in the mechanisms that compute or process information (Levy & Baxter, 1996; Lennie, 2003; Sengupta et al., 2013, 2014; Sterling & Laughlin, 2015; Engl & Atwell, 2015).

Like other explanatory constraints, metabolism is more *fixed* than other factors in cognitive modeling. The global energy availability is fixed by evolution, which favors brains that produce and use energy efficiently (section 2.1.2). Brain metabolism also does not vary substantially between task and rest (Clarke & Sokoloff, 1999; Raichle 2015). And while short-term increases to local energy are possible, they require diverting energy from other regions (section 4.2). Further, as constraints, metabolic processes *structure* or *guide* information processing and computation, but they do not trigger it. In cognitive models, information processing is triggered by external inputs or internal information-bearing states of the system (de Charms & Zador, 2000; Bechtel, 2008; Kriegeskorte & Douglas, 2018).

Metabolic explanatory constraints are similar to but more specific than resource rational analyses. This approach uses metabolism as one type of resource to select "the best biologically feasible mind out of the infinite set of bounded rational minds" (Lieder & Griffiths, 2020, p. 4). Resource rational analysis is evaluative because it uses resources, limitations, constraints, and costs to improve cognitive modeling. However, one problem of this approach that it lacks an account that unifies metabolism and other computational or informational constraints not directly related to metabolism (Ma & Woodward, 2020; Theriault et al., 2020). Our account avoids this difficulty by focusing directly on metabolic resources while not denying that a unified account can be given. It also distinguishes evaluative from generative roles of metabolism, which to our eyes is not included in resource rational analysis.

3.1.2. Core example: Evaluating mathematical models of synaptic memory

A clear example of evaluative metabolic considerations is the comparison between mathematical models of synaptic memory (Benna & Fusi, 2016; Karbowski, 2019). These models often face issues in adequately balancing the different synaptic activities that putatively underwrite memory formation, consolidation, storage, and retrieval. A crucial problem these models must resolve is the "plasticity–rigidity dilemma," that is, the problem that plastic synapses are good at storing new memories but bad at keeping old ones. One model put forward to solve this problem is a *linear cascade model* (Benna & Fusi, 2016), which represents memory storage as a chain of biochemical processes with different timescales.

The model represented in Figure 4a works analogously to a set of beakers that are filled with liquid, interconnected with tubes, and varying in diameter. The first beaker represents the strength of synaptic weight and is the most plastic one, while all others represent biochemical processes. The last beaker is connected to a reservoir that is always at equilibrium. Long-term potentiation is represented by adding liquid, while long-term depression is represented by removing liquid. Because of bidirectional interactions, a beaker's liquid level tends to settle into equilibrium at the level of its nearest neighbors. This allows for flexibility of memory because "a single synaptic modification is remembered as long as the liquid levels remain significantly different from equilibrium" (Benna & Fusi, 2016, p. 1699).

This cascade model allows that branches can split off or merge randomly from the initial linear chain model and shows that when properly tuned, both have the same memory performance. Therefore, both the linear and the complex branching models solve the plasticity–rigidity dilemma by storing the memory signal in fast variables first (e.g., synaptic strength) and then transferring it to increasingly slower variables (e.g., biochemical processes).

Though Karbowski also constructs a cascade model of synaptic learning and memory, he criticizes that almost all previous modeling "neglects the energetic aspect" of storing memories in synapses (2019, p. 1473). The crucial difference between them is Karbowski's inclusion of *cyclic motifs* (Figure 5B, C).

Whereas the complex branching in Benna and Fusi's model might incidentally contain cyclic motifs, Karbowski claims that they are necessary. In linear cascade models, as Benna and Fusi (2016) demonstrate, all biochemical processes become balanced over time. Metabolically speaking, balance means that synapses do not exchange any energy with their environment. Karbowski argues:

> Such a condition corresponds to a "thermal death" when no biological function can be performed, and thus it cannot represent a baseline or resting synaptic state, during which the synapse keeps track of its prior plasticity events, i.e., keeps their memory (Karbowski, 2019, p. 1481).

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

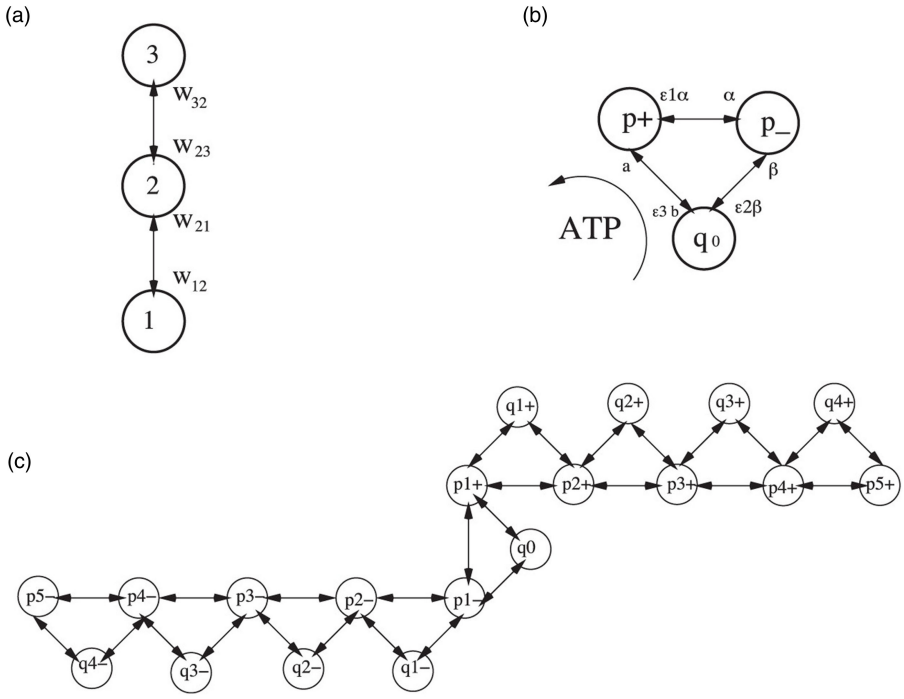


Figure 5. Noncyclical (A) and cyclical models (B, C) of synaptic memory consolidation. Adapted from Karbowski (2019, Figure 1).

While linear cascade models may be able to solve the plasticity–rigidity dilemma, they cannot solve the problem of thermal death because they assume that a synapse at equilibrium would be alive and capable of storing memories. Once metabolism is considered, this equilibrium state becomes unrealistic, as it would correspond to a dead synapse. Arguably, models of synaptic memory should avoid explaining how synapses keep their memories by appealing to states that no living synapse can enter. This is a deficiency of these sorts of models.

To solve this problem, cyclic motifs can be added to cascade models (Karbowski, 2019). Cyclically connected biochemical processes with bidirectional interactions remain at a non-equilibrium steady state, which requires energy and thus prevents the system from thermal death. For synapses, the most important cyclical motifs are phosphorylation–dephosphorylation cycles of proteins, such as actin or AMPA receptors. These cycles are the energetically dominant factor during the initial phase of LTP, when a memory is learned. To prevent equilibrium, one factor in the cyclic motif must be much smaller than the others. Synapses fulfill this condition because phosphorylation is much cheaper than dephosphorylation.

When comparing these models of synaptic memory, metabolic considerations play an evidential and constraining role. The demand that synapses must keep exchanging metabolites with their environment *prunes* the space of possible models underlying synaptic memory. It excludes acyclical models that allow biochemical cascades to reach equilibrium while including cyclical models in which balance is prevented. This evidential function fulfills the first feature of explanatory constraints: it is impossible to store a memory in a biochemically balanced synapse, as the synapse would face thermal death.

These metabolic constraints are *external* to the parts and operations underlying synaptic memory consolidation. Figure 5B shows that ATP hydrolysis powers the transition from a substrate complex q_0 to an activated protein p_+ (such as phosphorylation). But the arrow in that figure indicates that the process is external to the synaptic memory mechanism. ATP hydrolysis can power other processes, whose values remain fixed while the synapse consolidates a memory (Karbowski, 2019, Table 1). These costs *guide* and *structure* the outcome of synaptic memory consolidation. For example, since the lifetime of a memory trace grows faster than its metabolic cost, longer memories are relatively cheaper than shorter ones (Karbowski, 2019, p. 1483). The cost of memories structures the consolidation process because it favors longer memories.

In sum, metabolic considerations show that synaptic memory storage should not be modeled as a deviation from the equilibrium state but instead must be understood as a perturbation that increases energy dissipation from the nonequilibrium steady state. Cyclical models incorporate the constraint that biological systems must avoid thermal death, whereas linear cascade models do not. This makes the latter deficient in accounting for memory storage, and this deficiency ties to the model's lack of fit with these metabolic considerations.

3.1.3. Supporting example: Evaluating neuronal model dynamics

While the core example shows how metabolic considerations may categorically evaluate cognitive models as possible or impossible, researchers often seek to gradually improve their models by adding evidential constraints. This section discusses such an example of gradually improving the dynamics of a model by including metabolic constraints.

The example is the inclusion of metabolism in a leaky integrate-and-fire (LIF) model (Jaras et al., 2021). A LIF model neuron generates an action potential once it integrates enough synaptic input to push the membrane potential beyond a fixed threshold. The membrane potential itself dissipates over time because of a leak conductance. After the action potential, a refractory period describes the time to restore the resting potential. Available energy influences the dynamic of this restoration process, among other adaptation behaviors. Yet, "the model and its consecutive

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

extensions neglect metabolic dependence in their dynamic" (Jaras et al., 2021, p. 5250).

The *energy-dependent* LIF model, or EDLIF model, remedies this neglect. Crucially, spike dynamics are energy-dependent, as an ATP deficiency temporarily disables the Na^2-/K^+ pump after an action potential occurs. The result is slight depolarization, which makes stimulus causing another action potential more likely. The EDLIF model captures that ATP differences make a difference to the reset value of the membrane potential. When compared to the standard LIF model, EDLIF reproduces a greater dynamic range of action potentials in line with available evidence. It also reproduces a linear relationship between ATP consumption and firing known from earlier studies.[8]

The EDLIF model incorporates metabolism as explanatory constraint. First, ATP concentration is relatively *fixed* because homeostatic mechanisms balance neuronal energy supply and consumption (Jaras et al., 2021, p. 5251). Second, ATP concentration *structures* neural dynamics because it influences the probability of generating an action potential, but it does not trigger action potentials.[9] Finally, the EDLIF model does not contrast possible with impossible outcomes, but rather *gradually improves* the existing, energetically unconstrained LIF model. Evaluating a model's metabolic plausibility can thus be a gradual affair. That is, evaluative metabolic considerations can *partition* the space of possible models.

3.1.4. Guidelines how to use metabolism evaluatively

We can combine the general characterization with the details of the examples to formulate general guidelines for evaluative metabolic considerations in cognitive modeling (Box 1). We regard it as a strength if the guidelines strike modelers as sensible and general. If they are sensible, then they should hopefully lead to the widespread inclusion of metabolism in model evaluation. If they are general, then they should be equally applicable to models of different cognitive capacities, of different biological detail, of different species, etc.

At the center of evaluative metabolic considerations is the comparison of models in terms of their adequacy. To do so, researchers need to specify which models should be evaluated (E1). The common basis are models of the same cognitive capacity or the same neural system. In both cases, the models to be compared must be models of the same *target*. Ideally, there should be at least two models that differ by at least one component (an entity, operation, or organizational feature of the target system being modeled). With these models in hand, researchers should specify which metabolic processes or facts about energy consumption should constrain the models (E2). Researchers should choose metabolic processes or facts that are *relevant* to the cognitive capacity or neural system. This differs from case to case. Exemplary processes are oxidative phosphorylation, aerobic glycolysis, or fueling the Na^2-/K^+ pump, while exemplary facts are global energy budgets, the costs of synaptic transmission, or that metabolically active systems stay at nonequilibrium. As we have shown, many of these processes and facts have been uncovered, and they have been linked to paradigmatic cognitive capacities and neural systems.

In the next step, said metabolic factors are used to evaluate models and their components (E3). The relevance of the chosen processes or facts is determined by their utility for constraining the model. They should allow researchers to distinguish between metabolically possible or impossible models or components. This is exemplified by the cyclic and noncyclic motifs in synaptic learning models of the core example. Alternatively, they should allow researchers to assess the plausibility of models or components. This is exemplified by spike dynamics in LIF models of the supporting example. Importantly, this guideline does not entail that models that do not explicitly represent metabolic considerations are necessarily deficient. Models might adequately fit metabolic considerations even if they do not explicitly represent them. In our examples, deficiencies are revealed, but this not a guarantee. Nonetheless, researchers can determine whether a model fits these considerations via this guideline. For this reason, requiring evaluative metabolic considerations is not equivalent to demanding that models must always explicitly represent metabolism.

Through this evaluation, researchers should use metabolism to *prune* or *partition* the space of possible models (E4). Models of cognitive capacities or neural systems should be constrained by evidence about the elements of the target system. Evidence about metabolism and energy consumption can be used to rule out hypothesized structures of the elements that are impossible to implement in biological brains. This prunes the space of possible models underwriting a cognitive capacity. In addition, metabolic considerations can be used to assess the plausibility of a model. This amounts to a partition *within* the space of possible models into more or less plausible ones given available evidence. Finally, researchers should use these considerations to evaluate how energy availability and use structures and guides the flow of information in the system they target (E5). Even when assumed to be external to the computational operations they enable, metabolic constraints can limit the flow of information in cognitive systems. Researchers can therefore use metabolism to assess when the brain uses energetically expensive or cheap information channels (section 3.1.2) or distinguish between informational states that are energetically easy or difficult to undertake (section 4.1).

In sum, our account of evaluative metabolic considerations points out their role as explanatory constraints, which can be used to prune or partition the space of possible models. Although already used in current modeling practice, their use is not widespread. The benefit of our account is that it shows the logic of evaluative metabolic considerations, helping modelers to use them productively in the future.

Box 1. Guidelines for evaluative metabolic considerations

Guidelines for evaluative metabolic considerations

E1: Specify which models should be evaluated.

E2: Specify which metabolic processes or facts about neural energy consumption should constrain the models.

E3: Use metabolic considerations to evaluate model components.

E4: Use metabolic considerations to prune or partition the space of possible models.

E5: Use metabolic considerations to evaluate how energy availability structures and guides the flow of information in target systems.

3.2. Generative metabolic considerations

3.2.1. General characterization

Generative metabolic considerations allow researchers to formulate or develop new cognitive models. Using metabolic knowledge to generate a model belongs to the discovery process

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

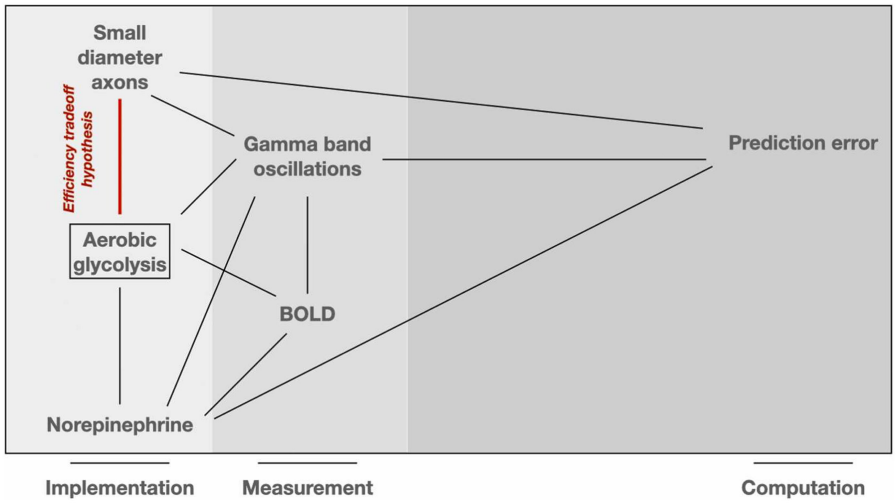


Figure 6. Theriault et al.'s conceptual model linking different features of neural structure and brain metabolism to measurable quantities and computational processes in the brain. Adapted from Theriault et al. (2023), Figure 6.

because it *populates* the space of possible models. Incorporating metabolism during discovery thus has different roles than during model evaluation. This section develops and makes precise these generative roles.

Generating models is valuable when researchers want to *explore* a target system (Gelfert, 2016, ch. 6). In this process, generative metabolic considerations function like a puzzle piece when exploring the relation between brain structure and the operations underlying cognitive capacities or neural systems. Metabolism links these two aspects because computations use energy, and structure determines how neural systems produce and consume energy. Models that include metabolism should allow researchers to explore the relation between brain structure and computation with greater detail.

Generative metabolic considerations can fulfill functions of exploratory modeling found across the physical and life sciences (Gelfert, 2016; Haueis, 2022). They can provide a *starting point* for further inquiry, for example, into the relation between brain structure, metabolism, and information processing or computation. They can provide a *proof-of-concept* that metabolic knowledge is relevant to modeling a particular cognitive capacity. They can deliver *potential explanations*, for example, of how a particular type of computation is implemented in a neural system. Lastly, they can be used to *reassess the target system*, that is, they can change assumptions about what kind of system the model targets.

3.2.2. Core example: Exploring efficiency tradeoffs in neural systems

A clear case of a generative use of metabolic considerations is from Theriault et al. (2023), who use the metabolic process of aerobic glycolysis to develop the *efficiency tradeoff hypothesis*. This hypothesis links information processing in thin axons (structure) to the transmission of prediction errors (computation). During tasks and the experimental resting state, the brain produces energy both by oxidative phosphorylation in mitochondria (30 ATP per glucose molecule) and aerobic glycolysis, a process where glucose is converted to energy while oxygen is present (2 ATP per glucose molecule). This fact presents a puzzle: why does the brain meet increased energy demand with the production inefficient process of aerobic glycolysis, when it could use the more production-efficient process of oxidative phosphorylation?

The efficiency-tradeoff hypothesis is a possible solution to this puzzle. It first uses principles of neural design (send at lowest rate, minimize wire) to state that thin axons are *information-energy*-efficient. Thin axons fire less often (spikes/sec), making each spike more informationally rich (bits/spike). Their small diameter makes them cheaper to repolarize (Na+ ions have greater effects on membrane voltage). Taken together, this means that thin axons maximize bits/ATP. By doing so, however, axons become too thin to house the mitochondria needed for oxidative phosphorylation. Therefore, this *production-efficient* process cannot fuel firing in thin axons, particularly during peak energy demand. The solution to this problem is aerobic glycolysis, which does not require mitochondria and thus can fuel spike transmission in thin axons. This solution pairs with the efficiency tradeoff because a production-inefficient process (aerobic glycolysis) is used to fuel a process that maximizes bits sent per ATP used.

The efficiency-tradeoff hypothesis motivates a larger attempt to develop a cognitive model using a predictive coding framework, where prediction errors are computations of deviations from an internal model constructed by the system in question.[10] Figure 6 shows that metabolic considerations function like a puzzle piece that links aspects of brain structure to measurable quantities and to computational capacities of the brain.

The link to prediction error brings out how generative uses of metabolism contribute to modeling cognitive capacities and neural systems. Figure 6 links prediction error, among other things, to firing in thin axons, since the firing characteristics of the latter match the computational profile of the former. In this model, aerobic glycolysis fuels the forwarding of prediction errors via information-energy-efficient thin axons.

The efficiency-tradeoff hypothesis and subsequent model formulation showcase exploratory functions of metabolism in cognitive modeling. First, Figure 6 represents a *starting point* for further inquiry. It suggests experiments for testing the decoupling between glucose and oxygen metabolism and gamma oscillations (Theriault et al., 2023, 16). Second, the link between aerobic glycolysis and prediction error provides a *proof-of-concept* that metabolic considerations are relevant to modeling. Aerobic glycolysis is used in executive function (Madsen et al., 1995), visual and tactile processing (Bednařík et al., 2015; Díaz-García et al., 2017), and spatial memory acquisition (Harris et al., 2019).[11] This suggests that the efficiency-tradeoff hypothesis can be tested by experiments with tasks that operationalize these cognitive capacities, measurement tools that

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

record activity in thin axons, and models that explain their activity in terms of prediction error transmission.

Third, the efficiency trade-off hypothesis provides a *potential explanation* to the puzzle of stimulation-based aerobic glycolysis. The existence of a tradeoff between two senses of efficiency (energy production and information rate efficiency), if corroborated, would explain why the brain uses an inefficient way to produce ATP when energy demand is high. Fourth, Figure 6 suggests that researchers should *reassess the target system* they are measuring, from one in which BOLD intensity increases indicate feature detection, to one where they indicate prediction error processing (cf., Theriault et al., 2023, pp. 23–24).

We analyze this example in detail not because we support every facet of this model. Rather, it shows that generative uses of metabolism occur. The core example shows how a model can be developed and adapted based on metabolic considerations.

3.2.3. Supporting example: Exploring energy-efficient learning strategies

Whereas the core example shows how metabolic considerations link known aspects of target systems, researchers can also use them to generate *novel strategies* of how neural systems execute cognitive capacities. An example is how energy efficiency can shape artificial neural network models of learning (Pache & Rossum, 2023). Assuming that animals have evolved and developed to learn by using as little energy as possible (section 2.1.2), Pache and Rossum explored energy-saving strategies in learning in neural networks trained with a backpropagation algorithm and tested on the MNIST digit classification task. In doing so, they find that "incorporating metabolic demands of learning leads to modifications of learning rules that are in line with experimental data" (Pache & Rossum, 2023, p. 5).

This exploratory study resulted in posits of three energy-saving strategies. First, a system saves energy by not updating on correct classifications. The saving scales with the size of the training dataset, though this increased consumption efficiency pairs with decreased robustness of the network to input variation. Second, a system saves energy by only updating synapses that require the largest magnitude of change. The saving scales with the number of synapses available given the relevant complexity of the task: if a network is far more complex than what is needed for the task, the savings are greater. Third, a system saves energy by using multiple forms of plasticity, allowing for some information to be cached in temporary, transient forms and other information to be consolidated into comparatively permanent forms. Unlike the others, this strategy is informed by higher energy costs of persistent, consolidated synaptic forms. The saving scales with the quantity of information updates that can be temporarily stored and minimally consolidated, though this must be balanced with potential information loss.

This work suggests learning rules that would be hard if not impossible to discover if modelers would ignore the brain's need to save energy. It thus explores a topic rarely addressed in cognitive models: the metabolic costs of plasticity (see Gershman, 2023 for an exception). The example thus illustrates our view that generative metabolic considerations can *populate* the space of possible cognitive models.

3.2.4. Guidelines: How to use metabolism to generate cognitive models

We can combine the general characterization with the details of the examples to formulate general guidelines for generative metabolic considerations in cognitive modeling (Box 2).

Box 2. Generative guidelines

Guidelines for generative metabolic considerations

G1: Posit or determine a cognitive capacity or a target system that one wants to model.

G2: Specify metabolic considerations that can inform a parameter in a model of this cognitive capacity or neural system.

G3: Construct (or adapt from an existing model) other parameters, fitting with those from (G2).

G4: Explore variations in parameters and their effects on other parameters of the model, with comparison to the target.

As this strategy involves generating a model, what one aims to model must be specified. The starting point of these guidelines is positing or determining a cognitive capacity or target system to be modeled (G1). Typically, the choice will be between a cognitive capacity and a neural system that might underwrite multiple cognitive capacities. One might, for instance, focus on modeling the activities of the hippocampus, even if one does not intend to model solely memory formation or consolidation. What is key is that the target choice delimits all subsequent modeling choices and therefore the success of the strategy. Lack of precision in specifying the target can make subsequent choices more nebulous. With a capacity or a system specified, one should select some set of considerations that will inform the parameters of the model in question. This means that researchers should specify metabolic considerations that can inform one or more parameters in a model of the target (G2). Examples include specific details of the efficiency of brain metabolism, principles of neural design, the characteristics of energy budgets, and metabolic costs of information processing (section 2.1). These considerations should be made at the start of generative modeling and should build on established knowledge of metabolic considerations relevant to the modeled capacity or system.

With metabolic considerations in hand and a target chosen, researchers should either construct parameters that fit with these considerations or adapt parameters from existing models of the same or comparable targets (G3). Researchers should thus generate the model by constructing other parameters. Importantly, G3 accommodates the fact that these choices need not be made *ex nihilo*. Rather, researchers can adapt parameters from existing models and established metabolic considerations, drawing upon the specifications made in G2. This is exemplified by the adaptation of a predictive coding framework in the core example and by the adaptation of neural network architectures in the supporting example. This integration is facilitated by drawing on previous attempts to model the target in question.

Finally, researchers should explore variations in the chosen parameters and their effects on other parameters in the model, comparing and assessing these variations' effects on the measure or manipulation of the chosen target (G4). One should explore how variations in parameters, including but not limited to those parameters informed by metabolic considerations, inform variations in other system parameters. In this sense, generative metabolic considerations provide: (1) a starting point for inquiry into the relation between brain structure and information processing, (2) a proof-of-concept that metabolic knowledge is relevant to cognitive modeling, or (3) a potential explanation of how a computation is implemented. This exploration can inform new hypotheses to be tested empirically and explored computationally or via simulation.

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

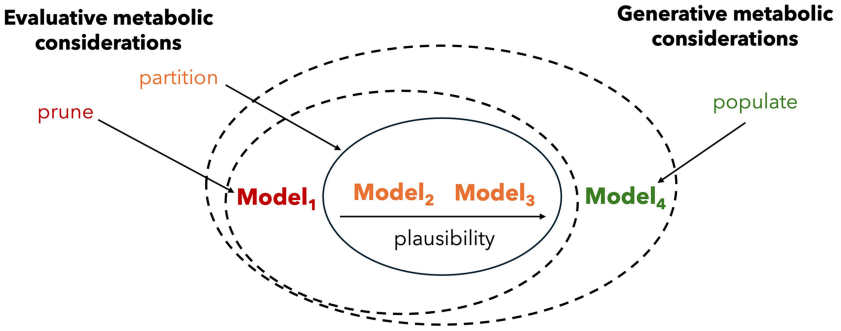


Figure 7. Schematic of the space of possible models for a cognitive capacity or neural system. Evaluative metabolic considerations prune the space by excluding models because they violate known metabolic constraints and partition the space by considering which models are more or less metabolically plausible. Generative metabolic considerations populate the space with new possible models, which can in turn be evaluated in terms of their plausibility.

This final generative guideline shows that the generation and evaluation of models that incorporate metabolic considerations can form a virtuous cycle (Figure 7). The former populates a space of possible models for cognitive capacities and neural systems. That is, generation gives us new models. Evaluation prunes or partitions this space. It allows us to assess the relative strengths and deficiencies of different models, and it also helps us determine which models are possible or plausible. This virtuous cycle, which entails generating and evaluating cognitive models via metabolism, can be iterative. Metabolic considerations provide a framework for taking in new metabolic knowledge, and assessing this knowledge reshapes how we assess existing models of cognitive capacities.

4. Implications of metabolic considerations for cognitive science, philosophy, and AI

Our account of metabolic considerations has far-reaching consequences for the sciences of the mind. For cognitive science (e.g., of mental effort), they provide a common standard to evaluate competing models (section 4.1). For philosophy, they show that even views like multiple realizability or medium independence of computation do not justify ignoring metabolism (section 4.2). And for AI, they push us to take energy into account when debating the parity between biological and artificial systems (section 4.3).

4.1. Metabolism and cognitive models of mental effort

In this section, we revisit the debate over models of mental effort and apply our account to it. Some cognitive tasks involve a conscious sensation of effort, and continuously exercising effortful tasks decreases one's performance over time (Ackerman, 2011). In this research, there is an active debate over whether metabolically inspired models or computational models best explain experimental findings (Holroyd, 2025). Our account suggests that while existing models may compete, there is no fundamental competition between appealing versus not appealing to metabolic resources. Rather, metabolic considerations can provide a common standard when evaluating any kind of model.

In an influential paper, Kurzban et al. (2013) distinguish *resource models* from *computational models* of mental effort (see also Holroyd, 2025, Box 1). Resource models explain decreases in task performance in mentally effortful tasks by positing a resource – either a domain-general information-processing resource or a local physical resource – that is depleted over time. An example of a domain-general resource model is Kahneman's capacity model of attention, while an example of a local resource model is the ego depletion model, according to which mentally effortful activities deplete glucose resources in the brain (Gaillot et al., 2007). By contrast, computational models explain mental effort as the result of brain mechanisms that solve a computational problem. In Kurzban et al.'s opportunity cost account, the computational problem is simultaneity – the organism cannot do everything at once. The solution to this problem is that the brain prioritizes a certain task by calculating the benefits and costs of continuing the task versus switching to another. The sensation of effort is the output of computational mechanisms that estimate these opportunity costs. Kurzban et al. argue that the empirical evidence speaks against local resource models like the ego depletion model and for computational models like their opportunity cost account.

Our account of metabolic considerations captures criticisms against local resource models while avoiding a false dichotomy between metabolism and computation in debates about mental effort. Pace the ego depletion model, there seems no empirical evidence that blood glucose levels vary between tasks with different levels of mental effort (Kurzban et al., 2013; Finley, Tang, & Schmeichel, 2019). Furthermore, a common theoretical assumption in neuroscience is that globally, brain metabolism does not vary between rest and mentally effortful tasks (Clarke & Sokoloff, 1999; Raichle, 2015; Zénon et al., 2019; Bermúdez, 2025). The criticism of the ego depletion model is thus is an evaluative metabolic consideration: the ego depletion model is inconsistent with empirical evidence and theoretical assumptions about brain metabolism. The criticism of local resource models thus prunes the space of models of mental effort to those that are consistent with the best empirical and theoretical evidence (Guideline E4).

Our account conflicts, however, with the idea that *either* a resource model *or* a computational model provides the best explanation of mental effort. Kurzban et al., for example, write that:

> We do not dispute that the brain is a biophysical system that obeys thermodynamic laws, nor do we dispute that the design of the brain is at some level constrained by energetic concerns. Where we disagree is with the notion that these principles are important to (or possibly even relevant to) explaining the phenomenology of subjective effort and task performance changes (2013, p. 714).

With their support of computational models, we take the position of these authors to be that modeling of cognitive effort (and perhaps of cognitive capacities in general) can safely ignore metabolism and energy costs, since they are not explanatorily relevant factors.

By contrast, our account implies that *any* model of cognitive effort (and by implication of any cognitive capacity) can benefit from accounting for the metabolic costs of the mechanisms that the model posits. So, even a computational model like the opportunity cost account can be evaluated based on metabolic knowledge about, for example, the energy costs of cost–benefit monitoring in the anterior cingulate cortex. Conversely, one can use metabolic knowledge to generate novel computational models, for example,

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

by positing computational mechanisms that satisfy an efficiency constraint (e.g., maximizing bit/ATP).

Our account implies that both critics and proponents of resource models of mental effort misconstrue the explanatory role of metabolic considerations. As explanatory constraints, metabolic resources are external to the process whose output they structure. They function as enabling conditions for cognitive mechanisms, rather than being operative in those mechanisms (Haueis, 2018). Enabling conditions typically have less specific effects on the outcomes that they enable (cf. Woodward, 2010, 317). Viewing metabolic resources as causally unspecific enabling conditions of mental effort thus explains why, for example, depleting glucose levels are only weakly correlated with decreases in task performance.

Our account, however, cautions critics to conclude that there is "little reason to explore hybrid accounts if computational accounts prove sufficient" (Kurzban et al., 2013, 714). Proving that any computational account is sufficient requires assessing, among other things, that it is consistent with evidence about brain metabolism. That a direct resource model is implausible thus does not speak for the plausibility of a computational model that is metabolically unconstrained. Once it is recognized that metabolic considerations provide a common standard for evaluating and generating models of mental effort, the supposed competition between resource models and computational models disappears.

While all models are partial and abstract or idealize aspects of the target system, there are benefits for including metabolism in cognitive modeling. First, including metabolism in *model evaluation* has benefits whether or not the models explicitly represent metabolic processes (Guideline E3). In the case of resource models, we learn whether the included metabolic facts are relevant to explain certain aspects of the phenomenon (e.g., phenomenology of effort), or whether they are best seen as enabling conditions. In the case of computational models, we learn whether the posited computational mechanisms are metabolically feasible or how we must modify them to make them so. In each case, we learn something about the models we cannot find out if we exclude metabolism from model evaluation.

Second, including metabolism when generating cognitive models might be hard, but it can be done. In our view, how relevant metabolism is for modeling particular cognitive capacities is a subject to be explored. It is not an *a priori* assumption to be made. As with any discovery process, model generation is risky. The benefit here is that including metabolism explicitly increases the chances of hitting upon a biologically plausible model of, for example, mental effort, whereas excluding it leaves modelers in the dark. We thus think that within the practical bounds of modeling projects, trying to include metabolism in evaluation or generation has benefits that justify the costs. We should try to explore these benefits as best as we can, even if it is not always possible to include metabolism.

The mental effort literature supports this view, as several models integrate computational and metabolic aspects of mental effort. One model assumes both that effortful tasks have high information costs, and that the invariance of metabolism constrains the number of subprocesses the brain can run in parallel (Zénon et al., 2019). Since effortful tasks require more information processing (and thus more energy), the brain would have to compensate them by allocating less energy to other tasks. Furthermore, local metabolic changes could structure transitions away from effortful tasks because "local information costs lead to progressive local metabolic alterations that accumulate over time" (Zénon et al., 2019, p. 13). Here, there is an explanatory link between the feeling of effort – a subjective experience of how fast metabolic changes accumulate – and fatigue – the anticipation or experience of metabolic accumulation.

Metabolic constraints can also be incorporated into computational models from a control theoretic perspective (Holroyd, 2025). Here, neural computations are conceptualized as low-dimensional transformations of representations – called "intrinsic manifold" – that occur in a high-dimensional state space of ensemble activity of neural populations. Around this low-dimensional state is the *controllosphere* – a part of state space which can only be reached by injecting large amounts of energy, for example, in the form of changing synaptic connectivity patterns. Mental effort occurs when the system expands from the intrinsic manifold into the controllosphere. This process could be constrained by the build-up of *metabolic waste*, which is itself caused by *computational mechanisms* involved in mental effort (ibid., 13).

The upshot is that Zénon et al.'s and Holroyd's model are neither purely resource-based nor purely computational models. Rather, they are models in which metabolism constrains which states a neural system can compute. These hybrid models indicate that biophysical and computational roles of mental effort are not competing explanatory posits. Instead, they are complementary elements that a model can integrate.

4.2. Metabolism, multiple realizability, and medium independence of computation

As physical system characteristics, metabolic considerations seem in tension with an outlook on cognitive modeling as focusing on the abstract, formal, or "computational" facets of a system. Metabolic considerations thus seem in tension with *multiple realizability* (MR) and *medium independence* (MI). In this section, we review tenets of these positions and argue that metabolic considerations do not *necessarily* violate them. Though there is tension between these positions and some views on metabolism (section 2.2.), these positions have little bite against our account of metabolic considerations in cognitive modeling.

MR is the position that mental kinds can be realized by physical kinds that share no features at this level of analysis (Putnam, 1967; Fodor, 1974; Polger & Shapiro, 2016). MI, with origins in AI (Haugeland, 1985), was reintroduced to philosophy by Garson (2003) and recently defended by Piccinini (2015, 2020) and Drayson (2025). We focus on Piccinini's construal of MI. Applied to a computational system,

> A vehicle is medium-independent just in case the rule (i.e., the input-output map) that defines a computation is sensitive only to differences between portions (i.e., spatiotemporal parts) of the vehicles along specific dimensions of variation – it is insensitive to any other physical properties of the vehicles (Piccinini, 2015, p. 122).

Piccinini continues, "the rules are functions of state variables associated with certain degrees of freedom, which can be implemented differently in different physical media" (2015, p. 122). By consequence, a computation "can be implemented in multiple physical media . . . provided that the media possess a sufficient number of dimensions of variation (or degrees of freedom)" (Piccinini, 2015, p. 122). On this view, "medium independence entails MR, because any medium independent property – such as performing a computation – can be realized by different kinds of medium" (Piccinini, 2020, p. 40, though see Williams, 2025).

It is debated whether metabolic considerations violate the tenets of MR or MI. On the one hand, proponents of MI, which for

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

Piccinini entails being a proponent of MR, address the importance of metabolism in modeling (Colombo & Piccinini, 2023, p. 32). However, they do not detail how metabolism relates to MI beyond generalist points (section 2.2.2). On the other hand, metabolic considerations are taken to challenge some construals of MR or MI (Thagard, 2022; Cao, 2022; Chirimuuta, 2022, 2024). For instance, Cao argues that "the point of metabolism is to exploit particular substrate-specific physical properties (e.g. of storing or releasing energy), rather than abstracting from them so as to implement some informational process" (Cao, 2022, p. 506). Consequently, "if metabolism is tied to a particular set of substrates, then it will anchor the information-processing stuff along with it" (ibid.).

The tension between these positions is thus as follows. Metabolic considerations might result in models of cognitive processing whose target vehicles or kinds are not medium-independent or multiply realizable, respectively. The metabolic demands of the brain differ from the energy consumption of other systems. By consequence, *changing* a physical medium or substrate might correspond to changes to the models in question. One interpretation is that the rule that defines the computational function should be only sensitive to computational variables and not to metabolic ones. Another interpretation is that the computational rules might be stable, but the metabolic ones are not, especially when implemented in different physical media.

There are two ways of addressing the potential tension between metabolic considerations and MR or MI. The first is objecting to MR or MI. One objection can be put as a question: should we fix the rules or the medium independence when assessing a computational system (Maley, 2023)? Rules might be determined first, after which we infer whether the system is medium-independent. Alternatively, medium independence might be determined first, after which we find systems whose rules abide by it. This "chicken-or-the-egg" scenario illustrates that it is unclear how to rule out certain physical processes while ruling in others, which is precisely what medium independence is supposed to do. This kind of position, sympathetic to medium *dependence*, has vocal supporters (e.g., Chirimuuta, 2022; Maley, 2025).

The second option is to accommodate metabolic considerations as part of MR or MI. Presuming physicalism, all non-hypothetical cognitive systems involve physical systems. Likewise, we can presume physical computational, information processing systems involve some degree of energy consumption. A characteristic of cognitive systems is that "input, output, and processing of information expend energy" (Thagard, 2022, p. 73), as "real world information processing depends on energy," and "energy depends on material substrates" (Thagard, 2022, p. 84).[12] Given this presumed commonality of physical systems, perhaps what is described by MR or MI should be sensitive to facets of energy consumption, which can be represented in modeling as variables where variation is tolerated. This would amount to "ruling in" metabolic considerations as medium-independent. MI, for instance, applies to "functions that are defined without reference to any *particular* physical medium" (Piccinini, 2020, p. 62, emphasis added), but in these cases we might not need to refer to anything more than that it *is* a physical medium. Thus, while the specifics of, for example, energy consumption vary between physical systems, the modeling of any physical system can involve the modeling of its energy consumption, variations of which can be tolerated just as MI tolerates other variations.

The second option also speaks to the challenge of squaring MR with metabolic considerations. Suppose that all physical systems consume energy, but the degree to which they consume energy can vary. Then, it is unclear whether we must deny that energy consumption may be multiply realizable, or that the same kind may be realized or implemented via systems with relevantly different energy uses.[13] Whether metabolic variations are consistent with MR will depend on whether these variations count as different *component realizers* of the same cognitive kind (Polger & Shapiro, 2016, p. 67). Because of this option, we are skeptical that appealing to metabolic considerations in cognitive modeling *necessarily* violates MR. The position leaves open that cognitive modelers can consider how energy consumption or regulation is realized in different systems. Even if we suppose that biological systems efficiently use energy, *how* efficiency manifests in the parameters of cognitive modeling remains open.

This option highlights that MR and MI have little bite against our call to use metabolism for cognitive modeling. At least some metabolic considerations tie to the physical characteristics of a system, which seems in tension with MR and MI when construed as substantive positions about the relation between cognitive and physical kinds. We do not deny this point. However, the second option highlights a critical detail that is not adequately discussed in the existing literature: *MR and MI are positions about (neuro) cognitive capacities or components thereof, but they are not positions about how one should develop and evaluate models of them.* Consider again generative metabolic considerations in modeling. While one might consider aspects of the realization or medium via these considerations, the model one develops based on them can nonetheless be consistent with MR or MI tenets. If all physical systems involve energy use, exploring how these systems use energy, which might initially be based on specific media, can offer insights into how different systems use energy, which, in turn, can be applied to other realizations or media.

One could, for instance, explore differences in energy consumption between systems when developing a model, with a focus on how different consumption levels relate to the computation in question. In essence, using metabolic considerations can produce better models of cognitive capacities, *even if* these capacities or components thereof are multiply realizable or medium-independent. This implies that some models might be deficient, but it is not a challenge to MR or MI *per se*. Rather, it is a challenge to a different – and less defensible – position: that MR or MI warrants us in *ignoring* metabolic considerations when modeling cognitive capacities or neural systems. This position, we wager, is not entailed by either MR or MI, and it is a position that we reject.

4.3. Energy costs and AI

Metabolic considerations have implications for assessing the perceived parity of AI and biological cognition. In this section, we show how our account helps to recast the debate that artificial systems like deep neural networks can achieve the same or close to the same cognitive function as humans.

AI systems – by which we mean the data, models and algorithms, and infrastructure needed to apply AI to a particular problem – are used to model cognitive and neural processes. Examples include using deep neural networks to predict the responses of visual neurons (Bashivan et al., 2019), generative

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

adversarial networks to model phonological learning (Beguš, 2020), or perceptron networks to estimate the efficiency of synaptic learning rules (Li & van Rossum, 2020). Running AI systems requires vast amounts of energy and other resources (International Energy Association, 2024), making their widespread use also a societal issue for sustainability and climate action (Schütze, 2024).

This vast difference in energy consumption makes clear that AI systems do not have the same energetic constraints as biological brains. This difference in energy use is telling. If one aims to defend a (near) parity between humans and AI, this functional parity is achieved via drastically different "metabolic" considerations. By contrast, if one is skeptical about this position, these energetic differences can help skeptics explore and find differences between these systems. Either way, incorporating metabolic considerations into cognitive modeling will result in our models of human cognitive capacities being different from our models of artificial "cognitive" capacities.

Including metabolic considerations opens novel avenues for comparing artificial and biological systems. Drawing from section 4.2, comparing systems with different energetic constrains highlights how the "same" function might be underwritten by different systems. Models of artificial and biological systems can be contrasted against one another. The use of these models helps make salient the differences between the systems, their operations, and their respective robustness, highlighting the degree to which metabolic constraints have shaped biological systems by comparing them to systems that have dissimilar constraints.

Metabolic considerations should also motivate proponents and skeptics alike to explicate their notion of parity. Is the parity simply behavioral output, for example, that AI systems provide responses with the same level of accuracy as humans (Yarkoni & Westfall, 2017)? Or does parity apply to the internal workings of the systems (Lilicrap & Kording, 2019; Cao & Yamins, 2024)? While differences in energy consumption do not challenge the former parity claim, they cast doubt on the latter one. If energy availability constrains the information processing and computations a system can perform (section 3.1.1), then different energy needs provide defeasible clues that artificial and biological systems perform different computations. Including metabolic considerations is thus in line with recent calls to constrain AI assessments by using prior theories and data about domain-specific principles of cognitive processing or neural architecture (Bowers et al., 2022). We argue that metabolic considerations are another constraint to further prune and partition the space of models. The example of finding energy-efficient synaptic learning rules shows that some researchers already use metabolism to constrain AI systems this way (section 3.2.3).

AI systems currently achieve many great things because they have practically unbounded amounts of energy available to them. Nonetheless, it is possible to prioritize information-energy efficiency more directly. An example are *analog AI chips,* which eliminate the von Neumann bottleneck, that is, the energy cost of transferring data between memory and central processing units (Ambrogio et al., 2023). Such chips use phase change memory (PCM), in which electrical pulses change between a crystalline and amorphous state (0 and 1). PCMs can encode in-between states (synaptic weights) through combinations of current and voltage, and they are non-volatile, that is, they save these states even when the chip is unpowered. These analog chips can implement the multiply-and-accumulate computation, and a fully connected network of PCMs can solve a keyword spotting tasks and the RNNT task at accuracy levels comparable to conventional deep learning networks on digital computers (Ambrogio et al., 2023). By avoiding the von Neumann bottleneck, this analog AI chip reaches 14-fold increase in information-energy efficiency compared to other systems (6704 tera-operations per seconds per Watt, cf. ibid., p. 775).

While analog AI chips can currently implement only one type of computation, some *neuromorphic computing systems* can implement entire deep learning algorithms, such as error back-propagation on the BrainScaleS-2 computing platform (Göltz et al., 2021). BrainScaleS-2 cuts energy costs by speeding up neural operations while preserving the temporal relations between them. A time-to-first spike code increases information-energy efficiency further since the timing of the neuron's first spike encodes the presence of a feature in the stimulus. This neuromorphic AI system achieves 96.9% accuracy on the benchmark MNIST classification task (10000 samples), while using 8.4 μJ per classification.[14]

The existence of analog AI and neuromorphic systems – although currently not scaled to industrial standards – shows that energetic considerations can drive technological innovation that is inspired by neural systems. Yet, we should be cautious not to overestimate these developments. AI developers may not prioritize energy efficiency as long as societal actors keep willing to prioritize the use of electricity, water, and other resources to AI systems, rather than allocating these resources to satisfy human needs directly (Crawford, 2021). It thus may be doubtful that biologically inspired computational hardware will supplant existing artificial hardware paradigms, unless societies set up stronger energy constraints for AI that force these developers' hands (Schütze, 2023). Though our doubt might be discouraging, it also highlights the stark differences between current artificial and biological systems. The former have the luxury, for lack of a better word, of achieving their outputs at the cost of a great deal of energy. The latter, shaped by evolution and the environment, do not.

5. Conclusion

In this paper, we have offered a preliminary yet systematic account of incorporating metabolic considerations into cognitive modeling. Philosophical and scientific investigations of metabolism prove the value of this approach, and recent empirical insights offer us the content needed to both evaluate and generate cognitive models. We argue that cognitive models should be consistent with the brain's metabolic limits, and modelers should assess how their models fit within these bounds. This does not mean that all models must explicitly represent metabolic considerations; rather, it means that models should be assessed in light of them. Our conclusion is in stark contrast with the widespread tendency to ignore what many scholars take to be – erroneously, in our view – superfluous or "noncognitive" details. We hope to have made a case for why these details are not superfluous. Metabolic considerations are vitally needed for evaluating and generating models of cognition.

Acknowledgments. We would like to thank Jordan Theriault and participants at the workshop "Concepts for understanding brain organization" (June 10 2024, ENS Paris), the SPAN meeting 2025 in St Louis, and the BSPS 2025 in Glasgow for providing helpful feedback on previous versions of this project. We would also like to thank Jacqueline Riffel, who reviewed some of the material analyzed in section 2.1 during an internship with Philipp Haueis in 2022.

Financial support. David Colaço is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project Number 528370787.

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

Competing interests. The authors declare no competing interest.

Notes

[1] See Stadler (2010, ch. 3, Stadler, 2011) for possible historical explanations for this neglect.

[2] Information rate is measured in bits/sec, where "one *bit* is the information needed to decide between two equally likely alternatives" (Sterling & Laughlin, 2015, p. 51, emphasis in original).

[3] Some signaling activities may not process behaviorally relevant information (Merker, 2013; Haueis, 2018). In that case, the energy for cognitive processing is limited to a subset of signaling functions.

[4] An intermediary step is estimating the bitrate per second. For example: assuming a firing rate of 4Hz and a vesicle release probability of 0.25, Harris et al. (2012) estimate that 6–32 bit/s are being transmitted at the synapse.

[5] The technical definition of information gain is the rate of the Kullback–Leibler divergence, where the latter specifies how much information an approximate probability distribution contains about an (empirical) probability distribution of states of a physical system. This definition suggests a close coupling between the energy spent for synaptic plasticity changes and the information gained about the physical state of the system (Karbowski, 2024).

[6] Karbowksi and Urban (2024, p. 303) provide their estimate in kT (Boltzmann constant times absolute brain temperature), where 20 kT is roughly 1 ATP. Note that there is an efficiency effect for increasing numbers of synapses (ibid., Figure 14).

[7] Thanks to Patrick McGivern for suggesting the phrase "metabolic turn."

[8] The group has extended this model to formulate an energy-dependent model of spike-time-dependent synaptic plasticity and dynamics in recurrent excitatory-inhibitory networks (Jaras et al., 2025).

[9] The authors may not view metabolism as external to behaviorally relevant information processing because they hold that behavioral consequences of neural activity directly result from neurons meeting their metabolic needs.

[10] A classic example of prediction errors is a difference between a predicted visual stimulus and the actual visual input (Rao & Ballard, 1999). They are the most informative signals in these frameworks because they occur infrequently and encode differences between the incoming and predicted signal.

[11] Prediction error models for these capacities also exist (Schultz, 2016; Colett, Mollick, & Kober, 2022; Kennedy et al., 2024).

[12] Our claims apply to dissipative systems, which are not at thermodynamic equilibrium and thus exchange energy with their environments (Karbowski, 2024). Whether a physical system that computes or processes information could be non-dissipative is a topic for discussion, but the idea seems more an idealizing assumption in modeling (e.g., Gallicchio, 2024). This posit is consistent with Landauer's principle (Anderson & Piccinini, 2024, 107).

[13] These different phrasings capture whether energy use is construed as the would-be realized or would-be realizer, respectively.

[14] Even for a given task, energy costs of AI systems are currently difficult to compare. Other system parameters vary (e.g. classifications/second, see Göltz et al., 2021 Table 2), and energy measures are inconsistent (Draguet & Frénay 2025). It may be hard to infer from reported differences in magnitude – ranging from picojoule over micro joule to megajoule – to the actual energy an AI system uses (see Hirtzlin et al., 2019; Göltz et al., 2021; Table 2, Wang et al., 2022 for further discussion).

References

Ackerman, P. L. (2011). *Cognitive fatigue: Multidisciplinary perspectives on current research and future applications*. American Psychological Association. https://doi.org/10.1037/12343-000

Allen, C. (2017). On (not) defining cognition. *Synthese*, *194*(11), 4233–4249. https://doi.org/10.1007/s11229-017-1454-4

Ambrogio, S., Narayanan, P., Okazaki, A., Fasoli, A., Mackin, C., Hosokawa, K., Nomura, A., Yasuda, T., Chen, A., Friz, A., Ishii, M., Luquin, J., Kohda, Y., Saulnier, N., Brew, K., Choi, S., Ok, I., Philip, T., Chan, V., … Burr, G. W. (2023). An analog-AI chip for energy-efficient speech recognition and transcription. *Nature*, *620*(7975), 768–775. https://doi.org/10.1038/s41586-023-06337-5

Amodei, D., Ananthanarayanan, S., Anubhai, R., Bai, J., Battenberg, E., Case, C., Casper, J., Catanzaro, B., Cheng, Q., Chen, G., Chen, J., Chen, J., Chen, Z., Chrzanowski, M., Coates, A., Diamos, G., Ding, K., Du, N., Elsen, E., … & Zhu, Z. (2016). Deep speech 2: End-to-end speech recognition in English and Mandarin. *Proceedings of Machine Learning Research*, *48*, 173–182. https://proceedings.mlr.press/v48/amodei16.html

Anderson, J. R. (2014). *Rules of the mind*. Psychology Press, Taylor & Francis Group. https://doi.org/10.4324/9781315806938

Anderson, N. G., & Piccinini, G. (2024). *The physical signature of computation: A robust mapping account*. Oxford University Press.

Attwell, D., & Laughlin, S. B. (2001). An energy budget for signaling in the grey matter of the brain. *Journal of Cerebral Blood Flow & Metabolism*, *21*(10), 1133–1145. https://doi.org/10.1097/00004647-200110000-00001

Barron, A. B., Halina, M., & Klein, C. (2023). Transitions in cognitive evolution. *Proceedings of the Royal Society B: Biological Sciences*, *290*(2002), 20230671. https://doi.org/10.1098/rspb.2023.0671

Bashivan, P., Kar, K., & DiCarlo, J. J. (2019). Neural population control via deep image synthesis. *Science*, *364*(6439). https://doi.org/10.1126/science.aav9436

Bechtel, W. (2008). *Mental mechanisms: Philosophical perspectives on cognitive neuroscience*. Psychology Press, Taylor & Francis Group.

Bednařík, P., Tkáč, I., Giove, F., DiNuzzo, M., Deelchand, D. K., Emir, U. E., Eberly, L. E., & Mangia, S. (2015). Neurochemical and BOLD responses during neuronal activation measured in the human visual cortex at 7 Tesla. *Journal of Cerebral Blood Flow & Metabolism*, *35*(4), 601–610. https://doi.org/10.1038/jcbfm.2014.233

Beer, R. D., & Williams, P. L. (2015). Information processing and dynamics in minimally cognitive agents. *Cognitive Science*, *39*(1), 1–38. https://doi.org/10.1111/cogs.12142

Beguš, G. (2020). Generative adversarial phonology: Modeling unsupervised phonetic and phonological learning with neural networks. *Frontiers in Artificial Intelligence*, *3*, 44.

Benna, M. K., & Fusi, S. (2016). Computational principles of synaptic memory consolidation. *Nature Neuroscience*, *19*(12), 1697–1706. https://doi.org/10.1038/nn.4401

Bermúdez, J. P. (2025). What is the feeling of effort about?. *Australasian Journal of Philosophy*, *103*, 88–105. https://doi.org/10.1080/00048402.2024.2351208

Bond, K., Rasero, J., Madan, R., Bahuguna, J., Rubin, J., & Verstynen, T. (2023). Competing neural representations of choice shape evidence accumulation in humans. *eLife*, *12*, e85223. https://doi.org/10.7554/eLife.85223

Bowers, J. S., Malhotra, G., Dujmovic, M., Llera Montero, M., Tsvetkov, C., Biscione, V., ... Blything, R. (2022). Deep problems with neural network models of human vision. *Behavioral and Brain Sciences*, *46*, e385. https://doi.org/10.1017/S0140525X22002813

Bruckmaier, M., Tachtsidis, I., Phan, P., & Lavie, N. (2020). Attention and capacity limits in perception: A cellular metabolism account. *Journal of Neuroscience*, *40*(35), 6801–6811. https://doi.org/10.1523/JNEUROSCI.2368-19.2020

Cao, R. (2022). Multiple realizability and the spirit of functionalism. *Synthese*, *200*(6), 506. https://doi.org/10.1007/s11229-022-03524-1

Cao, R., & Yamins, D. (2024). Explanatory models in neuroscience, Part 2: Functional intelligibility and the contravariance principle. *Cognitive Systems Research*, *85*, 101200. https://doi.org/10.1016/j.cogsys.2023.101200

Chirimuuta, M. (2022). The case for medium dependence: Comment on neurocognitive mechanisms by Gualtiero Piccinini. *Journal of Consciousness Studies*, *29*(7), 185–194. https://doi.org/10.53765/20512201.29.7.185

Chirimuuta, M. (2024). *The Brain abstracted: Simplification in the history and philosophy of neuroscience*. MIT Press.

Churchland, P. S., & Sejnowski, T. J. (1992). *The computational brain*. MIT Press.

Clarke, D. D., & Sokoloff, L. (1999). Circulation and energy metabolism of the brain. In George J. Siegel, Bernard W. Agranoff, R. Wayne Albers, Stephen K. Fisher, & Michael D. Uhler (Eds.), *Basic neurochemistry: Molecular, cellular and medical aspects* (pp. 637–669). Lippincott-Raven.

Colaço, D. (2025). On consistently assessing alleged mnemonic systems (or, why isn't immune memory "really" memory?). *Review of Philosophy and Psychology 16*, 1085–1103. https://doi.org/10.1007/s13164-025-00768-x

Colombo, M., & Piccinini, G. (2023). *The computational theory of mind*. Cambridge University Press.

Colombo, M., & Piccinini, G. (2024). *The computational theory of mind*. Cambridge University Press. https://doi.org/10.1017/9781009183734

Corlett, P. R., Mollick, J. A., & Kober, H. (2022). Meta-analysis of human prediction error for incentives, perception, cognition, and action. *Neuropsychopharmacology*, *47*(7), 1339–1349. https://doi.org/10.1038/s41386-021-01264-3

Craver, C. F., & Darden, L. (2013). *In search of mechanisms: Discoveries across the life sciences*. University of Chicago Press.

Crawford, K. (2021). *Atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. Yale University Press.

Dayan, P., & Abbott, L. F. (2005). *Theoretical neuroscience: computational and mathematical modeling of neural systems*. MIT Press.

deCharms, R. C., & Zador, A. (2000). Neural representation and the cortical code. *Annual Review of Neuroscience*, *23*(1), 613–647. https://doi.org/10.1146/annurev.neuro.23.1.613

Díaz-García, C. M., Mongeon, R., Lahmann, C., Koveal, D., Zucker, H., & Yellen, G. (2017). Neuronal stimulation triggers neuronal glycolysis and not lactate uptake. *Cell Metabolism*, *26*(2), 361–374.e4. https://doi.org/10.1016/j.cmet.2017.06.021

Dienel, G. A. (2014). Chapter 3—Energy metabolism in the brain. In J. H. Byrne, R. Heidelberger, & M. N. Waxham (Eds.), *From molecules to networks* (3rd ed., pp. 53–117). Academic Press. https://doi.org/10.1016/B978-0-12-397179-1.00003-8

Dienel, G. A. (2019). Brain glucose metabolism: Integration of energetics with function. *Physiological Reviews*, *99*(1), 949–1045. https://doi.org/10.1152/physrev.00062.2017

Draguet, N., & Frénay, B. (2025). *Making convolutional neural networks energy-efficient: An introduction*. European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, Bruges (Belgium). m http://www.i6doc.com/en/

Drayson, Z. (2025). Defending the medium-independence of computation, *40*(4),458–467. *Mind & Language*. https://doi.org/10.1111/mila.12536

Eliasmith, C. (2013). *How to build a brain: A neural architecture for biological cognition*. OUP USA.

Eliasmith, C. (2015). *How to build a brain: A neural architecture for biological cognition*. Oxford University Press.

Engl, E., & Attwell, D. (2015). Non-signalling energy use in the brain. *The Journal of Physiology*, *593*(16), 3417–3429. https://doi.org/10.1113/jphysiol.2014.282517

Finley, A. J., Tang, D., & Schmeichel, B. J. (2019). Sweet nothings: No effects of self-control exertion on blood glucose levels. *Social Psychology*, *50*(5–6), 322–331. https://doi.org/10.1027/1864-9335/a000376

Fodor, J. A. (1974). Special sciences (Or: The disunity of science as a working hypothesis). *Synthese*, *28*(2), 97–115. https://doi.org/10.1007/BF00485230

Gailliot, M. T., Baumeister, R. F., DeWall, C. N., Maner, J. K., Plant, E. A., Tice, D. M., Brewer, L. E., & Schmeichel, B. J. (2007). Self-control relies on glucose as a limited energy source: Willpower is more than a metaphor. *Journal of Personality and Social Psychology*, *92*(2), 325–336. https://doi.org/10.1037/0022-3514.92.2.325

Gallicchio, C. (2024). Euler state networks: Non-dissipative reservoir computing. *Neurocomputing*, *579*, 127411. https://doi.org/10.1016/j.neucom.2024.127411

Garson, J. (2003). The introduction of information into neurobiology. *Philosophy of Science*, *70*(5), 926–936. https://doi.org/10.1086/377378

Gelfert, A. (2016). *How to do science with models: A philosophical primer*. Springer.

Gershman, S. J. (2023). The molecular memory code and synaptic plasticity: A synthesis. *Biosystems*, *224*, 104825. https://doi.org/10.1016/j.biosystems.2022.104825

Gervais, R. (2015). Mechanistic and non-mechanistic varieties of dynamical models in cognitive science: Explanatory power, understanding, and the 'mere description' worry. *Synthese*, *192*(1), 43–66. https://doi.org/10.1007/s11229-014-0548-5

Godfrey-Smith, P. (2016). Mind, matter, and metabolism. *The Journal of Philosophy*, *113*(10), 481–506. https://www.jstor.org/stable/48568275

Göltz, J., Kriener, L., Baumbach, A., Billaudelle, S., Breitwieser, O., Cramer, B., Dold, D., Kungl, A. F., Senn, W., Schemmel, J., Meier, K., & Petrovici, M. A. (2021). Fast and energy-efficient neuromorphic deep learning with first-spike times. *Nature Machine Intelligence*, *3*(9), 823–835. https://doi.org/10.1038/s42256-021-00388-x

Harris, J. J., Jolivet, R., & Attwell, D. (2012). Synaptic energy use and supply. *Neuron*, *75*(5), 762–777. https://doi.org/10.1016/j.neuron.2012.08.019

Harris, R. A., Lone, A., Lim, H., Martinez, F., Frame, A. K., Scholl, T. J., & Cumming, R. C. (2019). Aerobic glycolysis is required for spatial memory acquisition but not memory retrieval in mice. *Eneuro*, *6*(1), ENEURO.0389–18.2019. https://doi.org/10.1523/ENEURO.0389-18.2019

Haueis, P. (2018). Beyond cognitive myopia: A patchwork approach to the concept of neural function. *Synthese*, *195*(12), 5373–5402. https://doi.org/10.1007/s11229-018-01991-z

Haueis, P. (2022). Descriptive multiscale modeling in data-driven neuroscience. *Synthese*, *200*(2), 129. https://doi.org/10.1007/s11229-022-03551-y

Haugeland, J. (1985). *Artificial intelligence: The very idea*. MIT Press.

Herculano-Houzel, S. (2011). Not all brains are made the same: new views on brain scaling in evolution. *Brain, Behavior and Evolution*, *78*(1), 22–36. https://doi.org/10.1159/000327318

Hirtzlin, T., Penkovsky, B., Bocquet, M., Klein, J.-O., Portal, J.-M., & Querlioz, D. (2019). Stochastic computing for hardware implementation of binarized neural networks. *IEEE Access*, *7*, 76394–76403. https://doi.org/10.1109/access.2019.2921104

Hofman, M. A. (1983). Evolution of brain size in neonatal and adult placental mammals: a theoretical approach. *Journal of Theoretical Biology*, *105*(2), 317–332. https://doi.org/10.1016/s0022-5193(83)80011-3

Holroyd, C. B. (2025). The controllosphere: The neural origin of cognitive effort. *Psychological Review132*(3), 603–631. https://doi.org/10.1037/rev0000467

Hosokawa, T., Xu, M., Katori, Y., Yamada, M., Aihara, K., & Tsutsui, K.-I. (2022). Monkey prefrontal single-unit activity reflecting category-based logical thinking process and its neural network model. *Journal of Neuroscience*, *42*(33), 6380–6391. https://doi.org/10.1523/JNEUROSCI.2286-21.2022

Howarth, C., Gleeson, P., & Attwell, D. (2012). Updated energy budgets for neural computation in the neocortex and cerebellum. *Journal of Cerebral Blood Flow & Metabolism*, *32*(7), 1222–1232. https://doi.org/10.1038/jcbfm.2012.35

International Energy Agency. (2024). *Electricity 2024—Analysis and forecast to 2026*.

Jaras, I., Harada, T., Orchard, M. E., Maldonado, P. E., & Vergara, R. C. (2021). Extending the integrate-and-fire model to account for metabolic dependencies. *European Journal of Neuroscience*, *54*(4), 5249–5260. https://doi.org/10.1111/ejn.15326

Jaras, I., Orchard, M. E., Maldonado, P. E., & Vergara, R. C. (2025). Unveiling the role of local metabolic constraints on the structure and activity of spiking neural networks. *PLOS Computational Biology*, *21*(6), e1013148. https://doi.org/10.1371/journal.pcbi.1013148

Karbowski, J. (2007). Global and regional brain metabolic scaling and its functional consequences. *BMC Biology*, *5*(1), 18. https://doi.org/10.1186/1741-7007-5-18

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

Karbowski, J. (2019). Metabolic constraints on synaptic learning and memory. *Journal of Neurophysiology*, *122*(4), 1473–1490. https://doi.org/10.1152/jn.00092.2019

Karbowski, J. (2024). Information thermodynamics: From physics to neuroscience. *Entropy*, *26*(9), 779. https://doi.org/10.3390/e26090779

Karbowski, J., & Urban, P. (2024). Cooperativity, information gain, and energy cost during early LTP in dendritic spines. *Neural Computation*, *36*(2), 271–311. https://doi.org/10.1162/neco_a_01632

Kennedy, N. G., Lee, J. C., Killcross, S., Westbrook, R. F., & Holmes, N. M. (2024). Prediction error determines how memories are organized in the brain. *eLife*, *13*, RP95849. https://doi.org/10.7554/eLife.95849

Klein, C. (2018). Mechanisms, resources, and background conditions. *Biology & Philosophy*, *33*(5), 36. https://doi.org/10.1007/s10539-018-9646-y

Klein, C. (2022). Explaining neural transitions through resource constraints. *Philosophy of Science*, *89*(5), 1196–1202. https://doi.org/10.1017/psa.2022.35

Koch, C. (2004). *Biophysics of computation: Information processing in single neurons*. Oxford University Press. https://doi.org/10.1093/oso/9780195104912.001.0001

Kriegeskorte, N., & Douglas, P. K. (2018). Cognitive computational neuroscience. *Nature Neuroscience*, *21*(9), 1148. https://doi.org/10.1038/s41593-018-0210-5

Kurzban, R., Duckworth, A., Kable, J. W., & Myers, J. (2013). An opportunity cost model of subjective effort and task performance. *Behavioral and Brain Sciences*, *36*(6), 661–679. https://doi.org/10.1017/s0140525x12003196

Lennie, P. (2003). The cost of cortical computation. *Current Biology*, *13*(6), 493–497. https://doi.org/10.1016/S0960-9822(03)00135-0

Levy, W. B., & Baxter, R. A. (1996). Energy efficient neural codes. *Neural Computation*, *8*(3), 531–543. https://doi.org/10.1162/neco.1996.8.3.531

Levy, W. B., & Calvert, V. G. (2021). Communication consumes 35 times more energy than computation in the human cortex, but both costs are needed to predict synapse number. *Proceedings of the National Academy of Sciences*, *118*(18)1–12. https://doi.org/10.1073/pnas.2008173118

Li, H. L., & Van Rossum, M. C. (2020). Energy efficient synaptic plasticity. *Elife*, *9*, e50804. https://doi.org/10.7554/eLife.50804

Lieder, F., & Griffiths, T. L. (2020). Resource-Rational analysis: Understanding human cognition as the optimal use of limited computational resources. *Behavioral and Brain Sciences*, *43*, e1. https://doi.org/10.1017/S0140525X1900061X

Lillicrap, T. P., & Kording, K. P. (2019). *What does it mean to understand a neural network?* (arXiv:1907.06374). arXiv. https://doi.org/10.48550/arXiv.1907.06374

Lisman, J. E. (1985). A mechanism for memory storage insensitive to molecular turnover: A bistable autophosphorylating kinase. *Proceedings of the National Academy of Sciences*, *82*(9), 3055–3057. https://doi.org/10.1073/pnas.82.9.3055

López-Jury, L., García-Rosales, F., González-Palomares, E., Kössl, M., & Hechavarria, J. C. (2021). Acoustic context modulates natural sound discrimination in auditory cortex through frequency-specific adaptation. *Journal of Neuroscience*, *41*(50), 10261–10277. https://doi.org/10.1523/JNEUROSCI.0873-21.2021

Ma, W. J., & Woodford, M. (2020). Multiple conceptions of resource rationality. *Behavioral and Brain Sciences*, *43*, 30–31. https://doi.org/10.1017/s0140525x19001754

Madsen, P. L., Hasselbalch, S. G., Hagemann, L. P., Olsen, K. S., Bülow, J., Holm, S., Wildschiødtz, G., Paulson, O. B., & Lassen, N. A. (1995). Persistent resetting of the cerebral oxygen/glucose uptake ratio by brain activation: Evidence obtained with the Kety—Schmidt technique. *Journal of Cerebral Blood Flow & Metabolism*, *15*(3), 485–491. https://doi.org/10.1038/jcbfm.1995.60

Maley, C. J. (2023). Medium independence and the failure of the mechanistic account of computation. *Ergo: An Open Access Journal of Philosophy*, *10*, 28. https://doi.org/10.3998/ergo.4658

Maley, C. J. (2025). Declaring independence from medium independence *40*(4), 448–457. *Mind & Language*. https://doi.org/10.1111/mila.12543

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. W.H. Freeman.

Meier, J. M., Perdikis, D., Blickensdörfer, A., Stefanovski, L., Liu, Q., Maith, O., Dinkelbach, H. Ü., Baladron, J., Hamker, F. H., & Ritter, P. (2022). Virtual deep brain stimulation: Multiscale co-simulation of a spiking basal ganglia model and a whole-brain mean-field model with the virtual brain. *Experimental Neurology*, *354*, 114111. https://doi.org/10.1016/j.expneurol.2022.114111

Merker, B. (2013). Cortical gamma oscillations: The functional key is activation, not cognition. *Neuroscience & Biobehavioral Reviews*, *37*(3), 401–417. https://doi.org/10.1016/j.neubiorev.2013.01.013

Mink, J. W., Blumenschine, R. J., & Adams, D. B. (1981). Ratio of central nervous system to body metabolism in vertebrates: its constancy and functional basis. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, *241*(3), R203–R212. https://doi.org/10.1152/ajpregu.1981.241.3.r203

Moore, C. I., & Cao, R. (2008). The hemo-neural hypothesis: On the role of blood flow in information processing. *Journal of Neurophysiology*, *99*(5), 2035–2047. https://doi.org/10.1152/jn.01366.2006

Moustafa, A. A. (Ed.). (2017). *Computational models of brain and behavior*. (1st ed.) Wiley. https://doi.org/10.1002/9781119159193

Pache, A., & Van Rossum, M. C. W. (2023). Energetically efficient learning in neuronal networks. *Current Opinion in Neurobiology*, *83*, 102779. https://doi.org/10.1016/j.conb.2023.102779

Padamsey, Z., & Rochefort, N. L. (2023). Paying the brain's energy bill. *Current Opinion in Neurobiology*, *78*, 102668. https://doi.org/10.1016/j.conb.2022.102668

Perge, J. A., Koch, K., Miller, R., Sterling, P., & Balasubramanian, V. (2009). How the optic nerve allocates space, energy capacity, and information. *Journal of Neuroscience*, *29*(24), 7917–7928. https://doi.org/10.1523/JNEUROSCI.5200-08.2009

Piccinini, G. (2015). *Physical computation: A mechanistic account*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199658855.001.0001

Piccinini, G. (2020). *Neurocognitive mechanisms: Explaining biological cognition*. Oxford University Press.

Polger, T. W., & Shapiro, L. A. (2016). *The multiple realization book*. Oxford University Press.

Pontzer, H. (2015). Energy expenditure in humans and other primates: A new synthesis. *Annual Review of Anthropology*, *44*(1), 169–187. https://doi.org/10.1146/annurev-anthro-102214-013925

Putnam, H. (1967). Psychological predicates. In W. H. Capitan, & D. D. Merrill (Eds.), *Art, mind, and religion* (pp. 37–48). University of Pittsburgh Press.

Raichle, M. E. (2015). The restless brain: How intrinsic activity organizes brain function. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1668), 20140172. https://doi.org/10.1098/rstb.2014.0172

Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*(1), 79–87. https://doi.org/10.1038/4580

Rieke, F., Warland, D., van Steveninck, R. de R., & Bialek, W. S. (1999). *Spikes: Exploring the neural code*. MIT Press.

Ritter, P., Schirner, M., McIntosh, A. R., & Jirsa, V. K. (2013). The virtual brain integrates computational modeling and multimodal neuroimaging. *Brain Connectivity*, *3*(2), 121–145. https://doi.org/10.1089/brain.2012.0120

Ross, L. N. (2023). The explanatory nature of constraints: Law-based, mathematical, and causal. *Synthese*, *202*(2), 56. https://doi.org/10.1007/s11229-023-04281-5

Salaj, D., Subramoney, A., Kraisnikovic, C., Bellec, G., Legenstein, R., & Maass, W. (2021). Spike frequency adaptation supports network computations on temporally dispersed information. *eLife*, *10*, e65459. https://doi.org/10.7554/eLife.65459

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

Samsonovich, A., & McNaughton, B. L. (1997). Path integration and cognitive mapping in a continuous attractor neural network model. *Journal of Neuroscience*, *17*(15), 5900–5920. https://doi.org/10.1523/JNEUROSCI.17-15-05900.1997

Schultz, W. (2016). Dopamine reward prediction-error signalling: A two-component response. *Nature Reviews Neuroscience*, *17*(3), 183–195. https://doi.org/10.1038/nrn.2015.26

Schütze, P. (2024). The problem of sustainable AI: A critical assessment of an emerging phenomenon. *Weizenbaum Journal of the Digital Society*, *4*(1) 1–28. https://doi.org/10.34669/WI.WJDS/4.1.4

Sengupta, B., Faisal, A. A., Laughlin, S. B., & Niven, J. E. (2013). The effect of cell size and channel density on neuronal information encoding and energy efficiency. *Journal of Cerebral Blood Flow & Metabolism*, *33*(9), 1465–1473. https://doi.org/10.1038/jcbfm.2013.103

Sengupta, B., Laughlin, S. B., & Niven, J. E. (2014). Consequences of converting graded to action potentials upon neural information coding and energy efficiency. *PLoS Computational Biology*, *10*(1), e1003439. https://doi.org/10.1371/journal.pcbi.1003439

Sengupta, B., & Stemmler, M. B. (2014). Power consumption during neuronal computation. *Proceedings of the IEEE*, *102*(5), 738–750. https://doi.org/10.1109/JPROC.2014.2307755

Shiu, P. K., Sterne, G. R., Spiller, N., Franconville, R., Sandoval, A., Zhou, J., Simha, N., Kang, C. H., Yu, S., Kim, J. S., Dorkenwald, S., Matsliah, A., Schlegel, P., Szi-Chieh, Y., McKellar, C. E., Sterling, A., Costa, M., Eichler, K., Jefferis, G. S. X. E., . . . & Scott, K. (2024). A *Drosophila* computational brain model reveals sensorimotor processing. *Nature*, 634, 210–219. https://doi.org/10.1038/s41586-024-07763-9

Stadler, M. (2010). *Assembling life. Models, the cell, and the reformations of biological science, 1920-1960*. [PhD Dissertation, Imperial College]. https://www.ultrabeige.de/pdf/2010_assembling_life.pdf

Stadler, M. (2011). Biophysikalisches Doppelleben, 1939–1946. Oder: Räume der Lange-weile. Zur (Nicht-)Zäsur des 'Informationsdiskurses.' *Berichte Zur Wissenschaftsgeschichte*, *34*(1), 27–63. https://doi.org/10.1002/bewi.201101504

Sterling, P., & Laughlin, S. B. (2015). *Principles of neural design*. MIT Press.

Stiefel, K. M., & Coggan, J. S. (2023). The energy challenges of artificial superintelligence. *Frontiers in artificial intelligence*, *6*, 1240653. https://doi.org/10.3389/frai.2023.1240653

Thagard, P. (2022). Energy requirements undermine substrate independence and mind-body functionalism. *Philosophy of Science*, *89*(1), 70–88. https://doi.org/10.1017/psa.2021.15

Theriault, J. E., Coleman, M., Feldman, M. J., Fridman, J. D., Sennesh, E., Barrett, L. F., & Quigley, K. S. (2020). Beginning with biology: "Aspects of cognition" exist in the service of the brain's overall function as a resource-regulator. *Behavioral and Brain Sciences*, *43*, 43–44. https://doi.org/10.1017/s0140525x19001705

Theriault, J. E., Shaffer, C., Dienel, G. A., Sander, C. Y., Hooker, J. M., Dickerson, B. C., Barrett, L. F., & Quigley, K. S. (2023). A functional account of stimulation-based aerobic glycolysis and its role in interpreting BOLD signal intensity increases in neuroimaging experiments. *Neuroscience & Biobehavioral Reviews*, *153*, 105373. https://doi.org/10.1016/j.neubiorev.2023.105373

Tsotsos, J., & Rothenstein, A. (2011). Computational models of visual attention. *Scholarpedia*, *6*(1), 6201. https://doi.org/10.4249/scholarpedia.6201

Van Rooij, I. (2008). The tractable cognition thesis. *Cognitive Science*, *32*(6), 939–984. https://doi.org/10.1080/03640210801897856

Wang, T., Ma, S.-Y., Wright, L. G., Onodera, T., Richard, B. C., & McMahon, P. L. (2022). An optical neural network using less than 1 photon per multiplication. *Nature Communications*, *13*(1), 123. https://doi.org/10.1038/s41467-021-27774-8

Williams, D. J. (2025). Two senses of medium independence. *Mind and Language*, *40*(4), 437–447. https://doi.org/10.1111/mila.12542

Woodward, J. (2010). Causation in biology: stability, specificity, and the choice of levels of explanation. *Biology & Philosophy*, *25*(3), 287–318. https://doi.org/10.1007/s10539-010-9200-z

Yarkoni, T., & Westfall, J. (2017). Choosing prediction over explanation in psychology: Lessons from machine learning. *Perspectives on Psychological Science*, *12*(6), 1100–1122. https://doi.org/10.1177/1745691617693393

Zénon, A., Solopchuk, O., & Pezzulo, G. (2019). An information-theoretic perspective on the costs of cognition. *Neuropsychologia*, *123*, 5–18. https://doi.org/10.1016/j.neuropsychologia.2018.09.013

Open Peer Commentary

From "neural activity" to "neural activities": a metabolic framework cultivates an integrative neuroscience by forcing a reexamination of foundational assumptions

Antonietta Alvarez-Hernandez* and Jordan Theriault-Brown

Northeastern University - Boston Campus: Northeastern University, USA,
alvarezhernandez.a@northeastern.edu
jo.theriault@northeastern.edu

*Corresponding author.

doi:10.1017/S0140525X2610627X, e454

Abstract

Haueis and Colaço avoid referencing "neural activity," instead of speaking "neural activities". This subtlety speaks to a strength of their framework: Their metabolically oriented account helps bypass abstractions and stimulus–response assumptions endemic to traditional neuroscience paradigms. We highlight two advantages of this pivot to "neural activities," and how they pave the way for a unified, metabolically grounded, systems neuroscience.

Haueis and Colaço's target article shows that brain metabolism is a critical and often overlooked factor in psychology, neuroscience, philosophy, and artificial intelligence. Among its many strengths, their perspective encourages researchers to get specific about how to integrate brain metabolism – as a source of empirical evidence – into traditional cognitive accounts, either evaluatively constraining them or generatively expanding them. Obviously, this is a long-term project, but even in the short term, it has an important effect: It encourages researchers to interrogate placeholder terms, like

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

"neural activity," which embed assumptions into research paradigms. As a term, "neural activity" is so ubiquitous as to seem innocuous; but it both lumps neural, hemodynamic, and metabolic events into a single construct (Singh, 2012), and embeds a traditional stimulus–response logic at the foundation of neuroscience (Hutchinson & Barrett, 2019; Keller & Mrsic-Flogel, 2018). The proposed metabolic perspective calls this term into question and points to an alternative.

Haueis and Colaço are likely aware that "neural activity" has these problems, as their paper deftly avoids it. What they say instead is worth highlighting: They generally refer to "neural activities." The difference is subtle but important. "Neural activity" implies a distinction between active and inactive brain states (Raichle, 2010, 2015). By contrast, "neural activities" imply that neural populations occupy many states, and that empirical measures – e.g., single unit recordings, local field potentials (LFPs), cerebral blood flow, and oxygen/glucose metabolism – are independent of theoretical interest, as opposed to all being better or worse approximations of an underlying, latent, and singular "neural activity."

One might object, saying that no one actually thinks these measures are interchangeable indices of latent "neural activity", but, in practice, they are often treated this way. This is most apparent in the blood-oxygen level-dependent (BOLD) response, the mainstay measure in fMRI, elicited by events like sensory stimulation or cognitive work. The event-elicited BOLD response is influenced by the ratio of oxygenated to deoxygenated hemoglobin, and BOLD increases are caused by localized increases in blood flow, drawing more oxygenated hemoglobin to the "active" region (Kwong et al., 1992; Ogawa et al., 1990, 1992). In 2001, Logothetis reported a key correlation between event-elicited BOLD increases and LFPs (Logothetis et al., 2001). This work fostered an interpretation of BOLD as an approximate measure of local synaptic transmission, licensing researchers to substitute references to BOLD signal with terms like "neural activity" or "neural processing" – despite cautions by Logothetis that "none of [these] techniques [e.g., BOLD fMRI, single unit recordings, LFPs] is a substitute for the others" (Logothetis, 2008, p. 877). This interpretation even carried over to resting state fMRI, despite evidence that resting state BOLD fluctuations are more weakly correlated with LFP (for review, see Drew, 2019), are caused by a variety of non-neural factors (see Fig. 9 in Haueis, 2017), and support a variety of non-neural functions, heterogeneously distributed throughout the brain (Drew, 2022). The point is that the BOLD signal follows from particular neural, metabolic, and hemodynamic activities, but these important details are lost when it is simplified into abstract "neural activity." This abstraction slots BOLD fMRI into a simple theoretical paradigm (specifically, a stimulus–response paradigm), but halts lines on inquiry that could trigger a paradigm shift.

Refocusing on "neural activities" has at least two advantages. First, it helps address a "cognitive myopia," where neural activities not directly related to cognition are dismissed as noise (Haueis, 2018). The brain is not just for thinking (Theriault et al., 2025): it also "maintains the body and its own structural integrity . . . [with] neural mechanisms that maintain and repair [its parts]" (Haueis, 2017, p. 252). For example, even the event-elicited BOLD response may support cognitive and noncognitive functions. Specifically, event-elicited BOLD increases cooccur with an atypical metabolic shift, a shift that metabolizes glucose nonoxidatively and threatens to acidify "active" brain regions (for review, see Dienel, 2019; Theriault et al., 2023). Beyond delivering oxygen, BOLD hemodynamics may flush lactic acid produced by this nonoxidative metabolic shift, regulating local pH levels within ranges for normal neural firing (DiNuzzo et al., 2024). The event-elicited BOLD increase, then, may not be a pure sign of neuro-cognitive processing; rather, it may, in part, stem from neuroprotective feedback mechanisms that buffering against the consequences of atypical and acidifying metabolic and neural activities.

Second, refocusing on "neural activities" may help contextualize the massive metabolic expense of the resting brain (~20% of the whole-body budget; Clarke & Sokoloff, 1999; Yu et al., 2018). Event-elicited metabolism, by contrast, is a metabolic blip, increasing energy consumption ~5% from baseline (Raichle, 2010). Further, the nonoxidative nature of event-elicited neural activities is distinct from the predominantly oxidative brain metabolism at rest (for review, see Dienel, 2019; Theriault et al., 2023). In other words, the brain does expensive work *all the time*. To understand why it pays this expense, it may be necessary to scrutinize the dominant stimulus–response paradigm (see Cisek, 1999; Dewey, 1896; Lashley, 1951; Logothetis, 2008; Powers, 1973; Yin, 2020) – a paradigm implicit in the project of brain mapping, where stimulus features are mapped to the "neural activity" they elicit (Hutchinson & Barrett, 2019). Alternatives posit that stimuli are context sensitive: their features mean nothing to an organism on their own, because their meaning is codetermined by how the organism is prepared to receive them (Shannon and Weaver (1964[1949]) – i.e., information encoding requires a preexisting neural context. Thus, the massive expense of the resting brain may go toward neural activities that maintain that context, whereas event-elicited neural activities may help propagate bottom-up perturbations of it (see Theriault et al., 2023).

The "neural activity" concept has encouraged stimulus–response paradigms that focus on a fraction of the many neural activities occurring in the brain. Haueis and Colaço's proposal may force neuroscientists to grapple with and resolve the contradictions that appear when "neural activities" are taken at face value. This will be a difficult project, but we believe it will ultimately pave the way for an integrative systems neuroscience, unified through a common metabolic thread.

Financial support. This article was supported by a grant from the National Institute of Mental Health (R21 MH129902).

Competing interests. None.

References

Cisek, P. (1999). Beyond the computer metaphor: Behavior as interaction. *Journal of Consciousness Studies*, *6*, 125–142.

Clarke, D. D., & Sokoloff, L. (1999). Circulation and energy metabolism in the brain. In G. J. Siegel, B. W. Agranoff, R. W. Albers, et al. (Eds.), *Basic neurochemistry: Molecular, cellular and medical aspects* Chapter 31 (6th ed., pp. 637–670). Philadelphia: Lippincott-Raven. https://www.ncbi.nlm.nih.gov/books/NBK20413/.

Dewey, J. (1896). The reflex arc concept in psychology. *Psychological Review*, *3*(4), 357–370. https://doi.org/10.1037/h0070405.

Dienel, G. A. (2019). Brain glucose metabolism: Integration of energetics with function. *Physiological Reviews*, *99*(1), 949–1045. https://doi.org/10.1152/physrev.00062.2017.

DiNuzzo, M., Dienel, G. A., Behar, K. L., Petroff, O. A., Benveniste, H., Hyder, F., Giove, F., Michaeli, S., Mangia, S., Herculano-Houzel, S., & Rothman, D. L. (2024). Neurovascular coupling is optimized to compensate for the increase in proton production from nonoxidative glycolysis and glycogenolysis during brain activation and maintain homeostasis of pH, $p\,CO_2$, and $p\,O_2$. *Journal of Neurochemistry*, *168*(5), 632–662. https://doi.org/10.1111/jnc.15839.

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

Drew, P. J. (2019). Vascular and neural basis of the BOLD signal. *Current Opinion in Neurobiology*, *58*, 61–69. https://doi.org/10.1016/j.conb.2019.06.004.
Drew, P. J. (2022). Neurovascular coupling: Motive unknown. *Trends in Neurosciences*, *45*(11), 809–819. https://doi.org/10.1016/j.tins.2022.08.004.
Haueis, P. (2017). *Meeting the brain on its own terms*. Otto-von-Guericke-Universität.
Haueis, P. (2018). Beyond cognitive myopia: A patchwork approach to the concept of neural function. *Synthese*, *195*, 5373–5402.
Hutchinson, B., & Barrett, L. F. (2019). The power of predictions: An emerging paradigm for psychological research. *Current Directions in Psychological Science*, *28*(3), 280–291. https://doi.org/10.1177/0963721419831992.
Keller, G. B., & Mrsic-Flogel, T. D. (2018). Predictive processing: A canonical cortical computation. *Neuron*, *100*(2), 424–435. https://doi.org/10.1016/j.neuron.2018.10.003.
Kwong, K. K., Belliveau, J. W., Chesler, D. A., Goldberg, I. E., Weisskoff, R. M., Poncelet, B. P., Kennedy, D. N., Hoppel, B. E., Cohen, M. S., & Turner, R. (1992). Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proceedings of the National Academy of Sciences*, *89*(12), 5675–5679. https://doi.org/10.1073/pnas.89.12.5675.
Lashley, K. S. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms inbehavior* (pp. 112–146). New York: Wiley.
Logothetis, N. K. (2008). What we can do and what we cannot do with fMRI. *Nature*, *453*(7197), 869–878. https://doi.org/10.1038/nature06976.
Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., & Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature*, *412*(6843), Article 6843. https://doi.org/10.1038/35084005.
Ogawa, S., Lee, T. M., Kay, A. R., & Tank, D. W. (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences, 87*(24), 9868–9872. https://doi.org/10.1073/pnas.87.24.9868.
Ogawa, S., Tank, D. W., Menon, R., Ellermann, J. M., Kim, S. G., Merkle, H., & Ugurbil, K. (1992). Intrinsic signal changes accompanying sensory stimulation: Functional brain mapping with magnetic resonance imaging. *Proceedings of the National Academy of Sciences*, *89*(13), 5951–5955. https://doi.org/10.1073/pnas.89.13.5951.
Powers, W. T. (1973). Feedback: Beyond behaviorism. *Science*, *179*(4071), 351–356. https://doi.org/10.1126/science.179.4071.351.
Raichle, M. E. (2010). Two views of brain function. *Trends in Cognitive Sciences*, *14*(4), 180–190. https://doi.org/10.1016/j.tics.2010.01.008.
Raichle, M. E. (2015). The restless brain: How intrinsic activity organizes brain function. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1668), 20140172–20140172. https://doi.org/10.1098/rstb.2014.0172.
Shannon, C., & Weaver, W. (1964). *The mathematical theory of communication* (10th ed.). The University of Illinois Press. (Original work published 1949).
Singh, K. D. (2012). Which "neural activity" do you mean? fMRI, MEG, oscillations and neurotransmitters. *NeuroImage*, *62*(2), 1121–1130. https://doi.org/10.1016/j.neuroimage.2012.01.028.
Theriault, J. E., Katsumi, Y., Reimann, H. M., Zhang, J., Deming, P., Dickerson, B. C., Quigley, K. S., & Barrett, L. F. (2025). It's not the thought that counts: Allostasis at the core of brain function. *Neuron*, *113*(24), 4107–4133. https://doi.org/10.1016/j.neuron.2025.09.028.
Theriault, J. E., Shaffer, C., Dienel, G. A., Sander, C. Y., Hooker, J. M., Dickerson, B. C., Barrett, L. F., & Quigley, K. S. (2023). A functional account of stimulation-based aerobic glycolysis and its role in interpreting BOLD signal intensity increases in neuroimaging experiments. *Neuroscience & Biobehavioral Reviews*, *153*, 105373. https://doi.org/10.1016/j.neubiorev.2023.105373.
Yin, H. (2020). The crisis in neuroscience. In W. Mansell (Ed.), *The interdisciplinary handbook of perceptual control theory* (pp. 23–48). Academic Press. https://doi.org/10.1016/b978-0-12-818948-1.00003-4.
Yu, Y., Herman, P., Rothman, D. L., Agarwal, D., & Hyder, F. (2018). Evaluating the gray and white matter energy budgets of human brain function. *Journal of Cerebral Blood Flow and Metabolism*, *38*(8), 1339–1353. https://doi.org/10.1177/0271678X17708691.

What does modeling neural function in its bodily and environmental context buy you?

Matthieu M. de Wit*

Department of Neuroscience, Muhlenberg College, Allentown, USA
matthieudewit@muhlenberg.edu

*Corresponding author.

doi:10.1017/S0140525X26106347, e455

Abstract

I agree with Haueis and Colaço that integrating metabolic considerations into cognitive modeling is important, given evolution by natural selection. However, I question their exclusive focus on neurocentric models, arguing that ecological models of neural function in embodied, situated contexts should be added to the space of possible models and that extraneural metabolic factors are important to consider as well.

Haueis and Colaço (H&C) provide an important corrective to the common tendency to background or ignore metabolic considerations in cognitive modeling. Given evolution by natural selection and the fact that, all else being equal, evolution favors more efficient solutions, metabolic factors should clearly be taken into account if one's goal is to construct biologically realistic models of cognition (Niven & Laughlin, 2008). Indeed, H&C discuss the relevance of evolution in several places (Sections 2.1.2 and 2.2.2).

It is therefore surprising that H&C focus only on the evaluation and generation of exclusively neural models. That is, nervous systems did not evolve in isolation but in the context of organisms that were already embodied and situated in structured environmental niches, and ignoring these facts is bound to distort our understanding of how behavior and cognition come about (Chiel & Beer, 1997; Jékely, Keijzer, & Godfrey-Smith, 2015). H&C furthermore focus, without substantive argument, only on computational models, in which cognition is characterized as the construction, processing, and transmission of information.

H&C's discussion of the implications of metabolic considerations in philosophical debates about multiple realizability could be understood to leave open the possibility of non-neural realizers of cognition (e.g., "The metabolic demands of the brain differ from the energy consumption of *other systems*," p. 14, emphasis added) but is not explicitly framed as such. Alternatives to computationalist cognitive modeling are mentioned but only in passing (Section 2.1.1), and the suggestion is made that the specific ontology of models (e.g., computational vs. dynamical) is not particularly relevant to their arguments (see also p. 15). I think that, on the contrary, this distinction is in fact crucial to evaluating and generating cognitive models in light of metabolic considerations.

Ecological psychology as well as related approaches like enactivism (Chemero, 2009) and its more recent offspring, ecological neuroscience, take a rather different approach to cognitive modeling (de Wit, 2024; Falandays et al., 2024; Favela, 2024). One motivation behind these approaches is the assumption that brain-body-environment-spanning solutions to cognitive problems are likely to be energetically cheaper than exclusively neural solutions

and therefore favored during natural selection whenever possible (Anderson, Richardson, & Chemero, 2012; Barrett, 2011; Gibson, 1966; Reed, 1996). Some or all of the information needed to perform a particular cognitive capacity may be found in organisms' richly structured environments and made available through bodily action and thus need not be reconstructed internally (Warren, 2021). Exploitation of a species' particular embodiment is another way in which the performance of a cognitive capacity can be simplified or even solved (Chiel et al., 2009; Pfeifer, Iida, & Lungarella, 2014).

In H&C's taxonomy of how metabolic considerations can be incorporated, from not addressing them at all to estimating the specific numeric values of the energetic costs associated with a model of a particular cognitive capacity, researchers in the ecological tradition can be classified as presenting substantive arguments for the importance of metabolic considerations when modeling cognitive capacities but otherwise largely backgrounding metabolism. In order to strengthen the competitiveness of ecological models relative to neurocentric, computational models, it would be beneficial to heed H&C's call and move from logical argument to an explicit integration of metabolic considerations in as specific a sense as possible (as discussed in Section 2.1.5). H&C provide clear guidelines for evaluative metabolic considerations in neurocentric, computational modeling. What would such an enterprise look like for ecological modeling?

Here, I can only speculate and gesture at some potentially relevant prior work. Collins et al. (2005) compared the efficiency of human walking with passive dynamic walkers, which solve walking entirely via interactions between body morphology and environment, and with Honda's ASIMO robot, which solves it via central control of joint angles. They found remarkable similarity between human walking and passive dynamic walking, both of which were an order of magnitude more efficient than ASIMO. MacIver, Patankar, and Shirgaonkar (2010) showed in electric fish that reducing swimming efficiency to improve sensing increases prey encounter rates, which more than makes up for the additional energy expended in swimming inefficiently. Finally, as with neural tissue, there are also methods for estimating metabolic cost in non-neural tissue, including skeletal muscle tissue (Javed et al., 2010) and even bone (Zhang, Modlesky, & McCully, 2021).

I end by raising two problems for cognitive modeling as a whole when broadening the space of possible models to also include more thoroughly evolutionary-informed, ecological models. First, for a variety of reasons, evolution by natural selection need not always produce perfectly optimal solutions (Withagen & Chemero, 2009; Withagen, de Wit, & van der Kamp, 2018). This complicates model evaluation. It is important to develop additional evaluative criteria that are sensitive to the possibility that a less efficient (but in theory metabolically viable) model may nevertheless be more plausible than a more efficient model. Second, once exclusively neural and ecological models are both included in the space of possible models, it will be important to develop ways to directly compare these model classes with very different metabolic considerations. For example, metabolic considerations related to the cost of deploying non-neural tissue in support of a cognitive capacity will by definition not be present in neurocentric models of cognition. Conversely, the number of bits sent through a neural pathway per ATP used will not be a meaningful metric for ecological models in which environmental ecological information is generated and detected via bodily action rather than reconstructed via internal processing.

Financial support. This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

Competing interests. None.

References

Anderson, M. L., Richardson, M. J., & Chemero, A. (2012). Eroding the boundaries of cognition: Implications of embodiment. *Topics in Cognitive Science*, *4*(4), 717–730. https://doi.org/10.1111/j.1756-8765.2012.01211.x

Barrett, L. (2011). *Beyond the brain: How body and environment shape animal and human minds*. Princeton University Press.

Chemero, A. (2009). *Radical embodied cognitive science*. MIT Press.

Chiel, H. J., & Beer, R. D. (1997). The brain has a body: Adaptive behavior emerges from interactions of nervous system, body and environment. *Trends in Neurosciences*, *20*(12), 553–557. https://doi.org/10.1016/S0166-2236(97)01149-1

Chiel, H. J., Ting, L. H., Ekeberg, Ö., & Hartmann, M. J. Z. (2009). The brain in its body: Motor control and sensing in a biomechanical context. *Journal of Neuroscience*, *29*(41), 12807–12814. https://doi.org/10.1523/JNEUROSCI.3338-09.2009

Collins, S., Ruina, A., Tedrake, R., & Wisse, M. (2005). Efficient bipedal robots based on passive-dynamic walkers. *Science*, *307*(5712), 1082–1085. https://doi.org/10.1126/science.1107799

de Wit, M. M. (2024). Separating minimal from radical embodied cognitive neuroscience. *Cognitive Neuroscience*, *15*(3–4), 102–103. https://doi.org/10.1080/17588928.2024.2405591

Falandays, J. B., Yoshimi, J., Warren, W. H., & Spivey, M. J. (2024). A potential mechanism for Gibsonian resonance: Behavioral entrainment emerges from local homeostasis in an unsupervised reservoir network. *Cognitive Neurodynamics*. *18*(4), 1811–1834. https://doi.org/10.1007/s11571-023-09988-2

Favela, L. H. (2024). *The ecological brain: Unifying the sciences of brain, body, and environment*. Routledge.

Gibson, J. J. (1966). *The senses considered as perceptual systems*. Houghton Mifflin.

Javed, F., He, Q., Davidson, L. E., Thornton, J. C., Albu, J., Boxt, L., Krasnow, N., Elia, M., Kang, P., Heshka, S., & Gallagher, D. (2010). Brain and high metabolic rate organ mass: Contributions to resting energy expenditure beyond fat-free mass. *The American Journal of Clinical Nutrition*, *91*(4), 907–912. https://doi.org/10.3945/ajcn.2009.28512

Jékely, G., Keijzer, F., & Godfrey-Smith, P. (2015). An option space for early neural evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1684), 20150181. https://doi.org/10.1098/rstb.2015.0181

MacIver, M. A., Patankar, N. A., & Shirgaonkar, A. A. (2010). Energy-information trade-offs between movement and sensing. *PLOS Computational Biology*, *6*(5), e1000769. https://doi.org/10.1371/journal.pcbi.1000769

Niven, J. E., & Laughlin, S. B. (2008). Energy limitation as a selective pressure on the evolution of sensory systems. *Journal of Experimental Biology*, *211*(11), 1792–1804. https://doi.org/10.1242/jeb.017574

Pfeifer, R., Iida, F., & Lungarella, M. (2014). Cognition from the bottom up: On biological inspiration, body morphology, and soft materials. *Trends in Cognitive Sciences*, *18*(8), 404–413. https://doi.org/10.1016/j.tics.2014.04.004

Reed, E. S. (1996). *Encountering the world: Toward an ecological psychology*. Oxford University Press.

Warren, W. H. (2021). Information is where you find it: Perception as an ecologically well-posed problem. *I-Perception*, *12*(2), 1–24. https://doi.org/10.1177/20416695211000366

Withagen, R., & Chemero, A. (2009). Naturalizing perception: Developing the Gibsonian approach to perception along evolutionary lines. *Theory & Psychology*, *19*(3), 363–389. https://doi.org/10.1177/0959354309104159

Withagen, R., van der Kamp, J., & de Wit, M. (2018). Suboptimalities for sure: Arguments from evolutionary theory. *Behavioral and Brain Sciences*, *41*, e247. https://doi.org/10.1017/S0140525X18001322

Zhang, C., Modlesky, C. M., & McCully, K. K. (2021). Measuring tibial hemodynamics and metabolism at rest and after exercise using near-infrared spectroscopy. *Applied Physiology, Nutrition, and Metabolism*, *46*(11), 1354–1362. https://doi.org/10.1139/apnm-2021-0135

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

Bridging metabolic-cognitive modelling: from synaptic cost to functional brain organisation

Hamish A. Deery[a,b], Rui Zheng[a,b] and Sharna D. Jamadar[a,b*]

[a]School of Psychological Sciences, Monash University, Melbourne, VIC 3800, Australia and [b]Monash Biomedical Imaging, Monash University, Melbourne, VIC 3800, Australia
hamish.deery@monash.edu
rui.zheng@monash.edu
sharna.jamadar@monash.edu

*Corresponding author.

doi:10.1017/S0140525X2610630X, e456

Abstract

Haueis and Colaço argue that metabolism is a generative principle for cognitive modelling. We extend this by bridging the "scaling gap" through functional PET (fPET) and network neuroscience. By measuring glucodynamics and the metabolic network of the brain, we can model how glucose metabolism enables network efficiency, temporal flexibility, information transfer, and ultimately the cognitive capacity of the human brain.

Haueis & Colaço's thesis brings welcome renewed attention to the importance of the energetic constraints of neural computation to the field of cognitive neuroscience. Here, we argue that the conceptual framework can be strengthened by considering a detailed network neuroscience perspective. To transform metabolic theory into a testable science, the following four pillars should be addressed.

The scaling gap: from synapses to network neuroscience

A primary challenge for the authors' framework is the "scaling gap" between microscopic metabolic estimates and macroscale neural signals. Haueis & Colaço's thesis relies heavily on bottom-up costs, such as the estimated 24,000 ATP/bit required for synaptic transmission. While these values are foundational, it remains unclear how they translate to metabolism measured in the human brain *in vivo* at the macroscale, such as regional cerebral metabolic rates of glucose (CMR_{GLC}), which reflect the aggregate activity of millions of neurones (Sokoloff, 1960).

There is a critical need to bridge local metabolic signals to the large-scale functional organisation of the brain. Cognition, affect, and behaviour emerge from the coherent activity of spatially distributed populations of neurones (Raichle et al., 2001; Yeo et al., 2011). This coherent and spatially distinct activity has primarily been modelled as a complex network of units (nodes or brain regions) and their interactions (edges or paths) using fMRI and M/EEG. Graph theory is often used to model how information is transferred throughout the system – the mechanism by which cognition, behaviour, and affect are thought to emerge (van den Heuvel & Sporns, 2011). Great strides have been made in mapping the haemodynamic (fMRI) and electromagnetic (M/EEG) characteristics of the functional brain network, but until recently, technical limitations have stymied the capacity to map the metabolic characteristics of this network within individuals (Sala et al., 2023). Advances in [^{18}F]-FDG-PET acquisition and modelling (Hahn et al., 2016; Jamadar et al., 2019; Villien et al., 2014; Volpi et al., 2021) now make it possible to characterise the macroscale metabolic network of the brain (Jamadar et al., 2021; Reed et al., 2023; Volpi et al., 2023). These methods provide the necessary "bridge" between micro- and macro-scale metabolic function, revealing how local glucose metabolism constrains macroscale organisation and how topological properties, such as network efficiency, directly relate to cognition (Deery et al., 2024; Deery et al., 2026).

Accounting for spatial and temporal variability in glucose metabolism

Models must further account for the fact that cerebral metabolism is not uniform across space nor time. Glucose is a rate-limiting substrate that sets biophysical limits on *when* and *where* neural activity occurs (Jamadar et al., 2025). As a result of these limits, the brain must operate as a highly responsive and scalable system that supports rapid neuronal activity. Glucose metabolism itself fluctuates as a dynamic signal – which we have termed "glucodynamics" – reflecting the time-variant energetic demands of underlying neuronal activity (Deery et al., 2026; Hahn et al., 2024; Hahn et al., 2020; Jamadar et al., 2025). Metabolic rates and glucodynamic signals vary across the cortex, shaped by the underlying cytoarchitecture, gene expression, structural constraints and neural functions (Deery et al., 2026; Jamadar et al., 2025). Greater glucodynamic variability and complexity determine the efficiency of information flow across the metabolic network of the brain, enable it to transition between a rich repertoire of network states, and are associated with more adaptive cognition (Deery et al., 2025). Cognitive models must therefore integrate these spatial *and* temporal glucodynamics to understand how the brain dynamically allocates resources to meet shifting internal and external demands.

Scaling information: from bits to complex cognitive processes

Haueis and Colaço define information rate in terms of bits per second: "one bit is the information needed to decide between two equally likely alternatives." While precise for signalling budgets, to be generative, models must quantify the cost or efficiency of information flow for complex operations, such as prediction error, executive control or memory binding. These complex operations emerge from parallel processing across distributed brain networks (Edelman, 1993; Goldman and Rakic, 1988; Mesulam, 1990) and cannot easily be reduced to their component parts (Miller et al., 2024), including bits per information unit processed. To truly understand the energetic demands of cognition, we must find a way to bridge the microscale energetic demands of single-unit firing to the macroscale coordinated dynamics from which complex mental operations emerge. One way to do this could be to develop biophysically informed models of the relationship between microscale and macroscale energetic demands. Such models, which provide a conceptual framework for bridging the micro to macroscale dynamics, exist for fMRI and M/EEG (Friston, Harrison, & Penny, 2003). While causal models do not currently exist for FDG-PET data, the field is well-positioned to develop these, given the elegant kinetic models that exist for FDG data

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

(Hahn et al., 2025; Reivich et al., 1985; Sokoloff et al., 1977). In the meantime, some insight can be gained from non-biophysically informed causal models, such as Granger causality (Deery et al., 2026). This shifts the focus from abstract energy-per-bit to the relative energetic capacity required to sustain and dynamically reconfigure neural systems within and across cognitive systems.

Energy as an enabler: optimisation, trade-offs, and alterations in health and disease

Haueis and Colaço argue that metabolism should not be treated merely as a background constraint but as a generative principle shaping brain structure and cognitive processes. We agree and emphasise that metabolism should be viewed as an *enabler* that actively structures large-scale brain network organisation and cognition. The brain is organised as a cost-efficient network that balances the metabolic expense of long-range integration against the economy of local segregation (Bullmore and Sporns, 2012). Evolution prioritises "high-cost, high-return" infrastructure, such as metabolic hubs and long-range connections, to facilitate the global integration required for complex behaviour (Barron, Halina, & Klein, 2023). This balance is an evolutionary "design choice": inter-species differences in brain architecture reflect different optimisations of the energy-information trade-off. For instance, axon diameter and myelin thickness affect the conduction velocity of action potentials and the "rate-efficiency" of information transfer and vary across species (Gong et al., 2025; Perge et al., 2009). From this perspective, states of altered brain metabolism and metabolic network topology (e.g., in ageing and neurodegeneration) are not merely "energy shortages" but perturbations that impair the brain's capacity to maintain these optimal, high-cost network states. Energetic cognitive models must be flexible to account for how these design trade-offs are made or break down and to establish states in which altered energetic supply, use, and dynamics lead to diminished network flexibility and poor functional outcomes.

Conclusion

We welcome Haueis and Colaço's position for elevating metabolism to a central role in cognitive theory. Because cognition is an emergent property of coherent and parallel firing of populations of neurones (Edelman, 1993; Goldman and Rakic, 1988), we argue that to properly characterise cognition, such models need to also account for macroscale glucodynamics. By bridging the scaling gap through macroscale neuroimaging approaches and recognising energy as a dynamic enabler of network organisation and information transfer, we can move toward a more mechanistic account of how energy shapes the complexity and efficiency of the human mind.

Financial support. Jamadar is supported by an Australian Research Council Future Fellowship FT250100206.

Competing interests. The authors declare no competing interests.

References

Bullmore, E., & Sporns, O., (2012). The economy of brain network organization. *Nature Reviews Neuroscience*, *13*(5), 336–49.

Barron, A. B., Halina, M., & Klein, C. (2023). Transitions in cognitive evolution. *Proceedings of the Royal Society B: Biological Sciences*, *290*(2002), 20230671.

Deery, H. A., Liang, E., Moran, C, Egan, G. F., Jamada, S. D., (2026). *Directed information flow across the metabolic network of the human brain.* bioRxiv.

Deery, H. A., Moran, C., Liang, E. X., Egan, G. F., & Jamadar, S. D., (2026). Metabolic Connectivity Gradients of the Human Brain. BiorXiv.

Deery, H.A., Liang, E., Moran, C., Egan, G.F., & Jamadar, S. D., (2025). Glucodynamic network transitions of the human brain. BiorXiv.

Deery, H. A., Liang, E. X., Siddiqui, N. M., Murray, G., Voigt, K., Paolo R. D., Moran C., Egan, G. F., & Jamadar, S. D. (2024). Reconfiguration of metabolic connectivity in ageing. *Communications Biology*, *7*(1), 1600.

Deery, H. A., Liang E., Moran, C., Egan, G. F., & Jamadar, S. D. (2026). Brain glucodynamic variability is an essential feature of the metabolism-cognition relationship. *Proceedings of the National Academy of Sciences of the United States of America*, *123*(3), e2510850123.

Edelman, G. M., (1993). Neural Darwinism: Selection and reentrant signaling in higher brain function. *Neuron*, *10*(2), 115–25.

Friston, K. J., Harrison, L., & Penny, W., (2003). Dynamic causal modelling. *Neuroimage*, *19*(4), 1273–302.

Goldman-Rakic, P. S., (1988). Topography of cognition: Parallel distributed networks in primate association cortex. *Annual Review of Neuroscience*, *11*, 137–56.

Gong, T., Maffei, C., Dann, E., Lee, H., Lee, H., Augustinack, J. C., Huang, S. Y., Haber, S. N., & Yendiki, A. (2025). Interplay between MRI-based axon diameter and myelination estimates in macaque and human brain. *Imaging Neuroscience*, *3*, imag_a_00576.

Hahn, A., Gryglewski, G., Nics, L., Hienert, M., Rischka, L., Vraka, C., Sigurdardottir, H., Vanicek, T., James, G. M., Seiger, R., Kautzky, A., Silberbauer, L., Wadsak, W., Mitterhauser, M., Hacker, M., Kasper, S., & Lanzenberger, R. (2016). Quantification of task-specific glucose metabolism with constant infusion of 18F-FDG. *Journal of Nuclear Medicine*, *57*(12), 1933–1940.

Hahn, A., Matej Murgas, p. F., Klug, S., Reed, M., Godbersen, G., de Leon, M. P., Milz, C., Hacker, M., Rujescu, D., Lanzenberger, R., (2025). *[18F] FDG Revisited: A New Perspective on the Temporal Dynamics of Brain Glucose Metabolism.*

Hahn, A., Reed, M. B., Vraka, C., Godbersen, G. M., Klug, S., Komorowski, A., Falb, P., Nics, L., Traub-Weidinger, T., Hacker, M., & Lanzenberger, R. (2024). High-temporal resolution functional PET/MRI reveals coupling between human metabolic and hemodynamic brain response. *European Journal of Nuclear Medicine and Molecular Imaging*, *51*(5), 1310–1322.

Hahn, A., Breakspear, M., Rischka, L., Wadsak, W., Godbersen, G. M., Pichler, V., Michenthaler, P., Vanicek, T., Hacker, M., Kasper, S., Lanzenberger, R., & Cocchi, L. (2020). Reconfiguration of functional brain networks and metabolic cost converge during task performance. *eLife*, *9*, e52443.

Jamadar, S. D., Behler, A., Deery, H., & Breakspear, M. (2025). The metabolic costs of cognition. *Trends in Cognitive Sciences*, *29*(6), 541–555.

Jamadar, S. D., Zhong, S., Carey, A., McIntyre, R., Ward, P. G. D., Fornito, A., Premaratne, M., Shah, N. J., O'Brien, K., Stäb, D., Chen, Z., & Egan, G. F. (2021). Task-evoked simultaneous FDG-PET and fMRI data for measurement of neural metabolism in the human visual cortex. *Scientific Data*, *8*(1), 267.

Jamadar, S. D., Ward, P. G., Li, S., Sforazzini, F., Baran, J., Chen, Z., & Egan, G. F. (2019). Simultaneous task-based BOLD-fMRI and [18-F] FDG functional PET for measurement of neuronal metabolism in the human visual cortex. *Neuroimage*, *189*, 258–266.

Mesulam, M. M., (1990). Large-scale neurocognitive networks and distributed processing for attention, language, and memory. *Annals of Neurology*, *28*(5), 597–613.

Miller, E. K., Brincat, S. L., & Roy, J. E. (2024). Cognition is an emergent property. *Current Opinion in Behavioral Sciences*, *57*, 101388.

Perge, J. A., Koch, K., Miller, R., Sterling, P., & Balasubramanian, V. (2009). How the optic nerve allocates space, energy capacity, and information. *The Journal of Neuroscience*, *29*(24), 7917–28.

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(2), 676–82.

Reivich, M., Alavi, A., Wolf, A., Fowler, J., Russell, J., Arnett, C., MacGregor, R. R., Shiue, C. Y., Atkins, H., & Anand, A. (1985). Glucose metabolic rate kinetic model parameter determination in humans: The lumped constants and rate constants for [18F]fluorodeoxyglucose and [11C]deoxyglucose. *Journal of Cerebral Blood Flow & Metabolism*, *5*(2), 179–92.

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

Reed, M. B., Ponce de León, M., Vraka, C., Rausch, I., Godbersen, G. M., Popper, V., Geist, B. K., Komorowski, A., Nics, L., Schmidt, C., Klug, S., Langsteger, W., Karanikas, G., Traub-Weidinger, T., Hahn, A., Lanzenberger, R., & Hacker, M. (2023). Whole-body metabolic connectivity framework with functional PET. *Neuroimage*, *271*, 120030.

Sala, A., Lizarraga, A., Caminiti, S. P., Calhoun, V. D., Eickhoff, S. B., Habeck, C., Jamadar, S. D., Perani, D., Pereira, J. B., Veronese, M., & Yakushev, I. (2023). Brain connectomics: Time for a molecular imaging perspective? *Trends in Cognitive Sciences*, *27*(4), 353–366.

Sokoloff, L., (1960). *The metabolism of the central nervous system in vivo.* Handbook of Physiology, Section I, Neurophysiology, 3, 1843–1864.

Sokoloff, L., Reivich, M., Kennedy, C., Des Rosiers, M. H., Patlak, C. S., Pettigrew, K. D., Sakurada, O., & Shinohara, M. (1977). The [14C]deoxyglucose method for the measurement of local cerebral glucose utilization: Theory, procedure, and normal values in the conscious and anesthetized albino rat. *Journal of Neurochemistry*, *28*(5), 897–916.

van den Heuvel, M. P., & Sporns, O. (2011). Rich-club organization of the human connectome. *The Journal of Neuroscience*, *31*(44), 15775–86.

Villien, M., Wey, H. Y., Mandeville, J. B., Catana, C., Polimeni, J. R., Sander, C. Y., Zürcher, N. R., Chonde, D. B., Fowler, J. S., Rosen, B. R., & Hooker, J. M. (2014). Dynamic functional imaging of brain glucose utilization using fPET-FDG. *Neuroimage*, *100*, 192–9.

Volpi, T., Silvestri, E., Corbetta, M., & Bertoldo, A. (2021). Assessing different approaches to estimate single-subject metabolic connectivity from dynamic [(18)F]fluorodeoxyglucose Positron Emission Tomography data. *Annual International Conference IEEE Engineering in Medicine and Biology Society*, *2021*, 3259–3262.

Volpi, T., Vallini, G., Silvestri, E., Francisci, M., Durbin, T., Corbetta, M., Lee, J. J., Vlassenko, A. G., Goyal, M. S., & Bertoldo, A. (2023). A new framework for metabolic connectivity mapping using bolus [(18)F]FDG PET and kinetic modeling. *Journal of Cerebral Blood Flow & Metabolism*, *43*(11), 1905–1918.

Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., Roffman, J. L., Smoller, J. W., Zöllei, L., Polimeni, J. R., Fischl, B., Liu, H., & Buckner, R. L. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, *106*(3), 1125–65.

The metabolic turn: Implications between the model and the system

Mads J. Dengsø*

University of Murcia: Universidad de Murcia, Spain
madsdengsoe@gmail.com

*Corresponding author.

doi:10.1017/S0140525X26106189, e457

Abstract

I offer further support for a *metabolic turn* in cognitive science and raise two questions for metabolic considerations in modeling: I first invite the authors to consider broader tensions between computational and metabolic models; I then discuss possible connections between the relevance attributed to metabolic considerations for modeling and ontological background commitments about cognitive systems.

This commentary will focus on Haueis and Colaço's notion of a *metabolic turn* in cognitive neuroscience and its broader implications for modeling cognitive systems. I find the core thesis of the relevance of metabolic considerations for cognitive modeling convincing, and my intention here is to provide an opportunity for further elaborating on this central thesis. To begin with, I will briefly present some potential compatriots of the metabolic turn in the cognitive sciences and the philosophy of biology.

In their target article, Haueis and Colaço note a growing metabolic turn in philosophy aligned with their core thesis (TA 2.2.3). But while the authors provide several illustrative examples of such metabolic approaches from cognitive neuroscience, less attention is given to existing literature on metabolism in philosophy of biology and cognitive science, much of which might add further support for their case.

Research from the philosophy of biology highlights the importance of metabolism in accounting for the development and functional role of various phenotypes – including cognitive capacities. Multiple accounts here align with the target paper in regarding metabolic considerations as critical to cognitive systems, their organization (Baedke, 2025; Bich & Bechtel, 2026; Fábregas-Tejeda et al., 2024; Nicholson, 2014; Walsh, 2015), development, and evolution of cognition (Ginsburg & Jablonka, 2021; Moreno & Peretó, 2026; Oyama, Griffiths, & Gray, 2001). Other accounts argue that metabolic systems may even be primary to substantial biological individuality (Dupré, 2024; Jefferey, Pollack, & Rovelli, 2019; Nicholson & Dupré, 2018). Additionally, biologically informed approaches in cognitive science likewise highlight interdependencies between cognition and broader organismic dynamics, such as metabolic processes (Barandiaran & Moreno, 2008; Huebner & Schulkin, 2022; Lyon et al., 2021). These various biologically informed approaches to cognitive systems strike me as conducive to the overall line of argument of the target article.

In addition to general support, these accounts also suggest a potential explanation for the neglect of metabolic considerations documented by the target article. Dynamical models based on biological and metabolic considerations, including ones listed above, are often seen as conflicting with standard computational models (Brette, 2022; Beer, 2023; Lyon, 2020). Indeed, some of the vocal supporters of medium dependence within the philosophy of neuroscientific modeling (TA 4.2) have recently argued that standard computationalist approaches encourage viewing biological features "as background support structures that are separable from the truly cognitive (i.e. computational) processes of the brain" (Chirimuuta, 2026, p. 4), suggesting there may be a conflict between metabolic considerations and standard computational modeling.

This brings me to my first question. While the authors seem to support the indivisibility of metabolism from cognition, they also reject the conflict between resource models and computational models (TA 4.2). *How does this rejection of conflict apply to wholesale critiques of computational approaches, such as claims that computationalism inappropriately divides metabolic hardware from cognitive software?* For example, are such wholesale contrasts overstating the limitations of computational analogues in modeling or the ramifications of a metabolic turn?

On the topic of how deep a metabolic turn in cognitive modeling should go, my second question concerns the potential connection between modeling considerations and ontological commitments. Throughout the target article, Haueis and Colaço

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

deliberately opt for neutrality about ontology, regarding the ontological status of metabolic considerations as secondary (TA 2.1.1, 2.2.3). They also note that, while metabolic considerations may contribute to cognitive modeling across the spectrum, determining the relevance of metabolic considerations for modeling is a subject to be explored (TA 4.1). While this choice of focus is quite reasonable, I want to raise the possibility that determining the relevance of various metabolic considerations may be inextricable from questions about our background ontological commitments about cognitive systems.

As an example, consider the conspicuous connection between computationalism and medium independence. As demonstrated by various accounts above, metabolically informed approaches often highlight the importance of accounting for a wider array of neurobiological factors as well as broader, organismic factors. In this sense, advocates of a metabolic turn tend to converge on an identification of cognitive systems with a wider array of substrate processes. By contrast, backgrounding of metabolic considerations often converges with more constrained conceptualizations of cognitive systems (Barrett, 2016; Hurley, 2001).

This raises the potential connection between the attribution of relevance in cognitive modeling, such as the backgrounding of metabolism (TA 2.1.5), and ontological commitments, such as a narrower conceptualization of the cognitive system. As some philosophical work of modeling argues, models may be importantly constrained by background assumptions, including ontological commitments about which processes and parts form part of the system in question (Longino, 2021; Potochnik & Sanches de Oliveira, 2020). If this is the case, ontological considerations may be unavoidable in specifying the relevance of various evaluative and generative metabolic considerations for modeling cognitive systems.

My second question is, *do the authors see any connection between ontological commitments and future work in determining the relevance of metabolic considerations for modeling?* For example, might ontological assumptions have influenced the present neglect of metabolic considerations in modeling cognitive systems? And could a metabolic turn in cognitive modeling potentially involve reconsidering some of our ontological background assumptions about cognitive systems?

Financial support. This article was funded by the Agencia Estatal de Investigación (AEI), Spain, Juan de la Cierva program (grant number JDC2024-053986-I).

Competing interests. Author declares no competing interests.

References

Baedke J. (2025). *The organism*. Cambridge University Press.

Barandiaran X., & Moreno A. (2008). Adaptivity: From metabolism to behavior. *Adaptive Behavior 16*(5), 325–344. https://doi.org/10.1177/1059712308093868

Barrett, L. (2016). Why brains are not computers, why behaviorism is not satanism, and why dolphins are not aquatic apes. *The Behavior Analyst*, *39*(1), 9–23.

Beer, R. D. (2023). On the proper treatment of dynamics in cognitive science. *Topics in Cognitive Science*, 1–14. Advance online publication. https://doi.org/10.1111/tops.12686

Bich, L., & Bechtel, W. (2026). Autonomy and heterarchy: Organizing control in biological organisms. In X. E. Barandiaran & A. Etxeberria (Eds.), *Outonomy: Fleshing out the concept of autonomy beyond the individual* (pp. 23–32). Springer Nature Switzerland.

Brette, R. (2022). Brains as computers: Metaphor, analogy, theory or fact?. *Frontiers in Ecology and Evolution*, *10*, 878729.

Chirimuuta, M. (2026). Précis of the brain abstracted. *Philosophy and the Mind Sciences*, *7*(1), 1–8.

Dupré, J. (2024). Speciation and species: A process perspective, *Evolutionary Journal of the Linnean Society*, *3*(1), 1–4. https://doi.org/10.1093/evolinnean/kzae020

Fábregas-Tejeda, A., Baedke, J., Prieto, G. I., & Radick, G. (2024). Organismal agency: A persistent riddle in the history and philosophy of the life sciences. In A. Fábregas-Tejeda, J. Baedke, G. I. Prieto, & G. Radick (Eds.), *The riddle of organismal agency* (pp. 1–20). Routledge.

Ginsburg, S, & Jablonka, E. (2021). Evolutionary transitions in learning and cognition. *Philosophical Transactions of the Royal Society B*, *376*, 20190766. https://doi.org/10.1098/rstb.2019.0766

Huebner, B., & Schulkin, J. (2022). *Biological cognition*. Cambridge University Press.

Hurley, S. (2001). Perception and action: Alternative views. *Synthese*, *129*(1), 3–40.

Jeffery, K., Pollack, R., & Rovelli, C. (2019). On the statistical mechanics of life: Schrödinger revisited. *Entropy*, *21*(12), 1211. https://doi.org/10.3390/e21121211

Longino, H. E. (2021). Scaling up; scaling down: What's missing?. *Synthese*, *198*(4), 2849–2863.

Lyon, P. (2020). Of what is "minimal cognition" the half-baked version? *Adaptive Behavior*, *28*(6), 407–424. https://doi.org/10.1177/1059712319871360

Lyon, P., Keijzer, F., Arendt, D., & Levin, M. (2021). Reframing cognition: Getting down to biological basics. *Philosophical Transactions of the Royal Society B*, *376*, 20190750. https://doi.org/10.1098/rstb.2019.0750

Moreno, Á., & Peretó, J. (2026). *An evolutionary story of agency: How life evolved to act on its own*. Springer Nature, p. 108.

Nicholson, D. J. (2014). The return of the organism as a fundamental explanatory concept in biology. *Philosophy Compass*, *9*(5), 347–359.

Nicholson, D. J., & Dupré, J. (2018). *Everything flows: Towards a processual philosophy of biology*. Oxford University Press.

Oyama, S., Griffiths, P. E., & Gray, R. D. (2001). Introduction: What is developmental systems theory? In S. Oyama, P. E. Griffiths, & R. D. Gray (Eds.), *Cycles of contingency: Developmental systems and evolution* (pp. 1–11). MIT Press.

Potochnik, A., & Sanches de Oliveira, G. (2020). Patterns in cognitive phenomena and pluralism of explanatory styles. *Topics in Cognitive Science*, *12*, 1306–1320. https://doi.org/10.1111/tops.12481

Walsh, D. M. (2015). *Organisms, agency, and evolution*. Cambridge University Press.

Metabolic constraints may both help and hinder cognitive models

Jake R. Embrey*

University of Chicago Booth School of Business, USA
jake.embrey@chicagobooth.edu

*Corresponding author.

doi:10.1017/S0140525X2610613X, e458

Abstract

Constraining cognitive models based on the metabolic realities of the brain is a laudable goal. However, imposing strong constraints on nascent theories risks prematurely discarding potentially productive models. Furthermore, despite models of cognitive fatigue being an intuitive test case for Haueis and Colaço's approach, the strictly medium-independent nature of non-resource models (e.g., opportunity cost models) highlights some limits to the proposed approach.

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

Metabolic feasibility is important. If our goal is to understand cognition, a biologically plausible model will always trump an infeasible alternative. The usefulness of this metabolic-centric approach, however, is best applied to fewer theories than proposed (i.e., all cognitive models), as there are some questions in cognitive science not *yet* aided by considering metabolic constraints. All sufficiently advanced models must eventually address the question of how the brain "does it," but forcing nascent theories to seriously consider their neural realisability effectively collapses Marr's levels of analysis (1982). Despite agreeing with the general framework, I discuss two examples that highlight some limitations of the proposed approach: (1) models of mental fatigue and (2) models where premature metabolic constraints may impede progress.

Models of cognitive fatigue

Much research on mental fatigue has focused on identifying neurometabolic substrates, such as neurotoxins that accumulate during effort (e.g., beta-amyloid and glutamate, see Holroyd, 2024; Wiehler et al., 2022) or depleting fuel sources (e.g., blood glucose; see Gailliot & Baumeister, 2007). A parallel line of research abstracts away from the brain, seeking to explain the cost of effort via motivational (Inzlicht, Schmeichel, & Macrae, 2014) or computational considerations (opportunity costs; Kurzban et al., 2013). Haueis & Colaço argue that *all* theories of cognitive effort – including those that are putatively indifferent to biological realisability – should be subject to metabolic constraints. For instance, the authors argue that opportunity cost models should consider the metabolic costs incurred by the brain when calculating opportunity costs during task engagement.

This conflates the question of whether a model is realisable in the brain with whether the brain's metabolism is relevant to the phenomenon being theorised. It is unreasonable that theories which argue energy expenditure is irrelevant to the experience of effort should themselves be subject to metabolic constraints. Consider the case of opportunity costs. While the brain is responsible for these calculations, the physical underpinnings of fatigue, according to these theories, are irrelevant to the reason fatigue arises. If fatigue is a signal that indicates one's attention is better spent elsewhere, it does not matter whether the attention being spent is that of a human being (using neurons) or a silicon-based computer; all that is required is that central processing is limited, that opportunity costs can be roughly estimated by the agent, and that opportunity costs give rise to fatigue. The theory is strictly medium independent.

The majority of research informed by such theories is presently concerned with whether opportunity costs are at all an input to mental effort costs (Dora et al., 2022; Embrey et al., 2025; Otto & Daw, 2019). If this turns out to be the case, then further work can be invested in determining the algorithms used to calculate opportunity costs and how such algorithms may be implemented in the brain; although this concern seems minor, given *roughly* estimating opportunity costs is computationally trivial relative to the complex processes the brain routinely performs (e.g., visual processing). As for determining whether neurometabolic resources underlie the subjective cost of effort, continued empirical work, aided by technological advances, is the only path forward.

Metabolic constraints restricting theory development

Unifying theories across Marr's levels of analysis is a laudable goal (Griffiths, 2026, p. 294), but it is not always immediately beneficial. At some point, metabolic realities must be faced, but putting such concerns into the "too hard basket" is often practical in the short term. The history of backpropagation in artificial neural networks illustrates this point.

Backpropagation – the algorithm used to train multilayered neural nets – is the bedrock of deep learning. Real neurons, however, lack the symmetric weights and global error signals that standard backpropagation requires (see Whittington & Bogacz, 2019), leading some to conclude that such models bear no relevance to the brain (Crick, 1989; Grossberg, 1987). Rather than abandoning such approaches to understanding the mind, however, many cognitive scientists continued using backpropagation-based models to develop models of human intelligence, either placing implementation concerns on the backburner or working in parallel on biologically feasible yet functionally comparable models (e.g., Hinton & McClelland, 1988; O'Reilly, 1996). Recently, several biologically plausible alternatives have been proposed that mimic the functional benefits of backpropagation (for reviews, see Lillicrap et al., 2020; Whittington & Bogacz, 2019).

In a sense, this brief history vindicates Haueis & Colaço's proposal in that theory was improved by adapting models to known biological constraints. However, if the field had entirely dismissed "connectionist" models because of their biological implausibility, we may have forgone the foundation upon which better theories were eventually built. Scepticism (much of it warranted) toward neural networks as models of the mind persists today, driven by the enormous differences in energy consumption between the brain and AI models (section 4.3). It is arguable, however, that the impressive capabilities of large-scale AI – which demonstrate that complex, human-like reasoning is achievable in artificial systems – while not intended to serve as models of the mind, have spurred interest in developing biologically constrained alternatives. Metabolic implausibility tells us that a model must be improved upon, but it should not lead to the abandonment of a particular approach at an early stage, especially when existing alternatives do a poorer job of explaining the data (e.g., category learning in symbolic vs. connectionist models).

The weight placed on whether a model aligns with the brain's biology should be considered with respect to the model's maturity. If a model is widely accepted, with differences between formalised mechanisms being *relatively* minor (e.g., differences between diffusion decision models; Ratcliff et al., 2016), then metabolic concerns will be useful in helping adjudicate between accounts. If, however, the theories being debated predominantly work at Marr's computational level (e.g., opportunity cost vs. resource-based accounts of fatigue), or the behaviour theorised is so complex that we have little idea how the brain may implement the hypothesised algorithms (e.g., models of social cognition, Alister & Perfors, 2025; Young & de-Wit, 2025), then biological concerns are best deferred until our theories are sufficiently advanced.

Acknowledgements. Thanks to Ross Otto and Matthieu Chidharom for providing feedback on this commentary.

Financial support. No funding to declare.

Competing interests. No competing interests to declare.

References

Alister, M., & Perfors, A. (2025). Stochastic search algorithms can tell us who to trust (and why). *Proceedings of the Annual Meeting of the Cognitive Science Society*, *47*, 3513–3520.

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

Crick, F. (1989). The recent excitement about neural networks. *Nature*, *337*(6203), 129–132.
Dora, J., van Hooff, M. L., Geurts, S. A., Kompier, M. A., & Bijleveld, E. (2022). The effect of opportunity costs on mental fatigue in labor/leisure trade-offs. *Journal of Experimental Psychology: General*, *151*(3), 695.
Embrey, J. R., Mason, A., Donkin, C., & Newell, B. R. (2025). On-task errors drive effort avoidance more than opportunity costs. *Journal of Experimental Psychology: General. 154(7),* 1916–1938. https://doi.org/10.1037/xge0001752
Gailliot, M. T., & Baumeister, R. F. (2007). The physiology of willpower: Linking blood glucose to self-control. *Personality and Social Psychology Review*, *11*(4), 303–327.
Griffiths, T. L. (2026). *The laws of thought: The quest for a mathematical theory of the mind.* Henry Holt and Co.
Grossberg, S. (1987). Competitive learning: From interactive activation to adaptive resonance. *Cognitive Science*, *11*(1), 23–63.
Hinton, G. E., & McClelland, J. L. (1988). Learning representations by recirculation. In D. Z. Anderson (Ed.), *Neural information processing systems, 1987* (pp. 358–366). American Institute of Physics.
Holroyd, C. B. (2024). The controllosphere: The neural origin of cognitive effort. *Psychological Review*, *132*(3), 603. https://doi.org/10.1037/rev0000467
Inzlicht, M., Schmeichel, B. J., & Macrae, C. N. (2014). Why self-control seems (but may not be) limited. *Trends in cognitive sciences*, *18*(3), 127–133.
Kurzban, R., Duckworth, A., Kable, J. W., & Myers, J. (2013). An opportunity cost model of subjective effort and task performance. *Behavioral and Brain Sciences*, *36(6),* 661–679.
Lillicrap, T. P., Santoro, A., Marris, L., Akerman, C. J., & Hinton, G. (2020). Backpropagation and the brain. *Nature Reviews Neuroscience*, *21(6),* 335–346.
Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. W.H. Freeman.
O'Reilly, R. C. (1996). Biologically plausible error-driven learning using local activation differences: The generalized recirculation algorithm. *Neural Computation*, *8(5),* 895–938.
Otto, A. R., & Daw, N. D. (2019). The opportunity cost of time modulates cognitive effort. *Neuropsychologia*, *123*, 92–105.
Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion decision model: Current issues and history. *Trends in Cognitive Sciences*, *20*(4), 260–281.
Whittington, J. C. R., & Bogacz, R. (2019). Theories of error back-propagation in the brain. *Trends in Cognitive Sciences*, *23*(3), 235–250.
Wiehler, A., Branzoli, F., Adanyeguh, I., Mochel, F., & Pessiglione, M. (2022). A neuro-metabolic account of why daylong cognitive work alters the control of economic decisions. *Current Biology*, *32*(16), 3564–3575.
Young, D. J., & de-Wit, L. H. (2025), The bias-and-expertise model: A Bayesian network model of political source characteristics. *Cognitive Science*, *49*, e70141. https://doi.org/10.1111/cogs.70141

Multiscale dynamical considerations for modeling (biological) cognition

Luis H. Favela*

Indiana University Bloomington, Bloomington, USA
lhfavela@iu.edu

*Corresponding author.

doi:10.1017/S0140525X26106268, e459

Abstract

Haueis and Colaço argue cognition must be modeled with metabolic constraints, challenging more abstract, computational views. This commentary supports their approach but identifies a key omission: multiscale dynamics. Cognition is fundamentally dynamic, emerging across coordinated timescales where metabolism is embedded, not merely constraining. Incorporating multiscale dynamics would strengthen their framework by aligning models with biologically realistic, time-dependent processes.

Haueis and Colaço's target article is timely and important. At a moment when artificial intelligence, neuroscience, and the intersection of the two (i.e., NeuroAI) increasingly position themselves as the primary explanatory domains for cognition, the authors' insistence that metabolic considerations constrain and generate cognitive models represents a necessary corrective. Their central claim – that cognition cannot be adequately modeled without accounting for the brain's energetic limitations and organization – directly challenges the long-standing emphasis on abstraction in cognitivist and other computational-oriented approaches that treat information processing as effectively costless. Though there is much to laud about this approach, there remains a glaring omission: time, specifically, multiscale dynamics. While the authors discuss dynamics a bit, temporality does not appear to be central to modeling or explaining cognition for them. The main aim of this commentary is to motivate the idea that temporality – specifically in the form of multiscale dynamics – ought to be treated as fundamental to understanding cognition, especially in biological systems. Accordingly, Haueis and Colaço's already-compelling project can be made more so by incorporating such temporal features.

A long-standing criticism of computational accounts of cognition is their atemporal nature (Brette, 2026; Favela & Raja, 2026; van Gelder, 1995). Here, "computational" refers to the idea that cognition is a form of information processing along the lines of a Turing machine (i.e., the algorithmic manipulation of symbols; Rescorla, 2025) and that brains instantiate such processing in some substantive way (e.g., Churchland & Sejnowski, 1992/2017; Rolls, 2021). This limitation has prompted attempts to reorient various relevant disciplines in ways that place temporal features at more central explanatory positions, including cognitive science (e.g., Ward, 2002) and neuroscience (e.g., Izhikevich, 2007). This shift has often been accompanied by increased attention on the brain's multiscale spatial organization (e.g., D'Angelo & Jirsa, 2022; Haueis, 2022; Thomas, Forrester, & Ronald, 2016), as well as its multiscale dynamics (e.g., Munn et al., 2024; Siettos & Starke, 2016; Zhang et al., 2020). One consequence of such disciplinary expansion upon the more traditional approaches is that cognition is becoming viewed as fundamentally a temporally extended, multiscale dynamical process. From this perspective, cognition is understood as emerging from the coordination of processes across multiple spatial and temporal scales, from single neurons to organism–environment interactions (Favela, 2024). In such a framework, metabolism is not an isolated constraint but is itself embedded within – and subordinate to – ongoing dynamical organization. Energy consumption, allocation, and efficiency are always in the service of maintaining and adapting these

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

coordinated dynamics in order to cause and constitute various cognitive processes.

This perspective reveals a deeper issue: metabolism without time risks reifying static constraints rather than capturing living processes. Haueis and Colaço treat metabolic factors as bounding and shaping computation, but they do not fully integrate the idea that these constraints are themselves dynamically modulated across timescales. For example, gene expression, synaptic plasticity, neural oscillations, and large-scale functional connectivity all unfold over distinct temporal regimes, yet must be coordinated for cognition to occur. Even metabolic processes – such as ATP production or glucose utilization – are temporally structured and context-sensitive, fluctuating with task demands and environmental coupling. Thus, from a multiscale dynamical viewpoint, metabolism should not be treated as a mere constraint on models; it should be understood as participating in temporally nested systems of interactions. The brain's energy use is not simply a budget to be respected, but a dynamically allocated resource that co-evolves with cognitive activity – not to mention nonneural physiology, whole bodies, and features of the environment. In terms of doing cognitive science with an emphasis on brains, explanations shift to tracking how patterns evolve over time (e.g., Buzsáki, 2006; Criscuolo, Schwartze, & Kotz, 2022) rather than identifying static components (e.g., representations) or computational constraints (e.g., algorithms).

Importantly, incorporating multiscale dynamics would strengthen, rather than undermine, Haueis and Colaço's project. Their evaluative and generative framework could be extended to include temporal constraints, where models would need to not only fit within energy budgets but also to reproduce the timing, coordination, and stability properties of real (i.e., biological) cognitive systems. For instance, a metabolically plausible model that fails to capture the temporal synchronization of neural populations during perception-action cycles would remain explanatorily deficient. Furthermore, emphasizing dynamics helps bridge the gap between metabolism (i.e., molecular) and more situated, organism-environment scales. While Haueis and Colaço focus on neural and computational modeling, other work that aims at more system-level targets of investigation aim to situate cognition within continuous organism–environment interactions, where timing and coordination are essential (e.g., Criscuolo, Schwartze, & Kotz, 2022; Kelso, 2012; additional examples in Favela, 2024). Metabolic processes, in this context, support the organism's ability to maintain these interactions over time. Thus, metabolism is best understood as enabling the ongoing dynamical coupling that constitutes cognition, rather than merely constraining internal computations.

In conclusion, Haueis and Colaço's article is a timely and important contribution that rightly foregrounds metabolic considerations in cognitive modeling, particularly in an era dominated by computational-centric approaches (cf. Favela, 2026; Favela & Amon, 2023). Their work pushes the field toward greater biological realism and offers valuable tools for constraining and generating models. However, from the perspective of more systems-level approaches (e.g., Chemero, 2009; Favela, 2024; Pessoa, 2022), their framework remains incomplete without a more explicit and central role for multiscale temporal dynamics. Biological cognition is not just energy-limited; it is dynamically organized across time, and metabolism itself is embedded within this organization. Recognizing this would not only address a key omission but also further enhance the explanatory power of their approach.

Financial support. None.

Competing interests. None.

Reference

Brette, R. (2026). *The brain, in theory*. Princeton University Press.
Buzsáki, G. (2006). *Rhythms of the brain*. Oxford University Press.
Chemero, A. (2009). *Radical embodied cognitive science*. The MIT Press.
Churchland, P. S., & Sejnowski, T. J. (1992/2017). *The computational brain (25th anniversary ed.)*. The MIT Press.
Criscuolo, A., Schwartze, M., & Kotz, S. A. (2022). Cognition through the lens of a body-brain dynamic system. *Trends in Neurosciences*, *45*(9), 667–677. https://doi.org/10.1016/j.tins.2022.06.004.
D'Angelo, E., & Jirsa, V. (2022). The quest for multiscale brain modeling. *Trends in Neurosciences*, *45*(10), 777–790. https://doi.org/10.1016/j.tins.2022.06.007.
Favela, L. H. (2024). *The ecological brain: Unifying the sciences of brain, body, and environment*. Routledge. https://doi.org/10.4324/9781003009955.
Favela, L. H. (2026). Understanding brains and minds on their own terms. In D. G. Kelty-Stephen, & M. Mangalam (Eds.), *Beyond the computer metaphor of mind and brain* (pp. 102–119). Cambridge University Press. https://doi.org/10.1017/9781009626286.007.
Favela, L. H., & Amon, M. J. (2023). Reframing cognitive science as a complexity science. *Cognitive Science*, *47*(4), e13280. https://doi.org/10.1111/cogs.13280.
Favela, L. H., & Raja, V. (2026). Dynamical cognitive science! Wherefore art thou? *Topics in Cognitive Science*. https://doi.org/10.1111/tops.70042.
Haueis, P. (2022). Descriptive multiscale modeling in data-driven neuroscience. *Synthese*, *200*(129), 1–26. https://doi.org/10.1007/s11229-022-03551-y.
Izhikevich, E. M. (2007). *Dynamical systems in neuroscience: The geometry of excitability and bursting*. The MIT Press.
Kelso, J. A. S. (2012). Multistability and metastability: Understanding dynamic coordination in the brain. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1591), 906–918. https://doi.org/10.1098/rstb.2011.0351
Munn, B. R., Müller, E. J., Favre-Bulle, I., Scott, E., Lizier, J. T., Breakspear, M., & Shine, J. M. (2024). Multiscale organization of neuronal activity unifies scale-dependent theories of brain function. *Cell*, *187*(25), 7303–7313. https://doi.org/10.1016/j.cell.2024.10.004
Pessoa, L. (2022). *The entangled brain: How perception, cognition, and emotion are woven together*. The MIT Press.
Rescorla, M. (2025). The computational theory of mind. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford encyclopedia of philosophy (fall 2025 ed.)*. Stanford University. https://plato.stanford.edu/archives/fall2025/entries/computational-mind/.
Rolls, E. T. (2021). *Brain computations: What and how*. Oxford University Press.
Siettos, C., & Starke, J. (2016). Multiscale modeling of brain dynamics: From single neurons and networks to mathematical tools. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, *8*(5), 438–458. https://doi.org/10.1002/wsbm.1348.
Thomas, M. S., Forrester, N. A., & Ronald, A. (2016). Multiscale modeling of gene-behavior associations in an artificial neural network model of cognitive development. *Cognitive Science*, *40*(1), 51–99. https://doi.org/10.1111/cogs.12230.
Van Gelder, T. (1995). What might cognition be, if not computation? *The Journal of Philosophy*, *92*(7), 345–381. https://doi.org/10.2307/2941061.
Ward, L. M. (2002). *Dynamical cognitive science*. The MIT Press.
Zhang, M., Kalies, W. D., Kelso, J. S., & Tognoli, E. (2020). Topological portraits of multiscale coordination dynamics. *Journal of Neuroscience Methods*, *339*(108672), 1–12. https://doi.org/10.1016/j.jneumeth.2020.108672.

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

Metabolic considerations are generic biological details without resource-rational analysis

Brendan Fleig-Goldstein*

Brown University, USA
brendan_fleig-goldstein@brown.edu

*Corresponding author.

doi:10.1017/S0140525X26106281, e460

Abstract

H&C leave unanswered whether metabolic considerations (MCs) are merely one neglected type of biological detail that constrains cognitive models or a distinctive kind uniquely poised to do so. Their examples give little reason to believe the latter. If the former, a methodology built around MCs specifically – rather than neuroscientific facts generally – is unjustified unless embedded within resource-rational analysis.

H&C convincingly point out that metabolic considerations (MCs) have been unduly ignored in part because multiple realizability and medium independence theses have led cognitive scientists to view them as inconsequential to characterizing cognitive processes, and because the general view of the brain as energy-efficient is sometimes taken to license their neglect. The authors make a valuable contribution by showing that these are bad reasons to ignore MCs and that such facts should be included in the general stockpile of evidence that constrains and indicates models of cognition.

However, one might wonder whether MCs are distinctive in their potential to inform cognitive theory. H&C's complaint that they are often invoked too generically is used to motivate the need for a systematic methodology. But if MCs are just another sort of biological detail, then there is no reason to suppose they require their own methodology distinct from one that utilizes biological details more generally.

Here is why they might be distinctive. All cognitive processes consume energy, so ATP serves as a common currency that every cognitive process must spend, forming a particularly informative bridge between low-level biological facts and higher-level theories of cognition. By contrast, the fact that, say, a synapse typically releases a single kind of neurotransmitter can constrain plausible cognitive models but can hardly constrain high-level theories, since it bears no direct connection to them. A Bayesian model pitched at the highest level of analysis, on the other hand, can still be ruled out if it requires more information transfer than ATP budgets could plausibly support. Because MCs bear on efficiency, the thought goes, they might constrain cognitive theories in unique ways, warranting a specialized methodology.

Interestingly, it is not clear that the examples H&C focus on actually illustrate MCs playing this kind of distinctive role. In the case of the core evaluative example, the linear cascade model, the point is simply that synapses at equilibrium are dead, and dead synapses cannot store memories. The constraint that cells be living applies equally to non-cognitive biological processes.

The supplementary example of the leaky integrate-and-fire models being assessed as less plausible is not due to considerations of resource efficiency but simply that EDLIF is more empirically accurate when it correctly models the Na^+/K^+ pump's activity as dependent on ATP. While modeling ATP's causal effect on other variables is in some sense ipso facto about energy use, there is no analysis of ATP per bit or the like; there is no bridge to cognition of the kind that would make MCs distinctive. The model is already cashed out at the biological-detail level and is being de-idealized by including a further causal variable, as could be done with any biological detail.

As for their core generative example, the efficiency-tradeoff hypothesis does not generate a novel cognitive model. Instead, it takes a pre-existing cognitive model (predictive processing) with already existent specified implementational details (thin interneuron axons transmitting error signals) and uses MCs merely to rationalize this design choice.

Part of the reason why it is difficult to use MCs in a non-generic way may have to do with the nature of efficiency considerations in general. Consider the oft-remarked point about the brain's incredible efficiency. Whereas frontier LLMs can draw up to 1000 watts during inference, brains run on around 20 watts. This seems to close the question of which is more efficient, but the story is complicated once one asks what one is getting for such a wattage. ChatGPT-5 can, according to estimates, use around 34 Wh on reasoning-heavy tasks, which is roughly what the human brain uses in two hours (Jegham et al., 2025). If one supposes that it would take a human more than two hours to produce equivalent output, then ChatGPT would be, in fact, more efficient than the human brain on that task. The point is similar to one that H&C rightly note: ATP/second is not the correct measure of efficiency. ATP/bit, as they suggest, is better. Shannon information, however, is famously semantically blind. All ChatGPT outputs of the same token length, produced with similar inference times, will have a similar electricity per bit, yet one output might take a human significantly less than two hours to replicate, and another significantly more.

The lesson to draw out here is that to assess resource efficiency, one needs to know the goal of cognition. Without knowing what the task is, what is being accomplished, and how well it's being done, it's not possible to assess whether ATP is being used wisely or wastefully. Icard (2023) makes the point that to know the value of a resource, and whether it's well spent, you want to know what else you could have spent those resources on.

This requirement – evaluating efficiency relative to task-specific opportunity costs – is exactly what resource rational analysis (RRA) systematizes. RRA begins by establishing a task-environment and deriving an optimal cognitive model. Discrepancies between predicted and actual behavior indicate psychological constraints, which are incorporated to derive a new model that is resource-optimal relative to bounds. This new model provides a baseline from which further discrepancies can be observed, and the methodology iterates.

Such a methodology provides exactly what H&C want, despite their claim that RRA lacks a unifying account of metabolic and informational constraints. In fact, RRA translates ATP and MCs, in general, into a common currency by treating them as limited

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

resources that constitute task-indexed constraints, which can then be traded off against other kinds of constraints. An explicit step is deriving optimal models from the set of identified constraints, which provides precisely the generative use that H&C claim is missing from RRA. This is the advantage of the top-down approach that RRA offers by initially deriving optimality directly from task-environments and working its way down, whereas a bottom-up approach that starts with MCs can only employ them as generic biological details. If one wants a systematic methodology that treats MCs as a distinctive kind of biological detail, one should do RRA.

Financial support. Preparation of this commentary was supported by a Mellon Postdoctoral Fellowship at the Cogut Institute for the Humanities, Brown University.

Competing interests. The author declares no competing interests.

References

Icard, T. F. (2023). Resource rationality. https://philarchive.org/rec/ICARRT

Jegham, N., Abdelatti, M., Koh, C. Y., Elmoubarki, L., & Hendawi, A. (2025). How hungry is AI? Benchmarking energy, water, and carbon footprint of LLM inference. https://arxiv.org/abs/2505.09598

Beyond energy: Expanding metabolic considerations in cognitive modeling

Angeles Garcia-Cazorla[a,b] and Fanny Mochel[c,d]*

[a]Neurometabolic Unit, Neurology Department, Institut de Recerca Sant Joan de Déu, CIBERER and MetabERN, Hospital Sant Joan de Déu, Barcelona, Spain; [b]Synaptic Metabolism and Personalized Therapies Lab, Institut de Recerca Sant Joan de Déu, Barcelona, Spain; [c]Team Metabolism, Immunity and NeuroDegeneration, INSERM U 1127, CNRS UMR 7225, Sorbonne Universités, UPMC Univ Paris 06 UMR S 1127, Paris Brain Institute (ICM), Paris, France and [d]Reference Center for Adult Neurometabolic Diseases, Department of Medical Genetics, AP-HP, Pitié-Salpêtrière University Hospital, Paris, France
angeles.garcia@sjd.es
fanny.mochel@icm-institute.org

doi:10.1017/S0140525X26106311, e461

Abstract

Integrating metabolism into cognitive science is essential. Beyond a mere constraint, metabolism may be constitutive of cognition itself. This requires expanding metabolism beyond neuronal energy use to include glial contributions, compartmentalized and dynamic processes, and broader brain–body interactions (e.g., interoception). Capturing this complexity will require more refined, integrative models, potentially leveraging artificial intelligence to unify multiscale biological data.

Haueis and Colaço provide a timely, original, and highly stimulating contribution by arguing that metabolic considerations should be incorporated into models of cognition. Evaluative metabolic considerations test whether a cognitive model is energetically plausible, while generative metabolic considerations use energy constraints to guide the construction of new models. For cognitive science, metabolic considerations provide an objective standard for evaluating models of mental effort. Rather than treating cognition as purely computational, this perspective emphasizes that cognitive processes are fundamentally energy-constrained and must be explained in relation to their biological costs. In an era increasingly shaped by comparisons between biological brains and artificial intelligence systems, their emphasis on energetic efficiency also highlights an often neglected but fundamental feature of cognition. Indeed, metabolism introduces energy efficiency as a central metric alongside performance. It complicates claims of parity between artificial and biological systems by highlighting their radically different energy regimes, and it encourages the development of more sustainable, biologically informed, humanized approaches to artificial intelligence design.

We strongly support the central thesis of the article. Nevertheless, several conceptual and metabolic issues may enrich the discussion further.

The argument depends critically on how *cognition* is defined. If cognition is conceived narrowly as symbolic or computational information processing, metabolism may appear as a supporting constraint. However, if cognition is understood more broadly as embodied, adaptive regulation of organism–environment interactions, metabolism may be constitutive of cognition itself rather than merely enabling it. Some processes that look purely metabolic – like blood flow – may actually contribute to information processing. Contemporary philosophy of biology and embodied cognition increasingly supports this broader view, emphasizing that informational and metabolic processes are deeply intertwined in living systems (Theparambil et al., 2024).

If metabolism is best understood as a resource rather than a mechanistic component, then metabolic considerations for cognitive modeling should not be restricted to energy metabolism. Likewise, we proposed that cognitive fatigue is not due to energy depletion but to the accumulation of neurochemical byproducts, notably glutamate, during sustained cognitive control (Wiehler et al., 2022). Using magnetic resonance spectroscopy, we showed that demanding tasks increase glutamate levels in the lateral prefrontal cortex, alongside behavioral signs of fatigue. Instead, metabolism is often reduced to neuronal energy production – primarily oxidative phosphorylation and glycolysis – and to the energetic costs of information.

While historically understandable, current neuroscience increasingly recognizes that glial cells – particularly astrocytes, but also oligodendrocytes and microglia – are active participants in cognition. Astrocytes regulate neurotransmitter recycling, lactate shuttling, ion buffering, blood-flow coupling, synaptic plasticity, and network synchronization. Several studies show that astrocytes actively contribute to fear memory retrieval, extinction, and memory engram modulation, challenging purely neurocentric models of cognition (Bukalo et al., 2026), as well as key regulators of signaling and behavior (Won, Patel, & Fields, 2024).

Cerebral metabolism is not a static variable but an extraordinarily dynamic and context-dependent process. It changes substantially across the lifespan: the infant brain displays metabolic patterns radically different from the adult or aging brain, with developmental windows of intense glucose use, synaptogenesis, and plasticity (Rava et al., 2026). Metabolism also differs across brain regions according to functional specialization, connectivity, and cellular composition. Furthermore, it varies by cell type: fast-spiking inhibitory interneurons have distinctive energetic

requirements compared with pyramidal neurons (Szegedi et al., 2026). Even within a single cell, metabolism is compartmentalized, with relevant distinctions between soma, dendrites, axons, synaptic terminals, and mitochondria-rich subdomains. Single-cell transcriptomics further demonstrate the remarkable spatial and temporal heterogeneity of astrocytic metabolic states across brain regions (Serrano-Pozo, 2024). These layers of heterogeneity suggest that invoking "brain metabolism" as a unitary explanatory variable may be insufficient.

The most fruitful integration of metabolism into cognitive science may come not only from energy budgets but also from a broader biochemical network perspective (Marano, Mariano, & Novarino, 2025), including brain–body interactions. Indeed, interoception suggests metabolism is not merely an external constraint but constitutive of cognition (Feldman, Bliss-Moreau, & Lindquist, 2024), because metabolic states are internally represented and used to drive processing. This means metabolism does not just limit computation – it becomes part of the causal and informational structure of cognitive mechanisms themselves.

Finally, metabolic considerations of cognition should be approached cautiously. Neurons show remarkable resilience and can preserve viability and partial function under low-energy or fluctuating metabolic states through homeostatic plasticity, adaptive prioritization, and network redundancy (Sobieski, Fitzpatrick, & Mennerick, 2018). This resilience suggests that metabolism may be one major dimension of cognition, but not necessarily a sufficient explanatory principle. Cognitive phenomena likely emerge from interactions among metabolism, cellular physical properties, circuit dynamics, developmental programs, and environmental inputs.

These considerations raise an important methodological question: how should metabolism be operationalized in future cognitive models? Should one model glucose uptake in prefrontal cortex during executive tasks, astrocytic lactate support during memory formation, mitochondrial constraints in hippocampal plasticity, or age-dependent shifts in cortical energetic allocation? The answer is likely pluralistic. Different cognitive domains may depend on different metabolic signatures, and no single scalar measure may suffice. Addressing these questions will be instrumental to apply the guidelines proposed by Haueis and Colaço on both evaluative and generative metabolic considerations.

For this reason, the future envisioned by Haueis and Colaço may require tools capable of integrating multiscale complexity. Artificial intelligence, machine learning, and systems biology approaches may be particularly useful for combining metabolomics, magnetic resonance methodologies, electrophysiology, developmental data, and behavioral phenotypes into unified predictive frameworks. Such approaches could identify hidden metabolic states linked to attention, learning, fatigue, resilience, or disease vulnerability. This would move the field from broad energetic function toward mechanistically testable, multilevel models.

In summary, this important article opens an interdisciplinary agenda that deserves broad attention. It convincingly argues that cognition cannot be fully understood while neglecting metabolism. Metabolic constraints enable researchers to rule out energetically infeasible models while also guiding the construction of biologically plausible ones. The next step may be to embrace a richer view of metabolism itself: distributed across neurons and glia, dynamic across time, heterogeneous across regions and compartments, and extending far beyond ATP production alone. Integrating this broader metabolic biology into cognitive theory models may substantially reshape our understanding of how minds emerge from living brains.

Financial support. AGC is supported by PI24/00469 Instituto de Salud Carlos III (ISCIII) and Fondo Europeo de Desarrollo Regional (FEDER). FM is supported by the Agence Nationale de la Recherche (ANR) and the Ministry of Health-Programme Hospitalier de Recherche Clinique (PHRC-APHP220801).

Competing interests. The authors declare that they have no competing interests.

References

Bukalo, O., O'Sullivan, R., Tanisumi, Y., Mendez, A., Weinholtz, C., Zimmerman, S., Offenberg, V., Carpenter, O., Bhagwat, H., Mosley, S., O'Malley, J. J., Lyons, K, Fang, Y., Goldschlager, J., Ostroff, L. E., Penzo, M. A., Wake, H., Halladay, L. R., & Holmes, A. (2026). Astrocytes enable amygdala neural representations supporting memory. *Nature*, *652*(8109), 434–441. https://doi.org/10.1038/s41586-025-10068-0

Feldman, M. J., Bliss-Moreau, E., & Lindquist, K. A. (2024). The neurobiology of interoception and affect. *Trends in Cognitive Sciences*, *28*(7), 643–661.

Marano, D., Mariano, V., & Novarino, G. (2025). Fueling the mind: Brain metabolism in health and neurodevelopmental disorders. *Annual Review of Genetics*, *59*(1), 415–434.

Rava, V., Cannone, E., Francolini, M., & Taverna, E. (2026). Metabolism controls the timing of human brain development and maturation. *Journal of Inherited Metabolic Disease*, *49*(3), e70180.

Serrano-Pozo, A., Li, H., Li, Z., Muñoz-Castro, C., Jaisa-Aad, M., Healey, M. A., Welikovitch, L. A., Jayakumar, R., Bryant, A. G., Noori, A., Connors, T. R., Hu, M., Zhao, K., Liao, F., Lin, G., Pastika, T., Tamm, J., Abdourahman, A., Kwon, T., Bennett, R. E., Woodbury, M. E., Wachter, A., Talanian, R. V., Biber, K., Karran, E. H., Hyman, B. T., & Das, S. (2024). Astrocyte transcriptomic changes along the spatiotemporal progression of Alzheimer's disease. *Nature Neuroscience*, *27*(12), 2384–2400.

Sobieski, C., Fitzpatrick, M. J., & Mennerick, S. (2018). Ambient but not local lactate underlies neuronal tolerance to prolonged glucose deprivation. *PLoS One*, *13*(4), e019552.

Szegedi, V., Douida, A., Hutóczki, G., Novák, L., Lamsa, K. (2026). Evolutionary neuroeconomic adaptations of fast-spiking neurons in the human neocortex. *Frontiers in Synaptic Neuroscience*, *18*, 1741452.

Theparambil, S. M., Kopach, O., Braga, A., Nizari, S., Hosford, P. S., Sagi-Kiss, V., Hadjihambi, A., Konstantinou, C., Esteras, N., Gutierrez Del Arroyo, A., Ackland, G. L., Teschemacher, A. G., Dale, N., Eckle, T., Andrikopoulos, P., Rusakov, D. A., Kasparov, S., & Gourine, A. V. (2024). Adenosine signalling to astrocytes coordinates brain metabolism and function. *Nature*, *632*(8023), 139–146. https://doi.org/10.1038/s41586-024-07611-w.

Wiehler, A., Branzoli, F., Adanyeguh, I., Mochel, F., & Pessiglione, M. (2022). A neuro-metabolic account of why daylong cognitive work alters the control of economic decisions. *Current Biology*, *32*(16), 3564–3575.e5.

Won, W., Patel, R. A., & Fields, R. D. (2024). Astrocytes as key regulators of neural signaling in health and disease. *Annual Review of Neuroscience*, *47*, 251–276.

What the func? Multiple realizability need not be vague

Olivia Guest[a,b*], Mark Blokpoel[a,b] and Iris van Rooij[a,b,c]

[a]Department of Cognitive Science and Artificial Intelligence, Radboud University, Nijmegen, Netherlands; [b]Donders Institute for Brain, Cognition, and Behaviour,

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

Radboud University, Nijmegen, Netherlands and [c]Department of Linguistics, Cognitive Science, and Semiotics, Aarhus University, Denmark
olivia.guest@donders.ru.nl
mark.blokpoel@donders.ru.nl
iris.vanrooij@donders.ru.nl

*Corresponding author.

doi:10.1017/S0140525X26106128, e462

Abstract

Multiple realizability (MR) is not necessarily unclear nor does it purely operate at the computational level. To understand potential relationships between MR and other constraints, such as metabolic, we formalize possible meanings of function in cognitive science. We build on these to formalize MR, thus resolving its apparent vagaries. Importantly, MR formalisms meaningfully guide and constrain theory building.

Haueis and Colaço rightly point out that resource considerations constrain the space of possible and plausible models (viz. Blokpoel, 2018; van Rooij, 2008; van Rooij & Baggio, 2021; van Rooij et al., 2019). Yet, multiple realizability (MR) remains an intrinsic feature of explanatory models. We offer a complementary but different view on MR to demonstrate that it does not deal in vagaries, nor exist solely at the computational level, and is a productive part of modeling.

First, notions of MR have existed for millennia. Voula Tsouna (2023) explains: "Epicurus' conception of the *pollachos tropos* [method of multiple explanations] is rich, nuanced, fairly coherent, and strongly motivated by methodological and scientific considerations" (p. 240). Epicureans contemplated celestial phenomena, but the argument applies to cognitive neuroscience: "there are in-principle reasons why explanatory pluralism is to be expected[:] the computational description is no more and no less realistic and fundamental than the mechanistic/implementational one" (Chirimuuta, 2018, p. 406).

Second, we define four distinct notions of function: φ, T, G, and τ. A function $F : I \rightarrow O$ maps elements in the set I to elements in the set O. Importantly, functions are not guaranteed to be computable. And for any given computable function, there exist many algorithms. This is an instance of MR at the algorithmic level to which we return below. Before that, what could a function be?

Functional capacity $\boldsymbol{\varphi}$: A system $\mathbb{S}$ (modeled by $\mathbb{M}$, see Figure 1) may be held to have some purported functional capacity, denoted by $\varphi : I_\varphi \rightarrow O_\varphi$. For instance, a person, brain, or component may be believed to have the capacity for addition, φ.

Functional theory T: A theorist may specify a functional theory for a purported capacity φ using a well-defined function, $T : I_T \rightarrow O_T$ (Egan, 2017). For instance, $T : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$ with $T(a, b) = a + b$ where $a, b \in \mathbb{N}$ may model the functional capacity for addition φ.

Functional goal G: An algorithmic model A may be designed or tuned (e.g., so-called "training" of artificial neural network models; Guest & Martin, 2023, 2026) to perform a given goal capacity, denoted $G : I_G \rightarrow O_G$. We say A (approximately) computes functional goal G if and only if $A(i) \approx G(i)$ (cf. Blokpoel, 2018; van Rooij & Wareham, 2012). When $G = T$, then A could be an algorithmic-level explanation for functional theory T.

Functional role τ: A functional capacity φ (modeled by T) may be hypothesized to play a functional role $\tau : I_\tau \rightarrow O_\tau$. For instance, adding may aggregate synaptic inputs or determine the cost of groceries. A capacity φ (modeled by T) fulfils a functional role if and only if it can behave consistently with the teleological specification, i.e., $\varphi(i) \equiv T(i) = \tau(i)$, for all $i \in I_\varphi = I_T \supseteq I_\tau$.

All models φ, T, G, and τ, are defined by the scientist. Knowledge obtained from experiments may be used to *abduce* candidate theories T of φ, but not to *deduce* them. In contrast, τ is definitionally *prima facie* assigned to $\mathbb{S}$'s φ and only unassigned if the system is deemed to have broken down (Egan, 2017, 2025). An example from biology for τ is that the function of the heart is to pump blood; perhaps less useful for cognitive science under computationalism as it results in truism, such as the purpose of vision is to see.

Next, we offer MR formalisms.

MR_φ: A functional role τ is said to be multiply φ-realizable if and only if there can in principle exist multiple distinct functional capacities $\varphi_1, \varphi_2, \ldots$ (modeled by $T_1, T_2, \ldots$) that fulfil that role τ in some possible world. We furthermore say τ is multiply φ-realized if there exists more than one such capacity in our world. For example, if one wants to soft-boil an egg for 6 minutes, and knows there are 6 minutes left on a video one is currently watching, then one can use the video player as a timer.

MR_m: A functional capacity φ is said to be multiply m-realizable if and only if there can, in principle, exist multiple distinct mechanisms $m_1, m_2, \ldots$ that have that capacity φ in some possible world. We furthermore say φ is multiply m-realized if there exists more than one such mechanism in our world; recall Figure 1 and bear in mind that substrates do not (straightforwardly) cut it as mechanisms under computationalism (Guest & Martin, 2025). For example, the capacity for addition can be performed in the head, using a calculator, or on an abacus (Guest, 2026). Here, m may refer to the algorithm, the physical implementation, or both.

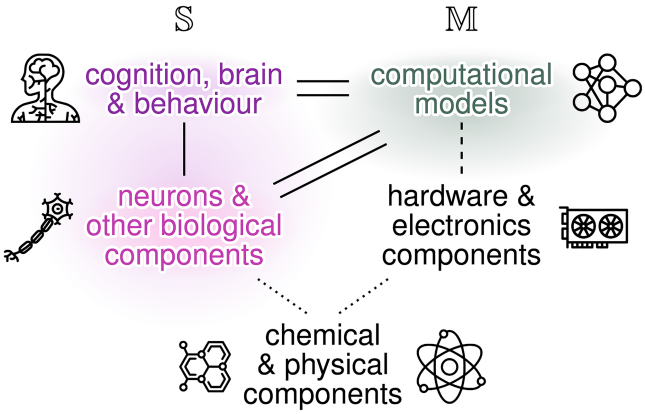


Figure 1 (Guest et al.). Relationships between the system under study $\mathbb{S}$ and models $\mathbb{M}$. The single solid line is relationships within the neurocognitive system. Double solid lines represent the theoretical mapping proposed between $\mathbb{S}$ to $\mathbb{M}$. The dashed line links $\mathbb{M}$ to its digital substrate. Dotted lines represent relationships to lower substrates.

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

By abuse of terminology, we may also say that a functional role τ can be multiply m-realizable if there exist mechanisms $m_1, m_2, \ldots$ that can realize one or more φ that can realize functional role τ. For example, a timer can be instantiated in a wall clock, a software stopwatch, or as above by the video player.

Finally, these meanings of MR demonstrate that it need not be vague nor operating only at the computational level (also see: Chirimuuta, 2018; Egan, 2017; Figdor, 2010; Hardcastle, 1995; Litch, 1997; Polger & Shapiro, 2016; Ross, 2020). MR operates at each level. That is, a particular functional role can be performed by different functional capacities. Functional capacities can be realized by different algorithms. A particular algorithm can be physically realized in different ways.

Any serious proposals for metabolic additions under computationalism will need to contend with MR – at the level and type the modeler defines – not because MR and metabolism have something specific to say about each other but because *both* are relevant. We thus urge and support theorists to take seriously their commitments to MR by providing these formal tools, which help think through what our formal models imply.

Acknowledgements. None.

Financial support. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Open access funding provided by Radboud University Nijmegen.

Competing interests. None.

References

Blokpoel, M. (2018). Sculpting computational-level models. *Topics in Cognitive Science*, *10*(3), 641–648. https://doi.org/10.1111/tops.12282

Chirimuuta, M. (2018). Marr, Mayr, and MR: What functionalism should now be about. *Philosophical Psychology*, *31*(3), 403–418. https://doi.org/10.1080/09515089.2017.1381679

Egan, F. (2017). Function-theoretic explanation and the search for neural mechanisms. https://doi.org/10.1093/oso/9780199685509.003.0007

Egan, F. (2025). *Deflating mental representation*. MIT Press. https://doi.org/10.7551/mitpress/15666.001.0001

Figdor, C. (2010). Neuroscience and the multiple realization of cognitive functions. *Philosophy of Science*, *77*(3), 419–456. https://doi.org/10.1086/652964

Guest, O. (2026). What does 'human-centred AI' mean? *Behavioral Sciences*, *16*(4), 583. https://doi.org/10.3390/bs16040583

Guest, O., & Martin, A. E. (2023). On logical inference over brains, behaviour, and artificial neural networks. *Computational Brain & Behavior*, *6*(2), 213–227. https://doi.org/10.1007/s42113-022-00166-x

Guest, O., & Martin, A. E. (2025). Are neurocognitive representations 'small cakes'? https://philsci-archive.pitt.edu/24834/

Guest, O., & Martin, A. E. (2026). A metatheory of classical and modern connectionism. *Psychological Review*, *133*(3), 719–736. https://doi.org/10.1037/rev0000591

Hardcastle, V. G. (1995). Computationalism. *Synthese*, *105*, 303–317. https://doi.org/10.1007/BF01063561

Litch, M. (1997). Computation, connectionism and modelling the mind. *Philosophical Psychology*, *10*(3), 357–364. https://doi.org/10.1080/09515089708573225

Polger, T. W., & Shapiro, L. A. (2016). *The multiple realization book*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198732891.001.0001

Ross, L. N. (2020). Multiple realizability from a causal perspective. *Philosophy of Science*, *87*(4), 640–662. https://philsci-archive.pitt.edu/16476/

Tsouna, V. (2023). The method of multiple explanations revisited. In F. Masi, P. M. Morel, & F. Verde (Eds.), *Epicureanism and scientific debates. antiquity and late reception: Volume i. language, medicine, meteorology* (pp. 221–256). Leuven University Press. Retrieved March 21, 2026, from http://www.jstor.org/stable/j.ctv34h08cj.13

van Rooij, I. (2008). The tractable cognition thesis. *Cognitive science*, *32*(6), 939–984. https://doi.org/10.1080/03640210801897856

van Rooij, I., & Baggio, G. (2021). Theory before the test: How to build high-verisimilitude explanatory theories in psychological science. *Perspectives on Psychological Science*, *16*(4), 682–697. https://doi.org/10.1177/1745691620970604

van Rooij, I., Blokpoel, M., Kwisthout, J., & Wareham, T. (2019). *Cognition and intractability: A guide to classical and parameterized complexity analysis*. Cambridge University Press. https://doi.org/10.1017/9781107358331

van Rooij, I., & Wareham, T. (2012). Intractability and approximation of optimization theories of cognition. *Journal of Mathematical Psychology*, *56*(4), 232–247. https://doi.org/10.1016/j.jmp.2012.05.002

Not without cost: The philosophical and psychological drawback of hybrid effort models

Malte Hendrickx*

Philosophy, University of Michigan, USA
hmalte@umich.edu

*Corresponding author.

doi:10.1017/S0140525X26106098, e463

Abstract

Haueis and Colaço argue that motivational/computational models of effort should incorporate metabolic factors. But metabolic and motivational accounts predict the same behavioral patterns, making their contributions hard to distinguish empirically. Moreover, philosophy provides independent reasons favoring motivational accounts: effort is unified across domains and normatively significant in ways metabolic factors cannot capture.

Imagine your kid refuses to do his homework. "I can't, Mom, I'm too exhausted." You just read Haueis and Colaço's thought-provoking piece on metabolic considerations; you know your kid had a long day at school, and so you are sympathetic to their plea: maybe this one time, skipping homework is okay because it's genuinely impossible for your child to muster the resources for homework. Or so you think, until your partner walks into the scene and aptly motivates your son by promising them a chocolate bar. Your child, feeling a surge of motivation, finishes his homework.

This little anecdote illustrates a key problem in the literature on mental effort. It is incredibly hard to distinguish between metabolic depletion and motivational degradation: absent strong motivational variation, both predict decreases in task performance. Subjects in research settings typically have milquetoast motivation: most researchers can neither offer them large rewards nor threaten them with genuinely painful consequences if they do a half-hearted job.

This difficulty in differentiating metabolic and motivational factors might be one reason for why ego depletion was so successful (Baumeister et al., 1998): task performance decreasing as motivation wanes might have been wrongly attributed to local resource depletion, rather than to motivational factors (cf. Vohs et al., 2021). If that's true, there is a risk in invoking metabolic factors in the

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

explanation of effort: namely, that it serves as a tempting way to account for behavioral variations that are better explained in terms of the computational model itself.

Haueis and Colaço rightly reject local resource models of cognitive effort. Yet they argue that their main alternative, computational motivational models such as Kurzban's opportunity cost model (Kurzban et al., 2013) or Shenhav's Expected Value of Control model (Shenhav, Botvinick, & Cohen, 2013; Shenhav et al., 2017), should take on board metabolic factors without much explanatory cost. For example, one could leave the EVC model intact but invoke metabolic factors as an external enabling condition that affects outcomes only insofar as it modulates parameters within the model.

Haueis and Colaço point to two recent theoretical proposals as successful examples of integrating metabolic and computational elements (Holroyd, 2025; Zenon, Solopchuk, & Pezzulo, 2019). These are the strongest cases for hybridization in the effort literature. My concern is narrower. In both models, the metabolic component explains the same behavioral change that motivational/computational models already predict: gradual disengagement and performance decrements at costly tasks. This makes it difficult to determine, from behavioral data alone, whether the metabolic addition is doing independent explanatory work or tracking motivational dynamics under a different description. Metabolic data are not completely absent but scarce (though see Wiehler et al., 2022). My point is simply that until much stronger convergent data are in hand, the explanatory benefit of hybridization in this domain remains uncertain.

Hybridization does not come without cost. In both experimental and real-world settings, motivation is found to provide a better explanation of self-control behavior than metabolic depletion. Muraven and Slessareva (2003) showed that "depleted" participants who were offered cash rewards or told their efforts would help others performed just as well as nondepleted participants. Of course, in extreme cases, metabolic factors will matter: no money in the world will make you ace the LSAT when you've been awake for 72 hours straight. But in normal circumstances, the burden of proof for showing more than niche relevance for metabolic considerations is not on the advocates of motivational/computational models.

So far, I have mainly argued on methodological grounds that the literature on mental effort should not consider the hybridization of motivational/computational models and metabolic models cost-free. But there is also independent conceptual evidence that favors motivational/computational models over metabolic models.

Specifically, there is a nascent literature on the philosophy of effort (Bermúdez, 2025; Bradford, 2015; Massin, 2017; Nelkin, 2016; von Kriegstein, 2019). Two points from this literature matter in explaining model choice. The first is that effort strikes many as a unified phenomenon; the second is that effort has a normatively rich profile.

Philosophers observe that the functional role effort and difficulty play is quite independent of whether the effort is mental or physical. For example, we think both Magnus Carlsen and Usain Bolt have performed difficult achievements (Bradford, 2015), and we attribute agency to both mental and bodily efforts (Massin, 2017). The evolutionary function of the feelings of effort as a way of prioritizing tasks, too, is motivational and cuts across domains (Bermúdez, 2025; Kurzban et al., 2013). On a motivation-first picture, this datum is explained because motivation, like effort, cuts across domains (Hendrickx, 2026), whereas metabolic accounts usually must differentiate between mental and bodily effort as they draw on different resources and markers (say, glucose and lactate).

Another worry concerns the normative profile of effort. We care about the efforts people make. If my life is in danger and you make a lukewarm effort, I'll blame you even if you had a long day at work! Grounding this normative profile requires tying effort to something more than metabolites: glucose and ATP are poor candidates for blaming a person, whereas an inappropriately low state of motivation is a direct indicator of the agent's cares and values (Hendrickx, forthcoming; cf. Nelkin, 2016). At least prima facie, this normative profile is data in the sense that we usually take our social practices to be sensitive to descriptive phenomena.

Of course, these philosophical considerations might be trumped by future empirical evidence indicating a hitherto unknown strong role of metabolic considerations in explaining the psychology of effort. And Haueis and Colaço make a strong case that, among a variety of cognitive explanations, theorists should be more mindful to explicitly include metabolic factors in their models. But motivational/computational models have seen spectacular progress in predicting behavior and shedding genuine light on deep philosophical questions about effort and difficulty (Bermudez, 2025; Hendrickx, 2026, forthcoming). Introducing metabolic considerations as a free parameter in otherwise functional models does not come at no cost and should be done only with utmost caution.

Acknowledgements. I am grateful for comments by Lorenzo Manuali.

Financial support. I am grateful for funding received from the Institute for Humane Studies under Grant No. IHS019832 and IHS019115; as well as a Mercatus Fellowship from the Mercatus Center and a Weinberg Fellowship from the Weinberg Center for Cognitive Science.

Competing interests. The author declares no competing interests.

References

Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology*, *74*(5), 1252–1265. https://doi.org/10.1037/0022-3514.74.5.1252

Bermúdez, J. P. (2025). What is the feeling of effort about? *Australasian Journal of Philosophy*, *103*(1), 88–105. https://doi.org/10.1080/00048402.2024.2351208

Bradford, G. (2015). *Achievement*. Oxford University Press.

Hendrickx, M. (forthcoming). *The cognitive control account of effort*. Journal of Philosophy.

Hendrickx, M. (2026). *Difficulty*. Mind.

Holroyd, C. B. (2025). The controllosphere: The neural origin of cognitive effort.

Kurzban, R., Duckworth, A., Kable, J. W., & Myers, J. (2013). An opportunity cost model of subjective effort and task performance. *Behavioral and Brain Sciences*, *36*(6), 661–679. https://doi.org/10.1017/S0140525X12003196

Massin, O. (2017). Towards a definition of efforts. *Motivation Science*, *3*(3), 230–259. https://doi.org/10.1037/mot0000066

Muraven, M., & Slessareva, E. (2003). Mechanisms of self-control failure: Motivation and limited resources. *Personality and Social Psychology Bulletin*, *29*(7), 894–906. https://doi.org/10.1177/0146167203029007008

Nelkin, D. K. (2016). Difficulty and degrees of moral praiseworthiness and blameworthiness. *Noûs*, *50*(2), 356–378. https://doi.org/10.1111/nous.12079

Shenhav, A., Botvinick, M. M., & Cohen, J. D. (2013). The expected value of control: An integrative theory of anterior cingulate cortex function. *Neuron*, *79*(2), 217–240. https://doi.org/10.1016/j.neuron.2013.07.007

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

Shenhav, A., Musslick, S., Lieder, F., Kool, W., Griffiths, T. L., Cohen, J. D., & Botvinick, M. M. (2017). Toward a rational and mechanistic account of mental effort. *Annual Review of Neuroscience*, *40*, 99–124. https://doi.org/10.1146/annurev-neuro-072116-031526

Vohs, K. D., Schmeichel, B. J., Lohmann, S., Gronau, Q. F., Finley, A. J., Ainsworth, S.E., Alquist, J. L., Baker, M. D., Brizi, A., Bunyi, A., & Butschek, G. J. (2021). A multisite preregistered paradigmatic test of the ego-depletion effect. *Psychological Science*, *32*(10), 1566–1581. https://doi.org/10.1177/0956797621989733

Von Kriegstein, H. (2019). On being difficult: Towards an account of the nature of difficulty. *Philosophical Studies*, *176*(1), 45–64. https://doi.org/10.1007/s11098-017-1005-3

Wiehler, A., Branzoli, F., Adanyeguh, I., Mochel, F., & Pessiglione, M. (2022). A neuro-metabolic account of why daylong cognitive work alters the control of economic decisions. *Current Biology*, *32*(16), 3564–3575.e5. https://doi.org/10.1016/j.cub.2022.07.010

Zenon, A., Solopchuk, O., & Pezzulo, G. (2019). An information-theoretic perspective on the costs of cognition. *Neuropsychologia*, *123*, 5–18. https://doi.org/10.1016/j.neuropsychologia.2018.09.013

The metabolic regulator: Why energy constraints imply generative models

Hamidreza Jamalabadi[a]* and Tim Hahn[b]

[a]Department of Psychiatry and Psychotherapy, Marburg University, Germany and [b]Institute for Machine Learning in Medicine (focus area Psychiatry), University of Münster, Germany
hamidreza.jamalabadi@uni-marburg.de
hahnt@uni-muenster.de

*Corresponding author.

doi:10.1017/S0140525X26106190, e464

Abstract

We argue that under minimal thermodynamic and control-theoretic assumptions, metabolic constraints imply generative cognition. We formalize cognition as a resource-bounded controller of probability flows, where metabolic budgets define path-space divergence constraints. This reframes inference as a physically instantiated process governed by metabolic regulation, implying that progress in the cognitive sciences requires grounding cognitive models in nonequilibrium thermodynamics and dynamical systems.

Haueis and Colaço argue that metabolic constraints are fundamental and underutilized in cognitive modeling. We agree but argue that the theoretical implications of their view go far beyond this insight. We propose that under three minimal assumptions – (i) that the brain admits an approximately informationally closed, Markovian description at an appropriate spatiotemporal scale; (ii) that its dynamics can be expressed in terms of variational free energy; and (iii) that cognitive behavior reflects the algorithmic expression of the brain's interaction with its environment – accounting for metabolic constraints entails a stronger conclusion: cognition can be formalized as a generative, control-theoretic process bounded by energy-dependent limits on controllability.

This follows directly from converging results in cybernetics and statistical physics that establish a link between energy dissipation, information processing, and regulation. While metabolism in the target article often refers to the biological implementation of energy transformations, we argue that its constraints arise from general thermodynamic principles governing nonequilibrium systems.

We begin from a minimal physical characterization of the brain as a nonequilibrium system maintaining organization against entropic decay (Prigogine, 1978). Survival, therefore, depends on regulating internal states within bounded ranges despite external perturbations. By the Good Regulator Theorem (Conant & Ross Ashby, 1970), any system capable of effective regulation must embody a model of the variables it regulates. If the brain must regulate its metabolic exchanges to remain within viability constraints, it must instantiate an internal model of those dynamics. This extends beyond the evaluative role proposed by Haueis and Colaço: energy-constrained cognition is necessarily model-based, even when implemented through physical dynamics rather than explicit computation (Haimerl, Rodrigues, & Paton, 2025).

A formal expression of this requirement is provided by the free energy principle (Friston, 2009), which characterizes self-organizing systems that resist entropy. Within this framework, the system minimizes variational free energy:

$$\begin{aligned} F &= \mathbb{E}_{q(\phi)}[\ln q(\phi) - \ln p(o,\phi)] \\ &= \underbrace{D_{KL}[q(\phi) \parallel p(\phi)]}_{complexity} - \underbrace{\mathbb{E}_{q(\phi)}[\ln p(o \mid \phi)]}_{accuracy} \end{aligned}$$

Here, ϕ denotes latent (hidden) states and o observations. Crucially, the complexity term – representing the divergence between the variational posterior and the prior – is not merely informational. When inference is physically implemented, updating internal states is expected to incur thermodynamic cost, and recent work has begun to make this link explicit (Fields, Goldstein, & Sandved-Smith, 2024). Landauer's principle establishes a lower bound on dissipation for logically irreversible operations (Landauer, 1961). More generally, results from stochastic thermodynamics show that maintaining nonequilibrium distributions requires continuous dissipation, with lower bounds related to their divergence from equilibrium (Horowitz & England, 2017; Sanders, Baldovin, & Muratore-Ginanneschi, 2024). In this sense, inference is not purely abstract computation, but a physically instantiated process constrained by thermodynamic cost.

A complementary control-theoretic formulation makes this constraint clearer. Let w_{ij}^0 denote passive transition rates and w_{ij}^π the transition rates under a control policy π. The energetic cost of imposing controlled dynamics over trajectories τ can be written as $\mathrm{Cost}(\pi) = D_{KL}(p(\tau \mid \pi) \parallel \mathrm{p}_0(\tau)) = \sum_\tau p(\tau \mid \pi) \ln \frac{p(\tau|\pi)}{p_0(\tau)}$. This expression shows that control cost arises from biasing controlled trajectories away from passive dynamics, thereby reshaping probability flow over state space (Kawakita et al., 2022). Available energy, therefore, defines a control budget that constrains reachable distributions, admissible transition structure, and the directional organization of dynamics. In neuroscience, this broader intuition is

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

consistent with work linking controllability-related quantities to brain structural properties (Haimerl et al., 2025; Jamalabadi et al., 2021; Takagi, 2021). In cognitive and psychiatric settings, it is also consistent with recent proposals that energetic constraints help shape the effective landscape over which cognition evolves (Kheirkhah et al., 2026; Markman & Otto, 2011). At the same time, energetic resources do not by themselves determine dynamics: the set of achievable states depends jointly on energy and system structure, including the geometry of the state space and the topology of passive transitions. Sustaining a nonequilibrium distribution, therefore, requires ongoing dissipation, linking energetic cost to the persistence of structured behavior (Horowitz & England, 2017). Energy thus constrains controllability without overriding structural limitations, implying that cognitive dynamics arise from the interaction between energetic resources and intrinsic system organization. Under these conditions, metabolic regulation naturally entails generative inference, insofar as agents must predict and selectively sample data so as to maintain their structural integrity (Sajid et al., 2021).

Importantly, the validity of this framework depends on the spatiotemporal scale at which its assumptions remain a reasonable approximation. Formal treatments of control cost and causal structure are typically developed for Markovian systems, while informational closure provides an additional criterion for identifying scales at which variables behave as partially autonomous dynamical descriptions (Bertschinger et al., 2008; Hoel, Albantakis, & Tononi, 2013). These conditions are unlikely to hold uniformly across scales, but may be approximately satisfied at microscopic or mesoscopic levels of neural organization. At the same time, cognitive variables of interest are usually defined at more macroscopic and behaviorally interpretable levels (Jazayeri & Afraz, 2017), creating tension if closure weakens at these levels. Thus, our framework is empirically testable: by combining metabolic measures with neural and behavioral data across tasks that vary in control cost, one can test at which scales energy expenditure, state transitions, and performance covary in line with model predictions.

Taken together, from this framework, three main insights arise: (i) Progress in cognitive modeling will not come from assigning metabolic costs to isolated cognitive operations, but from identifying mesoscopic variables that exhibit approximate dynamical autonomy and modeling their evolution as generative processes under energetic constraints. (ii) Optimality of cognition is not sufficiently defined by accuracy alone, but by the best achievable inference or control under given energetic resources. Metabolic constraints, therefore, do not merely prune the space of admissible models; they shape the effective geometry of reachable trajectories in state space. Cognition, in this sense, is not an unconstrained accuracy-maximizing engine, but a resource-bounded controller of probability flow, whose feasible trajectories are limited jointly by energetic budget and system structure. (iii) Efficient cognitive agents – whether biological or artificial – are intrinsically energy-constrained generative controllers, with explicit costs on deviations from default dynamics. Efficient architecture aligns desired behaviors with low-cost trajectories, leveraging structured representations and selective computation, essentially redefining optimality as performance maximization under the hard physical limits of energy and resource constraints.

Acknowledgments. The authors thank Marc Vöhringer Carrera for commenting on an earlier version of this manuscript.

Financial support. This work was funded in part by the consortia grants from the German Research Foundation (DFG) SFB/TRR 393 (project grant no 521379614) and by the Excellence Cluster EXC-3066 "The Adaptive Mind." Hamidreza Jamalabadi receives compensation from Springer Nature for his role as an Associate Editor.

Competing interests. The authors declare no competing interests.

References

Bertschinger, N., Olbrich, E., Ay, N., & Jost, J. (2008). Autonomy: An information theoretic perspective. *Biosystems*, *91*(2), 331–345.

Conant, R. C., & Ross Ashby, W. (1970). Every good regulator of a system must be a model of that system. *International Journal of Systems Science*, *1*(2), 89–97.

Fields, C., Goldstein, A., & Sandved-Smith, L. (2024). Making the thermodynamic cost of active inference explicit. *Entropy*, *26*(8), 622.

Friston, K. (2009). The free-energy principle: A rough guide to the brain? *Trends in Cognitive sciences*, *13*(7), 293–301.

Haimerl, C., Rodrigues, F. S., & Paton, J. J. (2025). Time, control, and the nervous system. *Annual Review of Neuroscience*, *48*:465–489. https://doi.org/10.1146/annurev-neuro-112723-025348.

Hoel, E. P., Albantakis, L., & Tononi, G. (2013). Quantifying causal emergence shows that macro can beat micro. *Proceedings of the National Academy of Sciences*, *110*(49), 19790–19795.

Horowitz, J. M., & England, J. L. (2017). Information-theoretic bound on the entropy production to maintain a classical nonequilibrium distribution using ancillary control. *Entropy*, *19*(7), 333.

Jamalabadi, H., Zuberer, A., Kumar, V. J., Li, M., Alizadeh, S., Amani, A. M., Gaser, C., Esterman, M., & Walter, M. (2021). The missing role of gray matter in studying brain controllability. *Network Neuroscience (Cambridge, Mass.)*, *5*(1), 198–210. https://doi.org/10.1162/netn_a_00174

Jazayeri, M., & Afraz, A. (2017). Navigating the neural space in search of the neural code. *Neuron*, *93*(5), 1003–1014.

Kawakita, G., Kamiya, S., Sasai, S., Kitazono, J., & Oizumi, M. (2022). Quantifying brain state transition cost via Schrödinger bridge. *Network Neuroscience*, *6*(1), 118–134.

Kheirkhah, M., Shariatpanahi, B., Hahn, T., Tiruvadi, V., Koppe, G., Nozari, E., Hofmann, S. G., & Jamalabadi, H. (2026). Re-engineering the disordered mind: Clinical experimentation, dynamical systems, and AI for personalized psychiatry. *Neuropsychopharmacology*, *51*(4), 718-721.

Landauer, R. (1961). Irreversibility and heat generation in the computing process. *IBM Journal of Research and Development*, *5*(3), 183–191.

Markman, A. B., & Otto, A. R. (2011). Cognitive systems optimize energy rather than information. *Behavioral and Brain Sciences*, *34*(4), 207.

Prigogine, I. (1978). Time, structure, and fluctuations. *Science*, *201*(4358), 777–785.

Sajid, N., Ball, P. J., Parr, T., & Friston, K. J. (2021). Active inference: Demystified and compared. *Neural Computation*, *33*(3), 674–712.

Sanders, J., Baldovin, M., & Muratore-Ginanneschi, P. (2024). Optimal control of underdamped systems: An analytic approach: J. Sanders et al. *Journal of Statistical Physics*, *191*(9), 117.

Takagi, K. (2021). Energy constraints on brain network formation. *Scientific Reports*, *11*(1), 11745.

Energy cost per operation in brains and AI hardware: What is more efficient?

Jan Karbowski*

Institute of Applied Mathematics and Mechanics, University of Warsaw, Warsaw, Poland
jkarbowski@mimuw.edu.pl

*Corresponding author.

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

doi:10.1017/S0140525X26106141, e465

Abstract

Haueis and Colaço point out that mammalian brains are much more energy efficient than AI supercomputers because the former use orders of magnitude less energy. I argue this issue is more nuanced, and the "efficiency" is scale dependent. Contrary to popular belief, I show that the human brain uses much more energy per operation than the most powerful AI supercomputers.

It is often said that the human brain is much more energy efficient than AI systems because the brain, while doing similarly complex cognitive tasks as AI, uses globally orders of magnitude less energy than the best AI supercomputers. It seems that the same common belief is shared by in their target article, which also lacks a direct comparison of energy efficiency between neural and AI systems. The goal of this comment is to challenge this notion by presenting specific hard data on the costs of computations in mammalian brains and AI machines. Using "the energy per operation" as a measure of computational efficiency, I show that the issue is more complex and level dependent and that the human brain is more expensive in this regard than the best AI machines.

The human brain consumes about 20 W of metabolic energy (Aiello & Wheeler, 1995), contains 86×10^9 neurons (Azevedo et al., 2009), and neurons fire action potentials at the rate of 1 Hz (Karbowski, 2009, 2012). Assuming that an action potential corresponds to a single neural operation, this means that the human brain uses about 200 pJ/op. A much more accurate cost of a single action potential, based on detailed theoretical analysis and simulations, yields (Attwell & Laughlin, 2001; Karbowski, 2009): 80–130 pJ/op for human cortex and 40 pJ/op for mouse cortex. (For a mouse neuron, an action potential requires removal of about 15×10^8 sodium ions from the neuronal interior that costs 5×10^8 ATP molecules, and 1 ATP uses 0.083 aJ energy. A human neuron uses 2–3 times more energy because its volume is 2–3 times larger).

Let us compare these numbers to the corresponding figures in AI semiconductors. AI devices can be broadly divided into two classes: fast supercomputers with von Neumann architecture and slower neuromorphic systems without von Neumann architecture serving specific purpose computations. The first group has a very different structure from mammalian brains (with spatially separate memory and decision modules), whereas the second tries to mimic the brain architecture (spatially close memory and CPU). The fastest supercomputer in the first group, known as "El Capitan," contains 6×10^{15} transistors, performs 2.8×10^{18} operations/sec, and dissipates 35 MW of energy (Science & Technology Review 2024), which gives 12.5 pJ/op. This implies that this "brute force" machine is 6–10 times more energy efficient per operation than neurons per action potential in the human brain. This is the first sign that biological information processing is costly on a microscale.

The second class of AI devices with neuromorphic design is much slower than the above supercomputer but also more energy efficient. The biggest neuromorphic system in the world, known as "Hala Point" with 10^9 artificial neurons (128 artificial synapses per neuron) and 2.6×10^{12} transistors, uses 2600 W and performs 2×10^{16} operations/sec (Intel Newsroom 2024). This gives the efficiency of 0.13 pJ/op (10^7 op/sec per artificial neuron), which is 100 times better than for the fast supercomputers and 600–1000 times better than the cost of operation in a human neuron. This means that current neuromorphic systems can achieve much better performance on a microscale in terms of energy/op than the human brain. However, on the other hand, the energy used per artificial neuron is 13×10^5 pJ/sec, which is 10000 times larger than in a human neuron. This is another type of "energy efficiency" (per second), and that one is won by the biological neurons.

The last numbers suggest that one has to be careful with the notion of energy efficiency because the context, i.e., energy per what, matters. Consequently, let us analyze the efficiency on a more fundamental, smaller scale, corresponding to transistors (in AI) and synapses (in brains).

The El Capitan supercomputer and Hala Point neuromorphic computer are assembled with the same type of 7 nm transistors (FinFET technology) that use about 1 aJ/op (Larsh et al., 2025). Thus, on a nanoscale, the efficiency of computation increases dramatically by a factor of 10^5–10^7, and it is the same in both computers. These numbers indicate that 1.8×10^7 transistors are involved in one operation in El Capitan and 10^5 transistors in Hala Point. On the other hand, one synapse uses 2×10^5 ATP molecules per synaptic transmission (Laughlin, de Ruyter van Steveninck, & Anderson, 1998; Karbowski, 2012), which is 0.017 pJ/op. This is 4700 times more energy efficient than in a neuron, but still 10^4 less efficient than in a transistor. However, one can argue that a typical synapse is about 1000 times larger than the 7 nm transistor and contains roughly 10^4 proteins (Sheng & Hoogenraad, 2007), which also consume energy and take part in synaptic plasticity, responsible for learning and memory (Lisman, Yasuda, & Raghavachari, 2012). This suggests that the efficiency per synaptic protein is probably similar to the efficiency of a modern transistor.

What are the implications? First, a single action potential in the brain is much more costly than a single operation in AI hardware, and thus it is more valuable. That is why mammalian brains are ultraslow in comparison to AI machines and employ a "sparse coding" strategy (to do things using as little energy as possible). Second, efficiency becomes comparable for brains and AI hardware on the nanoscale, on a level of synaptic proteins and transistors. Despite these similarities for the energy cost of information processing in the nanoscale, the global "cognition" of brains and AI machines differs significantly, suggesting that other factors also matter. These might be related to differences in, e.g., connectivity architecture (denser connections in the brain; Karbowski, 2014) and information carriers (much more flexible synaptic proteins in the brains than electrons in transistors; Lisman et al., 2012).

Financial support. The work was supported by the Polish Narodowe Centrum Nauki (NCN) grant number 2021/41/B/ST3/04300.

Competing interests. None.

References

Aiello, L. C., & Wheeler, P. (1995). The expensive-tissue hypothesis: The brain and the digestive-system in human and primate evolution. *Current Anthropology*, *36*, 199–221.

Attwell, D., & Laughlin, S. B. (2001). An energy budget for signaling in the gray matter of the brain. *Journal of Cerebral Blood Flow & Metabolism*, *21*, 1133–1145.

Azevedo, F. A., Carvalho, L. R., Grinberg, L. T., Farfel, J. M., Ferretti, R. E. L., Leite, R. E. P., Filho, W. J., Lent, R., & Herculano-Houzel, S. (2009). Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. *Journal of Comparative Neurology*, *513*, 532–541.

Intel Newsroom. (2024). Intel builds world's largest neuromorphic system to enable more sustainable AI. https://newsroom.intel.com/artificial-intellige

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

nce/intel-builds-worlds-largest-neuromorphic-system-to-enable-more-sustainable-ai.
Karbowski, J. (2009). Thermodynamic constraints on neural dimensions, firing rates, brain temperature and size. *Journal of Computational Neuroscience*, *27*, 415–436.
Karbowski, J. (2012). Approximate invariance of metabolic energy per synapse during development in mammalian brains. *PLoS ONE*, *7*, e33425.
Karbowski, J. (2014). Constancy and trade-offs in the neuroanatomical and metabolic design of the cerebral cortex. *Frontiers in Neural Circuits*, *8*, 9.
Larsh, L., Siddique, R., & Banad, S. S. Y. M. (2025). Ultra-low-power spiking neurons in 7 nm FinFET technology: A comparative analysis of leaky integrate-and-fire, Morris-Lecar, and axon-hillock architectures. Arxiv:2505.03764.
Laughlin, S. B., de Ruyter van Steveninck, R. R., & Anderson, J. C. (1998). The metabolic cost of neural information. *Nature Neuroscience*, *1*, 36–40.
Lisman, J., Yasuda, R., & Raghavachari, S. (2012). Mechanisms of CaMKII action in long-term potentiation. *Nature Reviews Neuroscience*, *13*, 169–182.
Science and Technology Review (2024). Introducing El Capitan. https://str.llnl.gov/str-december-2024/introducing-el-capitan.
Sheng, M., & Hoogenraad, C. C. (2007). The postsynaptic architecture of excitatory synapses: A more quantitative view. *Annual Review of Biochemistry*, *76*, 823–847.

Blind ignorance or rational inattention? On the history and future of metabolic considerations for cognitive models

Ari Khoudary[a,b,c*], Aaron M. Bornstein[a,b,c] and Megan A. K. Peters[a,b,c,d,e,f]

[a]Cognitive Sciences, University of California, Irvine, CA, USA; [b]Center for Theoretical Behavioral Sciences, University of California, Irvine, CA, USA; [c]Center for the Neurobiology of Learning and Memory, University of California, Irvine, CA, USA; [d]Logic and Philosophy of Science, University of California, Irvine, CA, USA; [e]Experimental Psychology, University College London, London, UK and [f]Program in Brain, Mind, and Consciousness, Canadian Institute for Advanced Research, Toronto, ON, Canada
ari.khoudary@uci.edu
aaron.bornstein@uci.edu
megan.peters@uci.edu

*Corresponding author.

doi:10.1017/S0140525X26106293, e466

Abstract

We build upon Haueis and Colaço's account in two ways. First, we argue that the backgrounding of metabolism was a pragmatic response to a *data granularity problem* faced by traditional cognitive modelers. Then, we suggest that *equivalence demonstrations* offer a promising pathway for incrementally introducing metabolic considerations into extant modeling practice.

Haueis and Colaço's (H&C) account asserts a *normative* role for metabolism in cognitive modeling: in general, models that incorporate metabolic considerations ought to be preferred over those that don't. This raises an important question left unaddressed by the authors: *why* hasn't metabolism played a normative role in cognitive modeling? Our comment offers an answer – the *data granularity problem* – which subsequently informs our suggestion to use *equivalence demonstrations* as a method for integrating metabolic considerations into existing cognitive modeling research.

Cognitive modeling has its roots in mathematical psychology, which emerged as a subdiscipline of psychology in the 1950s. The goal of this field was to build formally specified theories describing the cognitive constructs and processes that generate human behavior in experimental and real-world settings (Batchelder, 2010). These theories are tested by application to behavioral data. Modelers first define a set of mathematical objects and relationships that are intended to represent cognitive constructs and processes (i.e., they specify a model *structure*). Then, they identify numerical values for model variables that minimize the difference between the model's predictions and measurements of human behavior (i.e., they *fit* free parameters of the model to data). Inferences about *which* constructs and processes best explain behavior are then typically conducted by comparing model performance metrics across competing models (i.e., model comparison) and/or performing statistical tests on fitted model parameters.

Contrasting this description of a conventional cognitive model with H&C's definition of "idealized, abstract systems whose elements stand in a relation to assumed elements of the target capacity" (p. 2) reveals an important conceptual gap. Namely, H&C have in mind the broad space of formal structures that *could* function as models of cognitive processes, rather than the narrower space of formal structures that have *actually functioned* as "cognitive models" for decades. Understanding the historical context within which the backgrounding of metabolism became a norm is essential for identifying concrete, tractable, and actionable proposals for future research.

We suggest that the "backgrounding" of metabolic considerations emerged as a pragmatic response to empirical constraints on the model-based inference process. For a parameter value to be meaningfully estimated via behavior, it has to stand in for a latent component that plausibly (i) changes as a function of task conditions, (ii) differs across individuals performing the same task, or (iii) is some combination of (i) and (ii). Adding terms that do not contribute to structural variability in behavioral data risks *overfitting* the parameter values to idiosyncrasies of the sample, resulting in poor out-of-sample generalization. Further, as model complexity increases, so too does the probability of two competing models *mimicking* each other: making identical predictions about behavior on the basis of qualitatively distinct model structures (Khodadadi & Townsend, 2015; Zhang et al., 2014). To defend against these perils, modelers relied heavily on the *principle of parsimony* to guide model generation and evaluation: all else being equal, simpler models should be preferred (Vandekerckhove, Matzke, & Wagenmakers, 2015).

These historical considerations thus reveal that a *data granularity problem* figured centrally in the backgrounding of metabolism: the parameters of cognitive models are estimated by application to behavior, but behavioral data are often too coarse-

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

grained to meaningfully constrain model parameters related to metabolism. One solution to this problem can be incorporating biological measurements into the data that cognitive models are tasked with fitting, such that parameter values must be estimated jointly using behavior and brain data (Turner, Forstmann, & Steyvers, 2019). These approaches have had some success in adjudicating among mimicking models (Kelly, Corbett, & O'Connell, 2021), but (1) it is not practical to require that all cognitive model parameters be estimated jointly to neural and behavioral data and (2) "joint" models face their own challenges related to generalizability, interpretability, etc. (Churchland & Kiani, 2016).

As an alternative, we propose that *equivalence demonstrations* – formal derivations or numerical simulations identifying equivalence between two differently specified models – are a practical and powerful way to incrementally incorporate metabolism into cognitive modeling. This approach is exemplified in a recent study by Wang and Donkin (2024), who demonstrated that relaxing an empirically supported assumption in a prominent cognitive model increases its similarity to a more complex model that asserts a specific biological mechanism for the data-generating process. While this modification does not change the original model's ability to fit behavior, it increases the validity of interpreting its fits in light of biological processes. The authors essentially *leveraged* model mimicry – or functional equivalence – to their advantage by identifying simple models that mimic more complex, biologically inspired alternatives.

In this vein, we suggest that the field engage in *re-expressing* existing models in information-theoretic terms. Identifying information-theoretic "conversions" of prominent cognitive models would allow for investigating the metabolic costs of different operations within those models, making the *models themselves* "legible" with respect to metabolic facts rather than making *metabolic knowledge* "legible" or compatible with a given model's formulation (as suggested by H&C in §2.1.5). Our suggestion exemplifies the *generative* and *evaluative* functions of metabolism proposed by H&C: it "populates" the space with new functional forms, which themselves serve to "partition" the space of *existing* models with respect to their metabolic feasibility. Crucially, because these information-theoretic models are *not* tasked with fitting behavior directly, they need not abide by the strict parsimony constraints shaping traditional cognitive models. Instead, modelers can focus their attention on exploring the vast space of functional forms that relate metabolic considerations to cognitive functions using the "common currency" of information theory.

In sum, the backgrounding of metabolism in cognitive modeling stems not from blind ignorance or unwarranted dismissal of metabolic considerations, but principled constraints on the types of inference warranted by the historical *explananda* of cognitive models: behavioral data. The norms guiding cognitive modeling structurally disincentivize direct representation of metabolic processes, which poses a challenge for implementing H&C's prescription. Building new classes of *equivalent* models that represent relevant biological information at varying levels of detail both offers relief from these structural constraints and directly facilitates the generative and evaluative functions of metabolic considerations.

Financial support. This work was funded by a grant from the Research Corporation for Science Advancement (to M.A.K.P.) and the Canadian Institute for Advanced Research (fellowship in the Brain, Mind, & Consciousness Program, to M.A.K.P.).

Competing interests. None.

References

Batchelder, W. H. (2010). Mathematical psychology. *Wiley Interdisciplinary Reviews. Cognitive Science*, *1*(5), 759–765.

Churchland, A. K., & Kiani, R. (2016). Three challenges for connecting model to mechanism in decision-making. *Current Opinion in Behavioral Sciences*, *11*, 74–80.

Kelly, S. P., Corbett, E. A., & O'Connell, R. G. (2021). Neurocomputational mechanisms of prior-informed perceptual decision-making in humans. *Nature Human Behaviour*, *5*(4), 467–481.

Khodadadi, A., & Townsend, J. T. (2015). On mimicry among sequential sampling models. *Journal of Mathematical Psychology*, *68–69*, 37–48.

Turner, B. M., Forstmann, B. U., & Steyvers, M. (2019). *Joint models of neural and behavioral data*. (1st ed.) [PDF]. Springer Nature.

Vandekerckhove, J., Matzke, D., & Wagenmakers, E. (2015). Model comparison and the principle of parsimony. In J. R. Busemeyer, Z. Wang, J. T. Townsend, & A. Eidels (Eds.), *The oxford handbook of computational and mathematical psychology* (pp. 300–319). Oxford University Press.

Wang, J.-S., & Donkin, C. (2024). The neural implausibility of the diffusion decision model doesn't matter for cognitive psychometrics, but the Ornstein-Uhlenbeck model is better. *Psychonomic Bulletin & Review*, *31*, 1–13.

Zhang, S., Lee, M. D., Vandekerckhove, J., Maris, G., & Wagenmakers, E.-J. (2014). Time-varying boundaries for diffusion models of decision making and response time. *Frontiers in Psychology*, *5*, 1364.

Why metabolism remains implicit in cognitive models

Kacper Kondrakiewicz[a]* and Pedro J. Gonçalves[a,b,c]

[a]VIB Center for Neuroscience Leuven, VIB, Leuven, Belgium; [b]VIB Center for AI and Computational Biology, VIB, Leuven, Belgium and [c]Departments of Computer Science and Electrical Engineering, KU Leuven, Leuven, Belgium
kacper.kondrakiewicz@kuleuven.be
pedro.goncalves@kuleuven.be

*Corresponding author.

doi:10.1017/S0140525X26106177, e467

Abstract

Explicitly modeling metabolic constraints remains rare in cognitive neuroscience due to methodological, empirical, and conceptual challenges. Although metabolic limits already shape the biological data that many models are fitted to, direct measurements of energetic costs remain scarce and difficult to relate to model variables. We argue that recent experimental and computational advances may soon enable more explicit integration of metabolism into cognitive models.

The target article makes a case that metabolic constraints should play a more prominent role in cognitive modeling. While we agree with this premise, we examine why explicit metabolic modelling remains uncommon, using Marr's levels of analysis as a clarifying framework. In particular, Marr's perspective helps highlight how metabolic constraints can shape models, even when they are not represented explicitly.

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

David Marr famously proposed that information processing systems can be understood at three levels: the computational level (what abstract function the system performs), the algorithmic level (what representations and elementary operations realize this function), and the implementational level (how these operations are physically implemented) (Marr, 1982). Importantly, in the case of brains, the implementation level can entail multiple biological dimensions, such as electrical activity, synaptic transmission, development, connectivity, and metabolism. Although this framework is not without limitations, it remains a useful way to separate different types of explanations in computational neuroscience (Lengyel, 2024; Poggio, 2012).

Equipped with this terminology, we can rephrase the main thesis of the target article: the metabolic aspects of implementation should be included in cognitive models more often, because they provide useful constraints on other processes, both within the implementation level (e.g., membrane and synaptic biophysics) and across levels (algorithm and computation). While we agree with the diagnosis, we find it useful to consider why metabolism is not modeled more often.

One reason is methodological: researchers tend to model what can be measured and fitted. During the last few decades, neuroscientists have routinely collected vast amounts of data on neuronal activity *in vivo* (spiking, local field potentials, hemodynamic response, etc.). Recording energy expenditure (e.g., using ATP sensors) remains much more challenging, even despite recent methodological advances (Zhao et al., 2025). As a consequence, the researchers willing to model metabolic cost often lack ground truth data. Let us consider a concrete example: estimating the energy cost of a reward-related burst in a dopamine neuron. Because direct metabolic data are not available, one might in principle use the ATP cost per spike obtained from cortical neurons (Howarth, Gleeson, & Attwell, 2012). However, dopamine neurons likely consume much more energy due to their complex morphology (Pissadaki & Bolam, 2013). More importantly, it is challenging to estimate how much energy could be saved by reducing spiking (Kondrakiewicz & Nawrocka, 2025), as recent models suggest that dopamine neurons may need to fire periodically to prevent toxic buildup of free radicals (Chintaluri & Vogels, 2023). This illustrates how basic metabolic questions remain unanswered, leading many researchers to avoid metabolic assumptions altogether.

Another reason for ignoring metabolism is conceptual. Computational models are often designed to capture simple, general principles that approximate the behavior of real-world complex systems, while remaining amenable to analysis and interpretation. A well-known example is the backpropagation algorithm, originally developed to update synaptic weights in deep neural networks for learning complex tasks (Rumelhart, Hinton, & Williams, 1986). Although long assumed to be biologically unrealistic, it became the standard approach for training neural networks due to its effectiveness. Only recently have researchers discovered biologically plausible learning rules that can approximate aspects of the same computation (Lillicrap et al., 2020). This illustrates how abstracting away from implementation details to extract general principles can yield insights relevant for understanding computation in the brain.

Despite these limitations, the relative scarcity of explicit metabolic models does not imply that metabolic constraints play no role in current modeling practice. Instead, metabolic constraints often influence models *implicitly*.

First, many computational models are fitted to biological data. Those data already reflect the energetic limits of real neurons. If a model replicates, for instance, spiking statistics from experimental data, we can conclude that the model output is broadly consistent with the metabolic constraints on spiking, even if these constraints are not modeled explicitly. However, such agreement with observed activity patterns does not uniquely determine the underlying mechanisms, as different parameter configurations can give rise to similar activity patterns (Prinz, Bucher, & Marder, 2004). This in turn means that these configurations may differ in their associated energy costs, or achieve similar costs through different mechanisms, further complicating the relationship between function and metabolism (Deistler, Macke, & Gonçalves, 2022). That said, model fitting remains a powerful approach for assessing whether a model is generally consistent with implementation-level constraints without having to commit to specific and potentially unknown quantitative assumptions about those constraints.

Second, computation- and algorithmic-level modeling often invokes efficiency as a guiding principle. In particular, the efficient coding hypothesis is built on the idea that, due to evolutionary pressure, biological systems use their limited resources to represent information as efficiently as possible, for example, by reducing redundancy in neural representations (Barlow & Rosenblith, 1961; Simoncelli & Olshausen, 2001). The target article correctly classifies these approaches as "generic," since energetic costs are rarely explicit. However, these approaches can remain effective even without ground-truth metabolic data. For example, efficient coding models correctly predicted that the early visual system reduces spatial and temporal correlations to improve coding efficiency (Atick & Redlich, 1990; Dan, Atick, & Reid, 1996).

To sum up, we argue that the common practice of omitting explicit metabolic details reflects not a lack of interest, but the difficulty of integrating metabolism with other levels of analysis. At the same time, many existing modeling approaches already capture metabolic constraints implicitly. Does this mean that explicitly incorporating metabolic constraints is unnecessary, or that it will remain challenging? We do not think so. Interest in brain metabolism is growing, which will hopefully lead to the collection of much needed data on the topic (Bergmann et al., 2025; Padamsey & Rochefort, 2023). Combined with advances in computational methods and the growing public sharing of resources, this progress should make metabolic modeling more accessible than ever before. Ideally, these technical advances will enable researchers to effortlessly add a metabolic layer to existing computational models. This, in turn, would open the door to addressing questions that directly link levels of analysis, such as the relationship between algorithmic and metabolic efficiency, and the extent to which metabolic constraints shape or restrict the space of possible computations (Jedlicka, Bird, & Cuntz, 2022; Pallasdies et al., 2021). We anticipate that the coming decade will bring meaningful progress toward answering these and related questions.

Acknowledgments. We thank Ludwika Kondrakiewicz and Janko Petkovic for discussions and comments on the manuscript.

Financial support. KK was funded by the European Innovation Council grant (NEUROGATE, project no. 101213906)

PJG is thankful for the financial support from the Flemish government through long-term structural funding Methusalem (grant METH/26/003).

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

Competing interests. The authors declare no conflict of interests.

References

Atick, J. J., & Norman Redlich, A. (1990). Towards a theory of early visual processing. *Neural Computation*, *2*(3), 308–320. https://doi.org/10.1162/neco.1990.2.3.308.

Barlow, H. B., & Rosenblith, W. A. (1961). Possible Principles Underlying the Transformations of Sensory Messages. 217–234.

Bergmann, C., Mousaei, K., Rizzoli, S. O., & Tchumatchenko, T. (2025). How energy determines spatial localisation and copy number of molecules in neurons. *Nature Communications*, *16*(1), 1424. https://doi.org/10.1038/s41467-025-56640-0.

Chintaluri, C., & Vogels, T. P. (2023). Metabolically regulated spiking could serve neuronal energy homeostasis and protect from reactive oxygen species. *Proceedings of the National Academy of Sciences*, *120*(48), e2306525120. https://doi.org/10.1073/pnas.2306525120.

Dan, Y., Atick, J. J., & Reid, R. C. (1996). Efficient coding of natural scenes in the lateral geniculate nucleus: experimental test of a computational theory. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *16*(10), 3351–3362. https://doi.org/10.1523/JNEUROSCI.16-10-03351.1996.

Deistler, M., Macke, J. H., & Gonçalves, P. J. (2022). Energy-efficient network activity from disparate circuit parameters. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(44), e2207632119. https://doi.org/10.1073/pnas.2207632119.

Howarth, C., Gleeson, P., & Attwell, D. (2012). Updated energy budgets for neural computation in the neocortex and cerebellum. *Journal of Cerebral Blood Flow and Metabolism: Official Journal of the International Society of Cerebral Blood Flow and Metabolism*, *32*(7), 1222–1232. https://doi.org/10.1038/jcbfm.2012.35.

Jedlicka, P., Bird, A. D., & Cuntz, H. (2022). Pareto optimality, economy–effectiveness trade-offs and ion channel degeneracy: Improving population modelling for single neurons. *Open Biology*, *12*(7), 220073. https://doi.org/10.1098/rsob.220073.

Kondrakiewicz, K., & Nawrocka, J. (2025). Brains are expensive, but cognition is often cheap. *Neuroscience and Biobehavioral Reviews*, *179*, 106450. https://doi.org/10.1016/j.neubiorev.2025.106450.

Lengyel, M. (2024). Crosstalk rebuttal: Marr's three levels of analysis are not useful as a framework for neuroscience. *The Journal of Physiology*, *602*(9), 1919. https://doi.org/10.1113/JP286426.

Lillicrap, T. P., Santoro, A., Marris, L., Akerman, C. J., & Hinton, G. (2020). Backpropagation and the brain. *Nature Reviews Neuroscience*, *21*(6), 335–346. https://doi.org/10.1038/s41583-020-0277-3.

Marr, D. (1982). *Vision*. W. H. Freeman.

Padamsey, Z., & Rochefort, N. L. (2023). Paying the brain's energy bill. *Current Opinion in Neurobiology*, *78*, 102668. https://doi.org/10.1016/j.conb.2022.102668.

Pallasdies, F., Norton, P., Schleimer, J., & Schreiber, S. (2021). Neural optimization: Understanding trade-offs with pareto theory. *Current Opinion in Neurobiology*, *71*, 84–91. https://doi.org/10.1016/j.conb.2021.08.008.

Pissadaki, E. K., & Paul Bolam, J. (2013). The energy cost of action potential propagation in dopamine neurons: Clues to susceptibility in Parkinson's disease. *Frontiers in Computational Neuroscience*, *7*(13), 1–17. https://doi.org/10.3389/fncom.2013.00013.

Poggio, T. (2012). The levels of understanding framework, revised. *Perception*, *41*(9), 1017–1023. https://doi.org/10.1068/p7299.

Prinz, A. A., Bucher, D., & Marder, E. (2004). Similar network activity from disparate circuit parameters. *Nature Neuroscience*, *7*(12), 1345–1352. https://doi.org/10.1038/nn1352.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, *323*(6088), 533–536. https://doi.org/10.1038/323533a0.

Simoncelli, E. P., & Olshausen, B. A. (2001). Natural image statistics and neural representation. *Annual Review of Neuroscience*, *24*, 1193–1216. https://doi.org/10.1146/annurev.neuro.24.1.1193.

Zhao, M., Tanourlouee, S. B., McCracken, S., & Williams, P. R. (2025). Genetically encoded biosensors of metabolic function for the study of neurodegeneration, a review and perspective. *Neurophotonics*, *12*(Suppl 2), S22805. https://doi.org/10.1117/1.NPh.12.S2.S22805.

The scale mismatch problem in metabolic cognitive modeling

Xiang-Zhen Kong[a,b,c*], Yuhao Duan[a,b,c] and Yi Pu[d*]

[a]Department of Psychology and Behavioral Sciences, Zhejiang University, China; [b]The State Key Lab of Brain-Machine Intelligence, Zhejiang University, Hangzhou, China; [c]Zhejiang Key Laboratory of Neurocognitive Development and Mental Health, Zhejiang University, Hangzhou, China and [d]Shanghai Key Laboratory of Brain Functional Genomics (Ministry of Education), School of Psychology and Cognitive Science, East China Normal University, China

xiangzhen.kong@zju.edu.cn
0924809@zju.edu.cn
ypu@psy.ecnu.edu.cn

*Corresponding author.

doi:10.1017/S0140525X26106165, e468

Abstract

Haueis and Colaço compellingly argue for integrating metabolic considerations into cognitive modeling. We suggest that current discussions suffer from a scale mismatch: metabolic constraints are typically framed at global or microscopic levels, whereas cognitive models may operate at a systems level. We propose mesoscale metabolic properties of functional neural networks as a more promising interface for constraining cognitive models.

Haueis and Colaço provide a timely and compelling case for integrating metabolic considerations into cognitive modeling. Their distinction between evaluative and generative uses of metabolism offers an important advance beyond metaphor toward substantive constraint. However, their account remains underspecified with respect to scale. In practice, metabolic knowledge is not equally actionable for cognitive modelers, because it is often articulated at scales that are mismatched with the explanatory target of cognitive science. We argue that a scale mismatch problem pervades current discourse: metabolic data are typically either too molecular or too global to directly constrain models of cognitive systems. We propose that mesoscale metabolism, the metabolic phenotypes of functional neural circuits and networks, provides a more promising interface for metabolically informed cognitive modeling.

Cognitive models may operate at the level of functional systems, circuits, and networks. By contrast, much of metabolic knowledge resides at extremes of scales. At the molecular level, precise metrics such as ATP cost for synaptic transmission (Harris, Jolivet, & Attwell, 2012) or plasticity (Karbowski, 2019) are indispensable for biophysical realism, yet they rarely map directly onto the

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

functional descriptions employed in cognitive models. At the global level, whole-brain energy budgets (Attwell & Laughlin, 2001; Engl & Attwell, 2015) provide important upper bounds on information processing but they are too coarse to adjudicate between competing models of specific cognitive capacities. This mismatch helps explain why metabolism is often treated as "background" in cognitive modeling: not because it is irrelevant, but because it is typically presented at scales incommensurate with cognitive explanation.

A more promising interface lies at the mesoscale: the level of functional neural circuits, pathways, and networks that directly implement cognitive operations. Certain metabolic phenomena are inherently mesoscale. Aerobic glycolysis, for example, is preferentially upregulated in specific functional networks and covaries with their computational roles (Theriault et al., 2023). Likewise, neurovascular coupling and region-specific cerebral blood flow track the metabolic demands of distributed cognitive operations rather than local spiking alone. Crucially, recent large-scale mapping of mitochondrial respiratory capacity demonstrates that different mitochondrial profiles, including mitochondrial density, energy-transformation capacity, and oxidative potential, are systematically heterogeneous across the human brain (Figure 1) (Mosharov et al., 2025). This network- and region-specific patterning implies that mitochondrial energetics constitute a mesoscale property of neural systems, aligning with functional specialization. These phenomena are therefore not merely cellular processes, but metabolic phenotypes of cognitive systems.

Mesoscale metabolic features therefore offer constraints that are both biologically grounded and cognitively specific. They allow evaluative metabolic considerations to move beyond generic efficiency claims toward capacity-specific adjudication. For example, models positing sustained neural activity can be evaluated against evidence for rhythmic or transient metabolic engagement at the circuit level. Likewise, generative modeling can treat mesoscale metabolic regimes as sources of architectural and dynamical

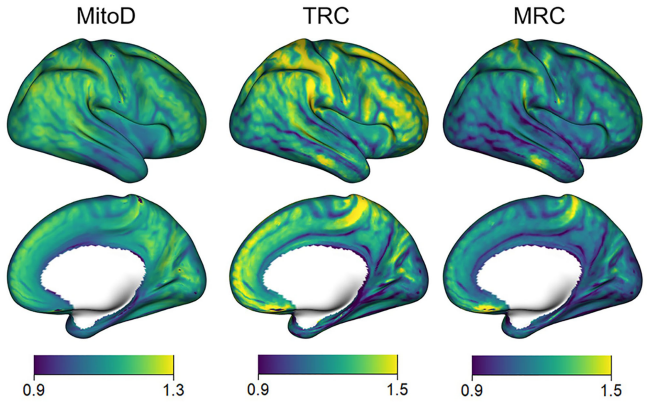


Figure 1 (Kong et al.). Mesoscale organization of mitochondrial energetics in the human brain. Brain-wide mapping reveals pronounced heterogeneity in mitochondrial respiratory capacity across the human brain. Data are referenced to the brain maps of mitochondrial profiles reported by Mosharov et al. (2025). MitoD denotes mitochondrial tissue density. TRC reflects tissue respiratory capacity. MRC (mitochondrial respiratory capacity) indexes TRC normalized for MitoD.

hypotheses, such as metabolically informed gating or gain control in predictive coding frameworks. In this way, metabolism becomes a direct driver of model structure, not merely a post hoc plausibility check.

This perspective operationalizes Haueis & Colaço's framework. For evaluative modeling, the specification of relevant metabolic constraints should prioritize mesoscale phenotypes tied to the target capacity. For generative modeling, network-level metabolic features provide natural parameters for model construction. Crucially, a mesoscale focus enables genuine multiscale integration: molecular cost and global budgets can be respected without collapsing explanation to either extreme.

We therefore agree that metabolism matters for cognitive modeling but argue that its impact depends critically on scale. Without an explicit mesoscale focus, metabolic considerations risk remaining either too abstract to constrain or too detailed to integrate. Mesoscale metabolism is where computation and energy genuinely meet, as cognitive algorithms are implemented in biological networks with finite, structured energetic resources. Future work should adopt a scale-sensitive framework that explicitly identifies which metabolic features constrain which levels of cognitive explanation. The path to metabolically constrained cognitive modeling runs through the mesoscale.

Financial support. XK was supported by the Fundamental Research Funds for the Central Universities (226-2025-00144). YP was supported by the National Natural Science Foundation of China (32400882).

Competing interests. The authors declare that there are no competing interests.

References

Attwell, D., & Laughlin, S. B. (2001). An energy budget for signaling in the grey matter of the brain. *Journal of Cerebral Blood Flow & Metabolism*, *21*(10), 1133–1145. https://doi.org/10.1097/00004647-200110000-00001

Engl, E., & Attwell, D. (2015). Non-signalling energy use in the brain. *The Journal of Physiology*, *593*(16), 3417–3429. https://doi.org/10.1113/jphysiol.2014.282517

Harris, J. J., Jolivet, R., & Attwell, D. (2012). Synaptic energy use and supply. *Neuron*, *75*(5), 762–777. https://doi.org/10.1016/j.neuron.2012.08.019

Karbowski, J. (2019). Metabolic constraints on synaptic learning and memory. *Journal of Neurophysiology*, *122*(4), 1473–1490. https://doi.org/10.1152/jn.00092.2019

Mosharov, E. V., Rosenberg, A. M., Monzel, A. S., Osto, C. A., Stiles, L., Rosoklija, G. B., Dwork, A. J., Bindra, S., Junker, A., Zhang, Y., Fujita, M., Mariani, M. B., Bakalian, M., Sulzer, D., De Jager, P. L., Menon, V., Shirihai, O. S., Mann, J. J., Underwood, M. D., Boldrini, M., de Schotten, M. T., & Picard, M. (2025). A human brain map of mitochondrial respiratory capacity and diversity. *Nature*, *641*(8063), 749–758. https://doi.org/10.1038/s41586-025-08740-6

Theriault, J. E., Shaffer, C., Dienel, G. A., Sander, C. Y., Hooker, J. M., Dickerson, B. C., Barrett, L. F., & Quigley, K. S. (2023). A functional account of stimulation-based aerobic glycolysis and its role in interpreting BOLD signal intensity increases in neuroimaging experiments. *Neuroscience & Biobehavioral Reviews*, *153*, 105373. https://doi.org/10.1016/j.neubiorev.2023.105373

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

Medium independence and cognitive ontology

Caitlin Mace[a], Zoe Drayson[b]* and Sarah Robins[c]

[a]University of Pittsburgh, USA; [b]University of California Davis, USA and [c]Purdue University, USA
caitlin.mace@pitt.edu
zdrayson@ucdavis.edu
skrobins@purdue.edu

*Corresponding author.

doi:10.1017/S0140525X26106256, e469

Abstract

This commentary emphasizes the importance of how we individuate cognitive capacities, operations, and vehicles. It challenges the target article's reliance on mechanistic notions of computation and medium-independence and uses examples from the memory literature to suggest that the role played by metabolic considerations in cognitive models will depend on questions of taxonomy and cognitive ontology.

In their target article, Haueis and Colaço acknowledge that there is a tension between the "abstract, formal" characterization of a computational physical system in terms of its medium-independence and a physical characterization of the same system which would include its metabolic properties. Their attempt to identify and address this tension is hampered, however, by their focus on Piccinini's idiosyncratic notion of physical computation. Piccinini (2015) specifically attempts to define and individuate the computational properties of a physical system in terms of concrete mechanisms, without invoking the notion of a formal property. It is widely acknowledged, however, that such an account of physical computation cannot make sense of Haugeland's (1985) concept of medium-independence, on which the formal procedure being executed by a physical system can be identified and described independently of any physical medium. (See Coelho Mollo (2017), Drayson (2025), Haimovici (2013), and Williams (2025) for further discussion).

The real tension for Haueis and Colaço, we take it, is that *if* a particular type of cognitive process can be individuated formally (e.g. algorithmically), then non-formal properties of a physical system which implements the formalism are irrelevant to the cognitive process *qua* formal computation. If Haueis and Colaço have shown that an appeal to metabolic properties of the physical system is necessary to identify the cognitive process in question, however, then they have demonstrated that there is a non-formal way to individuate that type of cognitive process. This would be an important result because it would show that the cognitive process is not computational (in the formal sense): details of its physical implementation are essential to its identity as that type of cognitive process.

Whether Haueis and Colaço will be able to achieve this result at an appreciable scope depends on their ability to secure a role for metabolic considerations in the evaluation of competing cognitive models. Haueis and Colaço show how this can be done when comparing models of synaptic memory storage. The case is compelling, but it is unclear how the lessons can be carried forward to broader models of memory. Their example is highly constrained: the cognitive capacity (memory), neural system (dentate gyrus), and computational operation (synaptic weighting) are fixed. This is rarely the case in the evaluative comparison of cognitive models. How to individuate capacities, operations, and vehicles is central to debates between cognitive models, and it is much less clear how to incorporate metabolism when one or more of these are in flux.

To illustrate this, consider memory models at a broader scale. Even granting that changes in synaptic weighting are the best model for synaptic memory storage, vehicle indeterminacy for memories remains. After all, a memory is not stored at a single synapse. How should changes in synaptic weighting be organized into vehicles of memory storage? Competing models often propose disparate structures or emphasize different properties of those structures, resulting in an indeterminacy of vehicles (Mace, 2025b). Spike trains, ensemble level properties, and field potentials are all contenders, and competing models of each emphasize different properties. For example, debates persist about when rate, timing, and synchrony of spike trains are important, as well as whether those properties can be sampled from individual neurons in different trials (Cariani & Baker, 2022; Ebitz & Hayden, 2021). Without having worked out vehicle indeterminacy issues in advance, it is difficult to determine which non-generic metabolic considerations matter and which *specific* neural, cognitive, or computational system to apply them to.

Similar issues arise for thinking about the operation being performed by these changes in synaptic weight. It is memory storage, but should this storage be thought of as storage of the memory, or storage of the index for information distributed elsewhere? Synaptic structures in the dentate gyrus may be interpreted as storing representations about the past for later retrieval or providing an index to cortically stored engrams so that they can be activated in instances of memory retrieval (see Borzello et al., 2023). Experimental contexts generally permit a number of disparate ascriptions for what is represented and how (Mace, 2025a). It is unclear how metabolic considerations can factor into adjudications about the operations of any particular structure, in which case, it is unclear whether metabolic processes should be considered.

The problem becomes even more pressing as the cognitive models under consideration scale up to the systems level and debates about cognitive ontology. Given Haueis and Colaço's remarks about cognitive effort, it is clear that they intend to have such reach. It is, however, unclear whether metabolic considerations can get a substantive grip in debates at this level of remove from neural implementation. Models of episodic memory, for example, are often divided by which feature they consider primary: the storage and/or recreation of rich, experientially detailed representations of past events, or the storage and reactivation of discrete, event-specific representations. As part of his general approach to cognitive ontology, Khalidi (2023) argues that these should be collapsed into a single cognitive kind. Others (Robins, 2024) argue for splitting them into two distinct kinds, episodic simulation and event memory, respectively.

It is difficult to see how metabolic considerations get much of an evaluative grip here. Neither denies that the underlying neural mechanisms are relevant for cognitive individuation. They instead

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

reflect the orientation of much recent work in cognitive ontology that resists aligning cognitive taxonomy too closely with neural details (Krickel, 2024). The emphasis is instead on the complex causal properties of the overall cognitive system, and the broader environmental and sociocultural context in which it operates.

In debates about specific cognitive models in neuroscience, it thus remains unclear how strongly metabolic considerations can serve as constraints and how decisively metabolic considerations can settle debates. While metabolic considerations may improve individual models, whether those considerations can come to bear on determining which higher level properties are relevant for some cognitive process like memory storage remains to be seen. Further study of this topic would benefit from discussions of the relationship between mechanistic and non-mechanistic cognitive models and the individuation of cognitive capacities, states, and processes.

Financial support. This work was supported in part by the National Science Foundation (grant no. 2422939) and a UC Davis Small Grant in Aid of Research.

Competing interests. The authors declare no competing interests.

References

Borzello, M., Ramirez, S., Treves, A., Lee, I., Scharfman, H., Stark, C., Knierim, J., & Rangel, L. (2023). Assessments of dentate gyrus function: Discoveries and debates. *Nature Reviews Neuroscience*, *24*, 502–517. https://doi.org/10.1038/s41583-023-00710-z

Cariani, P., & Baker, J. (2022). Time is of the essence: Neural codes, synchronies, oscillations, architectures. *Frontiers in Computational Neuroscience*, *16*, 898829. https://doi.org/10.3389/fncom.2022.898829

Coelho Mollo, D. (2017). Functional individuation, mechanistic implementation: The proper way of seeing the mechanistic view of concrete computation. *Synthese*, *195*(8), 3477–3497. https://doi.org/10.1007/s11229-017-1380-5

Drayson, Z. (2025). Defending the medium-independence of computation. *Mind and Language*, *40*(4), 458–467. https://doi.org/10.1111/mila.12536

Becket, E. R., & Hayden, B. (2021). The population doctrine in cognitive neuroscience. *Neuron*, *109*, 3055–3068. https://doi.org/10.1016/j.neuron.2021.07.011

Haimovici, S. (2013). A problem for the mechanistic account of computation. *Journal of Cognitive Science*, *14*(2), 151–181. https://doi.org/10.17791/jcs.2013.14.2.151

Haugeland, J. (1985). *Artificial intelligence: The very idea*. MIT Press.

Khalidi, M. A. (2023). *Cognitive ontology: Taxonomic practices in the mind-brain sciences*. Cambridge University Press.

Krickel, B. (2024). The new mechanistic approach and cognitive ontology—or: What role do (neural) mechanisms play in cognitive ontology? *Minds and Machines*, *34*(3), 1–19. https://doi.org/10.1007/s11023-024-09679-9

Mace, C. (2025a). On a new content indeterminacy problem in neuroscience. *Philosophical Psychology*, 1–18. https://doi.org/10.1080/09515089.2025.2512043

Mace, C. (2025b). The vehicle indeterminacy problem. [Preprint] URL: https://philsci-archive.pitt.edu/id/eprint/26416 (accessed 2026-04-07).

Piccinini, G. (2015). *Physical computation: A mechanistic account*. Oxford University Press UK.

Robins, S. (2024). Kinding memory: Commentary on Muhammad Ali Khalidi's "cognitive ontology". *Mind & Language*, *39*, 109–115. https://doi.org/10.1111/mila.12477

Williams, D. J. (2025). Two senses of medium independence. *Mind and Language*, *40*(4), 437–447. https://doi.org/10.1111/mila.12542

Metabolic constraints and the limits of cognitive modularity

David Moreau[a,b]* and Kristina Wiebels[a,b]

[a]School of Psychology, University of Auckland, New Zealand and [b]Centre for Brain Research, University of Auckland, New Zealand
d.moreau@auckland.ac.nz
k.wiebels@auckland.ac.nz

*Corresponding author.

doi:10.1017/S0140525X26106062, e470

Abstract

Building on Haueis and Colaço, we argue that metabolic considerations should also constrain cognitive architecture. Globally shared and dynamic resources make strong modularity difficult to justify. Rather than fixed features, modules may reflect local optima within a metabolically constrained landscape – stable under typical conditions but liable to break down under perturbation. Metabolic perturbations thus provide diagnostic tools for architectural assumptions.

Haueis and Colaço argue that metabolic considerations should play a central role in cognitive modeling, distinguishing between evaluative uses of metabolism, which function as explanatory constraints on models, and generative uses, which help formulate new models of cognitive capacities and neural systems. In this commentary, we extend their argument to a theoretical commitment that is often implicit in cognitive modeling but rarely scrutinized through a metabolic lens: the assumption that cognitive capacities can be cleanly partitioned into relatively autonomous components (see Barrett & Kurzban, 2006, for a review). Strong modular views have been most clearly articulated in evolutionary psychology and classical modularity theory (Barkow, Cosmides, & Tooby, 1992; Fodor, 1983; Pinker, 1997), but also appear in more implicit form in cognitive neuropsychology, where modular interpretations are often supported by dissociation logic (Caramazza, 1986; Coltheart, 2011). We argue that once metabolism is treated as a globally shared and dynamically allocated constraint, architectures that assume strong functional independence between cognitive subsystems become difficult to justify. Metabolic considerations, therefore, do not merely constrain which computations are feasible; they also place limits on which architectural descriptions remain stable across contexts.

The evaluative framework proposed by Haueis and Colaço emphasizes that metabolism constrains the computational operations biological systems can implement. An important implication concerns the degree of independence cognitive subsystems can plausibly exhibit. Many architectures assume that distinct functions (e.g., attention, memory, and decision-making) can be modeled as relatively autonomous mechanisms, justified on computational or behavioral grounds such as dissociations between tasks or selective impairments. However, if neural systems

operate under shared energetic constraints, functional autonomy becomes metabolically suspect.

The issue is not that shared resources eliminate modularity – many biological and engineered systems maintain functional specialization despite common resource pools. Rather, it pertains to whether cognitive subsystems can remain operationally independent under fluctuating energetic demands (Attwell & Laughlin, 2001). Energy in the brain is neither locally unlimited nor strictly compartmentalized: even when global metabolic expenditure remains relatively stable across tasks, local increases in energy demand must be balanced by redistributions across regions or processes (Lennie, 2003; Raichle & Gusnard, 2002; Raichle & Mintun, 2006). Although mechanisms such as neurovascular coupling allow for localized regulation of metabolic supply, they operate within a globally constrained system in which increases in demand must ultimately be balanced against overall energetic limits (Friston, 2010; Logothetis, 2008; Raichle, 2015). Under such conditions, sharp and stable boundaries between cognitive components are unlikely to hold across all contexts, and functional decompositions may represent approximate descriptions that hold under stable energetic conditions but degrade when metabolic constraints are perturbed (Moreau, 2025).

Empirical observations from metabolic perturbations illustrate this point. Acute energetic manipulations – such as fasting, fatigue, or sustained cognitive effort – tend to produce broad, graded changes across multiple cognitive domains rather than isolated deficits confined to a single function (Bamberg & Moreau, 2025; Chee & Chuah, 2008; van der Linden, Frese, & Sonnentag, 2003). Crucially, this is not consistent with the simple domain-general "resource models" that Haueis and Colaço rightly critique for failing to capture the complexity of metabolic regulation. Rather, the implication is architectural: when metabolic resources become scarce or must be redistributed, multiple processes are affected simultaneously, even when they rely on partially distinct neural circuits, suggesting the system is less cleanly partitioned than functional descriptions imply (Anderson, 2010; Sporns & Betzel, 2016).

One way to reconcile this with the widespread success of modular descriptions is to treat modular structures as local optima within a high-dimensional constraint space defined by task demands, neural organization, and metabolic efficiency (Moreau & Wiebels, 2022). Under stable energetic conditions, these optima may produce the appearance of well-defined modules, even though the underlying system remains metabolically and dynamically coupled. Perturbations in energetic availability – whether due to physiological states, prolonged cognitive effort, or developmental and aging processes – may shift the system away from these equilibria, revealing underlying couplings that modular descriptions obscure. From this perspective, metabolism does not merely constrain the implementation of computations; it shapes the landscape of possible cognitive architectures, and, like cognitive interventions, can be leveraged to formally test them (Moreau, 2022). Some architectural configurations may be computationally elegant yet metabolically unsustainable, whereas others may endure precisely because they satisfy energetic constraints.

These considerations extend the target article's framework in two ways. First, metabolic plausibility should be applied not only to the operations posited by cognitive models but also to the architectural assumptions those models rely upon: if a model presupposes strong independence between subsystems, that independence should be assessed under realistic energetic conditions. Second, and more concretely, metabolic perturbations provide a diagnostic tool for testing architectural claims: if two processes are truly modular, transient energetic stress should selectively impair the more metabolically demanding one; if both degrade in tandem, this reveals coupling that contradicts modular assumptions. Leveraging how cognitive performance changes under energetic stress, researchers can thus assess whether proposed decompositions reflect genuine architectural separations or merely stable operating points within a coupled system.

Acknowledgments. None.

Financial support. None.

Competing interests. None.

References

Anderson, M. L. (2010). Neural reuse: A fundamental organizational principle of the brain. *Behavioral and Brain Sciences*, *33*(4), 245–313.

Attwell, D., & Laughlin, S. B. (2001). An energy budget for signaling in the grey matter of the brain. *Journal of Cerebral Blood Flow and Metabolism*, *21*(10), 1133–1145.

Bamberg, C., & Moreau, D. (2025). Acute effects of fasting on cognitive performance: A systematic review and meta-analysis. *Psychological Bulletin*, *151*, 1147–1169.

Barkow, J. H., Cosmides, L., & Tooby, J. (Eds.). (1992). *The adapted mind: Evolutionary psychology and the generation of culture*. Oxford University Press.

Barrett, H. C., & Kurzban, R. (2006). Modularity in cognition: Framing the debate. *Psychological Review*, *113*(3), 628–647.

Caramazza, A. (1986). On drawing inferences about the structure of normal cognitive systems from the analysis of patterns of impaired performance: The case for single-patient studies. *Brain and Cognition*, *5*(1), 41–66.

Chee, M. W., & Chuah, L. Y. (2008). Functional neuroimaging insights into how sleep and sleep deprivation affect memory and cognition. *Current Opinion in Neurology*, *21*(4), 417–423.

Coltheart, M. (2011). Methods for modular modelling: Additive factors and cognitive neuropsychology. *Cognitive Neuropsychology*, *28*(3–4), 224–240.

Fodor, J. A. (1983). *The modularity of mind*. The MIT Press.

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, *11*(2), 127–138.

Lennie P. (2003). The cost of cortical computation. *Current Biology*, *13*(6), 493–497.

Logothetis, N. K. (2008). What we can do and what we cannot do with fMRI. *Nature*, *453*(7197), 869–878.

Moreau, D. (2022). How malleable are cognitive abilities? A critical perspective on popular brief interventions. *American Psychologist*, *77*(3), 409–423.

Moreau, D. (2025). Temporal stability in measurements matters only for stable constructs. *Nature Reviews Psychology*, *4*, 361.

Moreau, D., & Wiebels, K. (2022). Psychological constructs as local optima. *Nature Reviews Psychology*, *1*, 188–189.

Pinker, S. (1997). *How the Mind Works*. W. W. Norton.

Raichle M. E. (2015). The brain's default mode network. *Annual Review of Neuroscience*, *38*, 433–447.

Raichle, M. E., & Gusnard, D. A. (2002). Appraising the brain's energy budget. *Proceedings of the National Academy of Sciences*, *99*(16), 10237–10239.

Raichle, M. E., & Mintun, M. A. (2006). Brain work and brain imaging. *Annual Review of Neuroscience*, *29*, 449–476.

Sporns, O., & Betzel, R. F. (2016). Modular brain networks. *Annual Review of Psychology*, *67*, 613–640.

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

van der Linden, D., Frese, M., & Sonnentag, S. (2003). The impact of mental fatigue on exploration in a complex computer task: Rigidity and loss of systematic strategies. *Human Factors*, *45*(3), 483–494.

Having the landscape in the background: What can metabolic constraints explain about memory?

Jonathan Najenson*

Ruhr-Universität Bochum, Germany
jonathan.najenson@gmail.com

*Corresponding author.

doi:10.1017/S0140525X26106360, e471

Abstract

Haueis and Colaço claim that metabolic considerations function as explanatory constraints. I argue the explanatory contribution of metabolic constraints for synaptic models of memory is not well delineated. While metabolic constraints determine which synaptic architectures are biologically possible, they do not structure the outcome of synaptic consolidation. Metabolic constraints are explanatory since they account for the background conditions of memory formation.

Haueis and Colaço show that metabolic considerations function as explanatory constraints on cognitive models, demonstrating that ignoring them comes at a cost for cognitive modeling. But what is it that makes them explanatorily relevant, and what exactly do they explain? As we shall see, metabolic constraints on synaptic memory traces are explanatory since they identify dependency relations that account for the background conditions of memory formation.

Haueis and Colaço's core evaluative case – the comparison of linear cascade models (Benna & Fusi, 2016) with cyclical models incorporating non-equilibrium steady states (Karbowski, 2019) – is presented as demonstrating that metabolic considerations play "an evidential and constraining role" for synaptic memory models. Showing that a class of models is metabolically impossible is indeed an explanatory achievement but not one that bears directly on the phenomenon of memory formation itself. What Karbowski's model shows is which architectures are biologically unrealistic, thereby delimiting the space of metabolically feasible synaptic models. A synapse at thermodynamic equilibrium is a dead synapse. Models that presuppose otherwise were never live options. What needs clarifying, at this stage, is what metabolic constraints are explanatory *for*. Constraints of this kind are commonly understood to account for which outcomes are possible and which are impossible, not for why a system produces one outcome rather than another (Lange, 2016). These are genuinely different explananda. When metabolic constraints bear on synaptic memory, they are informative in a preliminary, eliminative sense. They tell us which models to bracket before the explanatory work begins. The evidential function of metabolic considerations here is thus to provide *evidence about background conditions*. Such evidence concerns the variables one holds fixed when evaluating models. The deficiency identified in linear cascade models is that these models fall outside the metabolically acceptable bounds that delineate the background conditions of memory formation.

Explanatory constraints are taken by Haueis and Colaço to do more than merely delimit. They argue that metabolic costs "guide and structure the outcome of synaptic consolidation" towards longer memory traces. Haueis and Colaço adopt this language from Ross (2023). The structuring function of constraints, according to Ross, usually operates through a form of channeling: vascular pathways spatially guide blood flow to specific regions, and molecular pathways direct a substrate through a specific sequence of enzymatic conversions. These are cases where the architecture of the constraining system determines the potential states the explanatory target will take on. In contrast, the synaptic architecture Karbowski's model describes does not *channel* memory traces toward longer durations (where the relevant notion of duration is decay rate rather than persistence of a consolidated trace). The duration of a memory trace depends on the stimulation the synapse receives. That longer memories are relatively cheaper does not mean shorter ones are impossible. A synapse that receives weak and brief stimulation will produce a short-lived trace despite the architecture favoring longer traces. What is more, the sense in which the architecture favors longer memories is that you get more *bang for your buck*. In certain parameter regimes of the model, the cost landscape favors longer traces per unit of energy spent (Karbowski, 2019, p. 1484). Unlike a vascular pathway, this is an efficiency gradient, not a dynamical bias of the system. Even so, an energetically favorable outcome is explanatorily irrelevant for a given phenomenon if it does not track the actual conditions under which that phenomenon obtains.

What dependency relation is there, then, between the structure of synaptic architecture and memory duration? At the moment of encoding, more energy would result in longer potential traces. Yet after encoding, memory duration decouples from energy expenditure (Karbowski, 2019, pp. 1482–3). This means duration *need not depend* on energy expenditure. The cost of memories does not, *pace* Haueis and Colaço, constrain the outcome of "synaptic consolidation." These dynamics are quite different from how the structuring function of constraints is supposed to work, where the constraining variable channels the process toward specific outcomes.

If metabolic constraints do not guide, where do they lead us? They delineate the background conditions of the target phenomenon (Craver, 2007). Haueis and Colaço point in this direction when they characterize metabolic resources as causally unspecific enabling conditions in the case of mental effort. If metabolic constraints on synaptic memory are similarly enabling conditions, then what determines explanatory relevance is the *invariance* of a causal relationship across a domain-specific range of such conditions (Woodward, 2010). Previously, I argued that synaptic models of memory draw their explanatory power from the fact that their variables are causally invariant across a range of background conditions (Najenson, 2024). For instance, synaptic mechanisms can be demonstrated to support memory across a wide range of behavioral tasks (Morris et al., 1986; Nabavi et al., 2014; Rioult-Pedotti, Friedman, & Donoghue, 2000). Interestingly, a non-equilibrium steady state is a highly stable but non-specific condition. It holds across all living synapses. It is an open empirical question, but memory traces may be largely invariant to metabolic

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

fluctuations within a fixed range of non-equilibrium steady states. So, coming back to where I started, what makes metabolic constraints on synaptic models of memory explanatorily relevant is that they *explain which synaptic models are biologically possible*. Their explanatory relevance lies in determining the background on which we evaluate how specific and invariant synaptic mechanisms are.

Financial support. This research was supported by a Minerva Stiftung fellowship.

Competing interests. None.

References

Benna, M. K., & Fusi, S. (2016). Computational principles of synaptic memory consolidation. *Nature Neuroscience*, *19*(12), 1697–1706. https://doi.org/10.1038/nn.4401

Craver, C. F. (2007). *Explaining the brain: Mechanisms and the mosaic unity of neuroscience*. Oxford University Press.

Karbowski, J. (2019). Metabolic constraints on synaptic learning and memory. *Journal of Neurophysiology*, *122*(4), 1473–1490. https://doi.org/10.1152/jn.00092.2019

Lange, M. (2016). *Because without cause: Non-causal explanations in science and mathematics*. Oxford University Press.

Morris, R. G., Anderson, E., Lynch, G. S., & Baudry, M. (1986). Selective impairment of learning and blockade of long-term potentiation by an N-methyl-D-aspartate receptor antagonist, AP5. *Nature*, *319*(6056), 774–776. https://doi.org/10.1038/319774a0

Nabavi, S., Fox, R., Proulx, C. D., Lin, J. Y., Tsien, R. Y., & Malinow, R. (2014). Engineering a memory with LTD and LTP. *Nature*, *511*(7509), 348–352. https://doi.org/10.1038/nature13294

Najenson, J. (2024). LTP revisited: Reconsidering the explanatory power of synaptic efficacy. *Review of Philosophy and Psychology*, *15*(4), 1281–1306. https://doi.org/10.1007/s13164-023-00694-w

Rioult-Pedotti, M. S., Friedman, D., & Donoghue, J. P. (2000). Learning-induced LTP in neocortex. *Science*, *290*(5491), 533–536. https://doi.org/10.1126/science.290.5491.533

Ross, L. N. (2023). The explanatory nature of constraints: Law-based, mathematical, and causal. *Synthese*, *202*(2), 56. https://doi.org/10.1007/s11229-023-04281-5

Woodward, J. (2010). Causation in biology: Stability, specificity, and the choice of levels of explanation. *Biology and Philosophy*, *25*(3), 287–318. https://doi.org/10.1007/s10539-010-9200-z

Metabolism beyond the neuron: system-level constraints on cognitive modelling

Olivier Penacchio[a,b*†] and Ana Clemente[c,†]

[a]Computer Vision Center, Bellaterra, Spain; [b]Computer Science Department, Universitat Autònoma de Barcelona, Bellaterra, Spain and [c]Department of Cognitive Neuropsychology, Max Planck Institute for Empirical Aesthetics, Frankfurt, Germany
penacchio@cvc.uab.cat
ana.clemente@ae.mpg.de

*Corresponding author.

†Equal contribution

doi:10.1017/S0140525X26106232, e472

Abstract

Survival depends on efficient energy management, and neural systems are highly metabolically demanding biological structures. We fully endorse Haueis and Colaço's call to integrate metabolism as a fundamental constraint on cognitive modelling. However, by focusing primarily on the neuronal level, their analysis underplays system-level approaches that have already incorporated metabolic constraints into successful and predictive accounts of cognition.

We concur with Haueis and Colaço that metabolism should be treated as a central constraint on cognitive modelling. We also support their constructive proposal that metabolic considerations should delimit what is biologically feasible (evaluative role) and guide the development of new models (generative role), thereby establishing an iterative process in which metabolism shapes and constrains the space of possible cognitive accounts. However, their review primarily emphasises the cellular and synaptic scale and pays comparatively little attention to system-level principles that have given rise to some of the most influential theoretical frameworks in cognitive neuroscience. These include sparse coding, predictive coding, and, more broadly, anticipatory schemes such as active inference – all global computational strategies that reflect and impose constraints on local neuronal activity.

Sparse coding builds on the efficient coding ideas of Attneave and Barlow by formalising a trade-off between information and metabolic cost in the encoding of relevant stimuli (Attneave, 1954; Barlow, 1961). This principle enforces that only a limited number of neurons are active at any given time (Olshausen & Field, 1996). Sparse coding is thus framed as a generative, iterative modelling scheme in which representations are successively evaluated at each stage using metrics that penalise excessive activity – serving as a proxy for the metabolic cost of spiking – while preserving representational accuracy (Olshausen & Field, 1996). Beyond its success in accounting for the receptive fields of sensory neurons in multiple modalities (Lewicki, 2002; Olshausen & Field, 1996; Perez-Orive et al., 2002), sparse coding has been observed in response to ecologically relevant stimuli in sensory (Machens et al., 2005; Vinje & Gallant, 2000), memory, and associative systems, suggesting a domain-general organisational rule (Rolls & Treves, 2011; Tallman, Steinmetz, & Wixted, 2025; Waydo et al., 2006). Sparse coding thus provides a clear illustration of how global constraints induce specific local patterns of neural activity, with direct implications for ATP consumption. More broadly, the empirical support for sparse coding indicates that system-level metabolic constraints directly shape cognitive representations, here giving rise to parsimonious selective codes that balance explanatory power against metabolic cost.

Another influential systems-level framework with direct metabolic implications is predictive coding, which is considered in the target article primarily from a local perspective. Predictive coding draws on Helmholtz's (von Helmholtz, 1866) view of perception as inference – where sensory input is interpreted against prior expectations – and on engineering approaches to efficient communication such as linear predictive coding (Makhoul, 1975), in which signals are predicted from past inputs and only residual errors, the difference between prediction and input, are

transmitted. Although not originally formulated in metabolic terms, its core mechanism – the local, distributed computation of prediction errors – minimises redundant signalling and reduces the need for sustained or long-range neural activity, thereby potentially lowering metabolic costs within a global computing scheme (Friston, 2010; Rao & Ballard, 1999), and suggesting a canonical organisational principle (Keller & Mrsic-Flogel, 2018). Supported by a substantial – albeit partial – body of empirical evidence (Bastos et al., 2012; Caucheteux et al., 2023) and central to many modelling efforts (Hodson, Mehta, & Smith, 2024), predictive coding has become a guiding principle in contemporary models of perception and cognition, offering a framework in which improved representational efficiency is achieved alongside reduced metabolic demand.

Building on predictive coding, active inference, a prominent framework in modelling cognition (Hodson, Mehta, & Smith, 2024), combines the advantage of prediction – which reduces unnecessary computations by encoding only deviations from expected sensory input – with a frugal inference engine that performs tractable, locally computable belief updates. Indeed, at its core, the free energy principle employs variational inference, replacing exact but intractable inference with a simplified optimisation based on principled approximations (Friston, 2010). Crucially, this notion of tractability – formulated as a system-level computational requirement and embedded within the architecture of active inference models – translates into metabolic constraints on neural systems, ensuring that inference remains feasible under limited energetic resources. In this sense, the free energy principle constrains the space of viable cognitive architectures: by relying on variational inference, it restricts cognition to tractable approximations, thereby placing principled bounds on cognitive complexity under metabolic constraints.

The examples above illustrate how global principles impose structured bounds on local neural activity and, consequently, on metabolic expenditure. These system-level restrictions complement the cellular and synaptic factors emphasised by Haueis and Colaço, together delineating a hierarchical organisation of metabolic requirements governing the viability of cognitive systems. These considerations naturally extend to even higher-level ecological constraints on cognitive modelling, which encompass and contextualise metabolic ones. In particular, they point to the necessity of grounding cognitive models in embodiment: cognition must be understood, and modelled, in relation to the sensory, motor, and environmental contingencies under which it operates. This perspective is essential both for explaining biological cognition and for guiding the development of more efficient artificial systems. In this respect, proposals such as the embodied Turing test (Zador et al., 2023) emphasise that models of intelligence should be evaluated not only in terms of internal performance, but in terms of their ability to operate under the physical, ecological, and energetic constraints that characterise real-world behaviour. In this hierarchy, neuronal metabolism should be understood as one crucial constraint among many that jointly determine viable cognitive architectures.

Financial support. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Competing interests. None.

References

Attneave, F. (1954). Some informational aspects of visual perception [Article]. *Psychological Review*, *61*(3), 183–193. https://doi.org/10.1037/h0054663.

Barlow, H. B. (1961). *Possible principles underlying the transformations of sensory messages*. A Symposium. M. I. T. Press, Cambridge 39, Mass., New York.

Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., & Friston, K. J. (2012). Canonical microcircuits for predictive coding. *Neuron*, *76*(4), 695–711. https://doi.org/10.1016/j.neuron.2012.10.038.

Caucheteux, C., Gramfort, A., & King, J.-R. (2023). Evidence of a predictive coding hierarchy in the human brain listening to speech. *Nature Human Behaviour*, *7*(3), 430–441. https://doi.org/10.1038/s41562-022-01516-2.

Friston, K. J. (2010). The free-energy principle: A unified brain theory? [Review]. *Nature Reviews Neuroscience*, *11*(2), 127–138. https://doi.org/10.1038/nrn2787.

Hodson, R., Mehta, M., & Smith, R. (2024). The empirical status of predictive coding and active inference. *Neuroscience and Biobehavioral Reviews*, *157*, 105473. https://doi.org/10.1016/j.neubiorev.2023.105473.

Keller, G. B., & Mrsic-Flogel, T. D. (2018). Predictive processing: A canonical cortical computation. *Neuron*, *100*(2), 424–435. https://doi.org/https://doi.org/10.1016/j.neuron.2018.10.003.

Lewicki, M. S. (2002). Efficient coding of natural sounds [Article]. *Nature Neuroscience*, *5*(4), 356–363. https://doi.org/10.1038/nn831.

Machens, C. K., Gollisch, T., Kolesnikova, O., & Herz, A. V. M. (2005). Testing the efficiency of sensory coding with optimal stimulus ensembles [Article]. *Neuron*, *47*(3), 447–456. https://doi.org/10.1016/j.neuron.2005.06.015.

Makhoul, J. (1975). Linear prediction: A tutorial review. *Proceedings of the IEEE*, *63*(4), 561–580. https://doi.org/10.1109/PROC.1975.9792.

Olshausen, B. A., & Field, D. J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images [Article]. *Nature*, *381*(6583), 607–609. https://doi.org/10.1038/381607a0.

Perez-Orive, J., Mazor, O., Turner, G. C., Cassenaer, S., Wilson, R. I., & Laurent, G. (2002). Oscillations and sparsening of odor representations in the mushroom body [Article]. *Science*, *297*(5580), 359–365. https://doi.org/10.1126/science.1070502.

Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects [Article]. *Nature Neuroscience*, *2*(1), 79–87. https://doi.org/10.1038/4580.

Rolls, E. T., & Treves, A. (2011). The neuronal encoding of information in the brain. *Progress in Neurobiology*, *95*(3), 448–490. https://doi.org/10.1016/j.pneurobio.2011.08.002.

Tallman, C. W., Steinmetz, P. N., & Wixted, J. T. (2025). Neuronal allocation and sparse coding of episodic memories in the human hippocampus. *Scientific Reports*, *15*(1), 40430. https://doi.org/10.1038/s41598-025-21967-7.

Vinje, W. E., & Gallant, J. L. (2000). Sparse coding and decorrelation in primary visual cortex during natural vision [Article]. *Science*, *287*(5456), 1273–1276. https://doi.org/10.1126/science.287.5456.1273.

Von Helmholtz, H. (1866). *Concerning the perceptions in general* (Vol. *3*). Dover.

Waydo, S., Kraskov, A., Quian Quiroga, R., Fried, I., & Koch, C. (2006). Sparse representation in the human medial temporal lobe. *The Journal of Neuroscience*, *26*(40), 10232–10234. https://doi.org/10.1523/jneurosci.2101-06.2006.

Zador, A., Escola, S., Richards, B., Ölveczky, B., Bengio, Y., Boahen, K., Botvinick, M., Chklovskii, D., Churchland, A., Clopath, C., DiCarlo, J., Ganguli, S., Hawkins, J., Körding, K., Koulakov, A., LeCun, Y., Lillicrap, T., Marblestone, A., Olshausen, B., Pouget, A., Savin, C., Sejnowski, T., Simoncelli, E., Solla, S., Sussillo, D., Tolias, A.S., & Tsao, D. (2023). Catalyzing next-generation artificial intelligence through neuroAI. *Nature Communications*, *14*(1), 7, Article 1597. https://doi.org/10.1038/s41467-023-37180-x.

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

Metabolic considerations point to specialised and analogue computation in the brain

Christian Percy*

University of Warwick, UK
chris@cspres.co.uk

*Corresponding author.

doi:10.1017/S0140525X26106153, e473

Abstract

Metabolic considerations are argued to drive cognitive models. I explore the related high-level hypothesis that the brain favours the most anatomically specialised system that still meets demands for computational flexibility in a given task. The hypothesis points to specific claims across a hierarchy of specialisation, from inter-neuronal connection structure to neuronal cell structure, through to non-neuronal cells and analogue mechanisms.

Neural signalling is metabolically expensive (Attwell & Laughlin, 2001), motivating hypotheses that metabolic considerations drive cognitive structures. Given the tendency for specialisation to track efficiency across diverse domains (e.g., neuroscience: Ito, 2006; algorithms: Chaitin, 1977; biology: Vogel, 1998), I propose a related high-level hypothesis: the brain favours the most anatomically specialised system that still meets demands for computational flexibility in a given task.

This hypothesis helps posit specific cognitive capacities or target systems that merit detailed metabolic modelling, connecting to guideline G1 in Haueis and Colaço. Taken seriously, this hypothesis motivates attention towards emergent research areas that could enrich connectionist models with metabolically and computationally relevant roles for non-neuronal cells and analogue mechanisms.

The high-level hypothesis

More anatomically specialised architectures are defined as those whose physical structures map to a narrower range of system-relevant computational tasks. For instance, the autoassociation capability of the hippocampus is more general than retinal cell edge detection but less computationally general than predictive processing in the neocortex. Detailed analysis can leverage methods to quantify dimensions of "computational generality," e.g., programmability (Zenil, 2013), VC dimension (Vapnik, 1998), and Rademacher complexity (Bartlett & Mendelson, 2002).

"Metabolic efficiency" can be operationalised by adapting "information-energy efficiency": average ATP count to process environmentally representative information in one functional input-output sweep on a system-relevant task (e.g., 1x image recognition, 1x learning update).

When two systems have been available in evolutionary terms and are both capable of a given computation, the hypothesis claims the brain will favour the more specialised system due to its greater efficiency. In practice, the brain and its evolutionary environment are highly complex, path-dependent systems (Jacob, 1977), such that other factors might dominate. In general, the hypothesis should be stronger under greater selection pressure – phylogenetically older systems, more frequent tasks, etc. – pointing to possible cases to explore first.

Specific cases in inter-neuronal networks to explore

The brain has evolutionary access to a highly general-purpose architecture and yet exploits it rarely. Fully connected single hidden-layer neural networks can approximate any function (Hornik, Stinchcombe, & White, 1989). Networks with identical neurones and uniform connectivity structure represent very low levels of specialisation. However, even 20%-30% connectivity qualifies as densely connected local circuits in the brain (Perin, Berger, & Markram, 2011; Song et al., 2005), and fully connected cliques occur only at very small scales (Reimann et al., 2017). Instead, the most computationally general brain structure, the neocortex, exploits layers and reciprocal dynamics in highly structured networks (Mountcastle, 1997). The hypothesis suggests that metabolic efficiency analysis would locate this level of structural fine-tuning in an approximate Goldilocks zone, given the computational generality of target tasks.

The hierarchy of anatomical specialism extends beyond structured connections between neurones to specialised neuronal cells. The cerebellum implements narrow adaptive filtering tasks with efficient specialised microcircuitry via Purkinje neurones and climbing fibre error signals (Albus, 1971; Kawato et al., 2021; Marr, 1969). Standard predictive coding in cortical microcolumns (Bastos et al., 2012) could implement these tasks in principle, but this hypothesis suggests the cerebellum is favoured (in part) for efficiency reasons, e.g., multiple microcolumns might be required for each Purkinje cell.

Cases beyond inter-neuronal synaptic connectivity?

Neuronal signalling is metabolically expensive, so offloading computational primitives to non-neuronal cells (where viable) is likely to be even more favourable under this hypothesis than neuronal specialism.

Computationally narrow tasks in phylogenetically older systems provide candidate cases to examine. For instance, Müller glia in the retina play roles in signal processing, guiding light to photoreceptors and regulating neurotransmitter dynamics (Newman & Reichenbach, 1996; Reichenbach & Bringmann, 2013), while Bergmann glia in the cerebellum interact with synaptic timing and plasticity (Bellamy, 2006; De Zeeuw & Hoogland, 2015).

A metabolic comparison of neuronal and non-neuronal mechanisms for implementing these computational primitives should favour the latter. In cases where a non-neuronal efficiency gain has not been exploited, we expect a countervailing reason: perhaps the efficiency gain is too modest relative to robustness priorities (Goldman et al., 2001), the task too infrequent for significant selection pressure, the required system transition too large to have occurred (yet) by chance mutation, or system reuse more efficient than specialisation overall (e.g., once construction and system interaction costs are considered; see Anderson, 2010).

This hypothesis also motivates attention to emerging, more speculative avenues of research on candidate analogue computational mechanisms, inspired also by the parallels to computing R&D in Haueis and Colaço.

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

Consider the emerging evidence on ephaptic coupling mechanisms in mammal brains, given how modest field strengths shape activation phases (e.g., 0.2 to 2 mV/mm) (Anastassiou et al., 2011; Anastassiou & Koch, 2015; Jefferys, 1995) and extracellular fields generated by neurones shape both activations (Pinotsis & Miller, 2023) and working memory (Pinotsis & Miller, 2022). Since EM fields are already produced by neurones, it is potentially highly efficient to leverage them computationally. Industrial R&D also leverages fields as energy-efficient analogue mechanisms in their own right – e.g., matrix multiplication (Bai et al., 2023), Fourier transforms (Heinz, Artman, & Lee, 1970), and combinatorial optimisation (e.g., Hamerly et al., 2019; US patent US11698945B2) – providing potential (if highly speculative) inspiration for applications to investigate in the brain.

Components have lower metabolic costs than their host systems, motivating an examination of intra-neuronal mechanisms. A classic example is dendritic computation enhancing neurone-level capabilities (London & Häusser, 2005; Ujfalussy et al., 2015). Turning to even smaller subsystems, a speculative example is leveraging extremely rapid path integral mechanics for predictive processing in cases where Hodgkin–Huxley neurones are too slow (Wiest & Puniani, 2025). Much more work is needed to assess any role that quantum phenomena might play in cognition, but the R&D parallel provides further motivation to investigate the possibility, given the energy efficiency argument for quantum computers (Meier & Yamasaki, 2025; Nielsen & Chuang, 2010; Preskill, 2018; US patent US11238000B2).

As appreciation of evolutionary complexity improves, it is plausible that efficiency gains motivating industrial R&D might have also been mined by evolution. Metabolic considerations provide a structured framework for investigating these possibilities and for converting intriguing speculations into disciplined science.

Acknowledgements. None.

Financial support. Work partially funded by a research grant from a philanthropic consortium led by David Langer that included Lionheart Ventures, Amphibian Capital, and Winslow Strong.

Competing interests. None.

References

Albus, J. S. (1971). A theory of cerebellar function. *Mathematical Biosciences*, *10*(1–2), 25–61.

Anastassiou, C. A., & Koch, C. (2015). Ephaptic coupling to endogenous electric field activity: Why bother? *Current Opinion in Neurobiology*, *31*, 95–103. https://doi.org/10.1016/j.conb.2014.09.002

Anastassiou, C. A., Perin, R., Markram, H., & Koch, C. (2011). Ephaptic coupling of cortical neurons. *Nature Neuroscience*, *14*(2), 217–223. https://doi.org/10.1038/nn.2727

Anderson, M. L. (2010). Neural reuse: A fundamental organizational principle of the brain. *Behavioral and Brain Sciences*, *33*(4), 245–266. https://doi.org/10.1017/S0140525X10000853

Attwell, D., & Laughlin, S. B. (2001). An energy budget for signaling in the grey matter of the brain. *Journal of Cerebral Blood Flow & Metabolism*, *21*(10), 1133–1145.

Bai, Y., Xu, X., Tan, M., Sun, Y., Li, Y., Wu, J., Morandotti, R., Mitchell, A., Xu, K., & Moss, D. J. (2023). Photonic multiplexing techniques for neuromorphic computing. *Nanophotonics*, *12*(5), 795–817.

Bartlett, P. L., & Mendelson, S. (2002). Rademacher and Gaussian complexities: Risk bounds and structural results. *Journal of Machine Learning Research*, *3*, 463–482.

Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., & Friston, K. J. (2012). Canonical microcircuits for predictive coding. *Neuron*, *76*(4), 695–711.

Bellamy, T. C. (2006). Interactions between Purkinje neurones and Bergmann glia. *The Cerebellum*, *5*(2), 116–126

Chaitin, G. J. (1977). Algorithmic information theory. *IBM Journal of Research and Development*, *21*(4), 350–359.

De Zeeuw, C. I., & Hoogland, T. M. (2015). Reappraisal of Bergmann glial cells as modulators of cerebellar circuit function. *Frontiers in Cellular Neuroscience*, *9*, Article 246. https://doi.org/10.3389/fncel.2015.00246

Goldman, M. S., Golowasch, J., Marder, E., & Abbott, L. F. (2001). Global structure, robustness, and modulation of neuronal models. *Journal of Neuroscience*, *21*(14), 5229–5238.

Hamerly, R., Inagaki, T., McMahon, P. L., Venturelli, D., Marandi, A., Onodera, T., Ng, E., Langrock, C., Inaba, K., Honjo, T., Enbutsu, K., Umeki, T., Kasahara, R., Utsunomiya, S., Kako, S., Kawarabayashi, K., Byer, R., Fejer, M., Mabuchi, H., … & Yamamoto, Y. (2019). Experimental investigation of performance differences between coherent Ising machines and a quantum annealer. *Science Advances*, *5*(5), eaau0823.

Heinz, R. A., Artman, J. O., & Lee, S. H. (1970). Matrix multiplication by optical methods. *Applied Optics*, *9*(9), 2161–2168.

Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, *2*(5), 359–366. https://doi.org/10.1016/0893-6080(89)90020-8

Ito, M. (2006). Cerebellar circuitry as a neuronal machine. *Progress in Neurobiology*, *78*(3–5), 272–303.

Jacob, F. (1977). Evolution and tinkering. *Science*, *196*(4295), 1161–1166. https://doi.org/10.1126/science.860134

Jefferys, J. G. (1995). Nonsynaptic modulation of neuronal activity in the brain: Electric currents and extracellular ions. *Physiological Reviews*, *75*(4), 689–723.

Kawato, M., Ohmae, S., Hoang, H., & Sanger, T. D. (2021). 50 years since the Marr, Ito, and Albus models of the cerebellum. *Neuroscience*, *462*, 151–174. https://doi.org/10.1016/j.neuroscience.2020.06.019

London, M., & Häusser, M. (2005). Dendritic computation. *Annual Review of Neuroscience*, *28*, 503–532.

Marr, D. (1969). A theory of cerebellar cortex. *The Journal of Physiology*, *202*, 437–470. https://doi.org/10.1113/jphysiol.1969.sp008820

Meier, F., & Yamasaki, H. (2025). Energy-consumption advantage of quantum computation. *PRX Energy*, *4*(2), 023008.

Mountcastle, V. B. (1997). The columnar organization of the neocortex. *Brain*, *120*(4), 701–722.

Newman, E. A., & Reichenbach, A. (1996). The Müller cell: A functional element of the retina. *Trends in Neurosciences*, *19*(8), 307–312. https://doi.org/10.1016/0166-2236(96)10040-0

Nielsen, M. A., & Chuang, I. L. (2010). *Quantum computation and quantum information*. (10th anniversary ed.) Cambridge University Press.

Perin, R., Berger, T. K., & Markram, H. (2011). A synaptic organizing principle for cortical neuronal groups. *Proceedings of the National Academy of Sciences*, *108*(13), 5419–5424. https://doi.org/10.1073/pnas.1016051108

Pinotsis, D. A., & Miller, E. K. (2022). Beyond dimension reduction: Stable electric fields emerge from and allow representational drift. *NeuroImage*, *253*, 119058. https://doi.org/10.1016/j.neuroimage.2022.119058

Pinotsis, D. A., & Miller, E. K. (2023). In vivo ephaptic coupling allows memory network formation. *Cerebral Cortex*, *33*(17), 9877–9895.

Preskill, J. (2018). Quantum computing in the NISQ era and beyond. *Quantum*, *2*, 79. https://doi.org/10.22331/q-2018-08-06-79

Reichenbach, A., & Bringmann, A. (2013). New functions of Müller cells. *Glia*, *61*(5), 651–678. https://doi.org/10.1002/glia.22477

Reimann, M. W., Nolte, M., Scolamiero, M., Turner, K., Perin, R., Chindemi, G., & Markram, H. (2017). Cliques of neurons bound into cavities provide a

missing link between structure and function. *Frontiers in Computational Neuroscience*, *11*, 34.
Song, S., Sjöström, P. J., Reigl, M., Nelson, S., & Chklovskii, D. B. (2005). Highly nonrandom features of synaptic connectivity in local cortical circuits. *PLoS Biology*, *3*(3), e68.
Ujfalussy, B. B., Makara, J. K., Branco, T., & Lengyel, M. (2015). Dendritic nonlinearities are tuned for efficient spike-based computations in cortical circuits. *eLife*, *4*, e10056.
Vapnik, V. N. (1998). *Statistical learning theory*. Wiley.
Vogel, S. (1998). *Cats' paws and catapults: Mechanical worlds of nature and people*. W. W. Norton & Company.
Wiest, M. C., & Puniani, A. S. (2025). Conscious active inference I: A quantum model naturally implements the path integral needed for real-time planning and control. *Computational and Structural Biotechnology Journal*, *30*, 108–121.
Zenil, H. (2013). A behavioural foundation for natural computing and a programmability test. *arXiv preprint arXiv:1303.5887*. https://arxiv.org/abs/1303.5887

Bridging metabolic and computational accounts of efficient information processing

Stephan Pohl*

Philosophy, New York University, USA
stephan.pohl@nyu.edu

*Corresponding author.

doi:10.1017/S0140525X26106220, e474

Abstract

Although abstract computational models in neuroscience omit biochemical detail, metabolic considerations still apply through abstract measures of efficient information processing. Against Haueis and Colaço's claim that such measures are too generic to discriminate among competing models, I argue that coding efficiency and related measures are commonly applied to evaluate and generate specific cognitive models.

While Haueis and Colaço successfully highlight the relevance of metabolic considerations in neuroscientific modeling, they focus, primarily, on intracellular biochemical processes (e.g., costs in terms of number of ATP molecules spent). Most modeling of cognitive phenomena in neuroscience operates at a different level of abstraction, framed in terms of computational processes across populations of neurons. Haueis and Colaço do consider information-theoretic measures of efficiency applicable to such models, but dismiss them as too generic to adjudicate between competing cognitive models. I argue that this dismissal is premature: abstract considerations about the efficiency of information processing are broadly applied to evaluate and generate specific cognitive models in neuroscience. Yet the rationale for expecting efficiency in biological cognitive systems often remains implicit. By foregrounding metabolic considerations, Haueis and Colaço make explicit one reason to expect efficient information processing in such systems.

A neural code is efficient for transmitting information about a stimulus feature if it conveys as much information as possible using as short a message as possible (e.g., with few spikes) (Shannon, 1948). This information-theoretic notion has been applied to computational neuroscience (Borst & Theunissen, 1999). Bialek et al. (1993) operationalized the coding efficiency in terms of the mutual information between a stimulus feature and a neural response, normalized by the entropy of the neural response. The normalization is key: coding efficiency measures what proportion of the neural variability (measured by its entropy) actually carries information about the stimulus. One natural motivation for considering this proportion is metabolic. Maintaining neural variability (e.g., the firing of neurons) is metabolically costly; variability that carries no information about the stimulus is therefore metabolically wasteful. An efficient neural code allows spending limited metabolic resources with the greatest benefit for information transmission.

Let us consider a few cases in which coding efficiency has been used to evaluate and generate specific cognitive models.

Rieke, Warland, and Bialek (1993) showed that receptors in the bullfrog inner ear, transmitting information about the acoustic waveform, and filiform hairs of the cricket, transmitting information about airflow velocity, both achieve coding efficiencies of 50–60%. This establishes that these systems exploit a substantial proportion of their response variability. Rieke, Bodnar, and Bialek (1995) used the same measure to show that neural coding is optimal for natural stimuli: primary auditory afferents in the bullfrog, transmitting information about acoustic waveform, respond with ~80% efficiency to stimuli resembling natural advertisement calls, but only with ~20% efficiency to white noise. This supports the hypothesis that the neural code is specifically organized around the statistics of ecologically relevant stimuli and rules out models in which the code is stimulus-agnostic.

Coding efficiency can also help to resolve questions about the temporal structure of the neural code. In a rate-based code, only the firing rate across a relatively large time window is computationally relevant. In a timing-based code, the spike timing down to a small temporal resolution is relevant. De Ruyter van Steveninck et al. (1997) and Strong et al. (1998) measured how much information H1 motion-sensitive neurons in the fly visual system transmit about visual flow as a function of the temporal resolution of spike counts. Efficiency remains near 50% even at 3 ms resolution, supporting a timing-based rather than rate-based code. Here, the measure is doing exactly the discriminative work Haueis and Colaço suggest it cannot: choosing between two specific, competing hypotheses about neural representation.

Considerations about the efficiency of information processing in cognitive neuroscience extend far beyond the literature on coding efficiency alone. As Pohl et al. (2026) note, standard goodness-of-fit measures for encoding models are closely related to coding efficiency. Most notably, the proportion of variance explained is a variance-based analogue of coding efficiency as defined by Bialek et al. (1993), where the neural variability not explained by the stimulus is measured in terms of the remaining variance of the error of an encoding model rather than the remaining entropy of the neural response given the stimulus. The interpretation of the measure remains largely the same: Since maintaining variability is costly, we can expect that a true model of the computational process leaves little variance unexplained. A model that fails to explain a large proportion of variance is probably missing something about what the neural response is doing.

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

Measures of the goodness-of-fit of encoding models are applied even more broadly across cognitive neuroscience than information-theoretic measures of coding efficiency. Pohl et al. (2026) argue that encoding model performance, under the label "specificity," is one of the four core characteristics of neural representations that structure how evidence for neural representations is gathered and link this explicitly to measures of coding efficiency.

What remains underexplored – and what Haueis and Colaço's framework usefully foregrounds – is *why* we should expect biological cognitive systems to implement efficient information processing in the first place. Metabolic cost is one compelling answer: unexploited variability is energetically expensive, and we may expect evolution to select against it. But this is unlikely to be the only answer. Efficient neural codes may also be faster, affording quicker reaction times (Gautrais & Thorpe, 1998), or they may be less complex and therefore favored by learning or ontogenesis (Grünwald, 2007; Simoncelli & Olshausen, 2001).

Notably, in the line of literature following Bialek et al. (1993) and Rieke, Warland, and Bialek (1993), the benefits of efficient coding are rarely discussed explicitly. Here, efficiency of coding may be seen as a mere theoretical virtue much like mathematical elegance or simplicity. In contrast, the line of literature on efficiency that Haueis and Colaço draw on (Attwell & Laughlin, 2001; Laughlin, 2001) does consider metabolic cost as an explicit motivation for efficiency. There are surprisingly few crossreferences between these lines of literature, which may be an instance of a broader sociological divide between more biology-adjacent and more physics-adjacent groups of neuroscientists (De Schutter, 2008). By drawing our attention to metabolic considerations, Haueis and Colaço help to bridge this gap.

Financial support. There was no funding for this work.

Competing interests. None.

Reference

Attwell, D., & Laughlin, S. B. (2001). An energy budget for signaling in the grey matter of the brain. *Journal of Cerebral Blood Flow & Metabolism*, *21*(10), 1133–1145. https://doi.org/10.1097/00004647-200110000-00001.

Bialek, W., DeWeese, M., Rieke, F., & Warland, D. (1993). Bits and brains: Information flow in the nervous system. *Physica A: Statistical Mechanics and Its Applications*, *200*(1–4), 581–593. https://doi.org/10.1016/0378-4371(93)90563-J.

Borst, A., & Theunissen, F. E. (1999). Information theory and neural coding. *Nature Neuroscience*, *2*(11), 947–957. https://doi.org/10.1038/14731.

de Ruyter van Steveninck, R. R., Lewen, G. D., Strong, S. P., Koberle, R. & Bialek, W. (1997). Reproducibility and variability in neural spike trains. *Science*, *275*(5307), 1805–1808. https://doi.org/10.1126/science.275.5307.1805.

De Schutter, E. (2008). Why are computational neuroscience and systems biology so separate? *PLoS Computational Biology*, *4*(5), e1000078. https://doi.org/10.1371/journal.pcbi.1000078.

Gautrais, J., & Thorpe, S. (1998). Rate coding versus temporal order coding: A theoretical approach. *Biosystems*, *48*(1–3), 57–65. https://doi.org/10.1016/S0303-2647(98)00050-1.

Grünwald, P. D. (2007). *The minimum description length principle*. The MIT Press. https://doi.org/10.7551/mitpress/4643.001.0001.

Laughlin, S. (2001). Energy as a constraint on the coding and processing of sensory information. *Current Opinion in Neurobiology*, *11*(4), 475–480. https://doi.org/10.1016/S0959-4388(00)00237-3.

Pohl, S., Walker, E. Y., Barack, D. L., Lee, J., Denison, R. N., Block, N., Meyniel, F., & Ma, W. J. (2026). Clarifying the conceptual dimensions of representation in neuroscience. *Nature Reviews Neuroscience*, *27*, 357–372. https://doi.org/10.1038/s41583-026-01030-8.

Rieke, F., Bodnar, D. A., & Bialek, W. (1995). Naturalistic stimuli increase the rate and efficiency of information transmission by primary auditory afferents. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *262*(1365), 259–265. https://doi.org/10.1098/rspb.1995.0204.

Rieke, F., Warland, D., & Bialek, W. (1993). Coding efficiency and information rates in sensory neurons. *Europhysics Letters (EPL)*, *22*(2), 151–156. https://doi.org/10.1209/0295-5075/22/2/013.

Shannon, C. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, *27*(3), 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x.

Simoncelli, E. P., & Olshausen, B. A. (2001). Natural image statistics and neural representation. *Annual Review of Neuroscience*, *24*(1), 1193–1216. https://doi.org/10.1146/annurev.neuro.24.1.1193.

Strong, S. P., Koberle, R., de Ruyter van Steveninck, R. R., & Bialek, W. (1998). Entropy and information in neural spike trains. *Physical Review Letters*, *80*(1), 197–200. https://doi.org/10.1103/PhysRevLett.80.197.

Metabolism, medium dependence, and the physicality of information processing

Wiktor Rorot[a]*, Urte Laukaityte[b] and David J. Harrison[c]

[a]Institute of Philosophy and Sociology, Polish Academy of Sciences, Warsaw, Poland; [b]Department of Philosophy and History of Ideas, Aarhus University: Aarhus Universitet, Denmark and [c]Department of Philosophy, Lingnan University, Hong Kong
wiktor.rorot.research@gmail.com
urte.laukas@berkeley.edu
davidharrison@ln.edu.hk

*Corresponding author.

doi:10.1017/S0140525X26106207, e475

Abstract

The target article underscores that metabolic considerations can accommodate the standard functionalist account of cognition. However, the energetic makeup of (biological) information processing systems in fact supports the claim of medium dependence of cognitive capacities. We explore two case studies from active matter physics and basal cognition to highlight the philosophical implications of the physicality of information processing.

The physicality of information processing systems has largely been sidelined by the standard functionalist picture. The primary philosophical upshot of Haueis and Colaço's analysis of metabolic considerations concerns the functionalist tenets of multiple realizability (Fodor, 1974; Polger & Shapiro, 2016; Putnam, 1967) and medium independence (Drayson, 2025; Garson, 2003; Haugeland (1993 [1985]); Piccinini, 2015, 2021). They submit that the metabolic underpinnings of cognition do not provide decisive arguments against these positions. However, in addition to metabolism, a naturalistic account of cognition requires engaging with the physicality of the information processing vehicles. Such considerations problematize the ontological agnosticism adopted by the

authors. They propel us to reject medium independence, while allowing for a degree of multiple realizability.

The formal view of computation, encoded in the medium independence thesis (see Drayson, 2025), holds that any physical implementation of a formal model of computation is equivalent to any other. From this perspective, the causally and explanatorily relevant features of physical computation can be abstracted from their implementation. Information, the primary resource of computation, becomes immaterial as well. This view has been captured in structural terms by positing that information is carried by the *differences* in the medium rather than the medium itself (e.g., Piccinini, 2015).

While this idealized picture of computation is powerful, grounding various empirical achievements within computational neuroscience, for the purposes of naturalization projects, it must be supplemented with the constraints that physical implementations introduce. The extensive research on aspects of physical computation (e.g., Chirimuuta, 2024; Copeland, 1996; Horsman et al., 2014; Miłkowski, 2013, 2018; Piccinini, 2015, 2021; Shagrir, 2022) contrasts with the relatively low interest in the physicality of information, with the notable exceptions of Collier (1996, 1999, 2000, 2008, 2011) and Deacon (2007, 2008, 2011, 2021; Deacon & García-Valdecasas, 2023; Deacon, García-Valdecasas, & Monzón-Marqués, 2026).

For Collier, information content is specified through the Negentropy Principle of Information, which compares the actual thermodynamic entropy of a system with its theoretical maximal entropy. This is closely related to the Gibbs free energy – the ability of a (closed) system to do useful work or, as Collier interprets it, the ability to control a physical process (Collier, 1999; see also Rorot, 2025). In a similar vein, Deacon (2007, p. 146) argues that "a sign/signal with any given Shannon information capacity has achieved it due to extrinsically imposed physical (thermodynamic) work," highlighting the relational, constraining character of physical information.

Consider the following examples of systems exhibiting cognition-like behaviors, which underline the role of physicality in information processing (see also Harrison, Rorot, & Laukaityte, 2022; Lyon, 2020; Lyon et al., 2021):

1. Oil droplets, non-biological active matter systems, are studied as precursors to biotic materials. As a result of physico-chemical instability, the droplets can move in complex ways. For instance, they can navigate mazes using chemical gradients (Čejková et al., 2014; Lagzi et al., 2010). This capability responds primarily to surface tension, but some droplet systems have been shown to exhibit a degree of autonomy or decoupling from the environment, including memory effects (Čejková et al., 2017; Hanczyc & Ikegami, 2010) and metabolism-like homeostasis (Hanczyc, 2011), on account of its energetic non-equilibrium state. The thermodynamic relationship between the system and its environment has informational character due to how the chemical gradient constrains and controls the system. These relationships can also be exploited by an external user, as the development of chemical computation platforms shows (e.g., Adamatzky, 2009).
2. Slime mould *Physarum polycephalum*, a unicellular amoeba, has complex navigational and decision-making capacities. It consists of a single, macroscopic, mobile cell with multiple nuclei and an intracellular cytoskeleton that scaffolds a network of veins. It exhibits oscillatory chemical signaling, which is hypothesized to underpin its ability to forage for food, as well as learning and memory (Boussard et al., 2021; Sims, 2025). As such, these oscillations are taken to enable information processing and storage, thereby serving as the focus of a functional description of *P. polycephalum's* behaviour. This motivates research exploring the connection to electrical oscillations in the nervous system (Boussard et al., 2021; see also Hanson, 2021) and the development of robotic models of slime mould navigation (Umedachi et al., 2013). Such models draw from the framework of morphological computation, underscoring the role of "physical form" (Paul, 2006; see also Rorot, 2022) and the medium-dependent character of information processing.

These examples bolster two crucial insights regarding multiple realizability: first, if information comprises certain thermodynamic constraints between the system and its environment, only systems with a particular energetic makeup (active matter systems, see Harrison, 2024, forthcoming; Harrison, Rorot, & Laukaityte, 2022; McGivern, 2020) can be considered to be processing information. Second, at least for some forms of (biological) computation, their implementation relies on the physical properties of the system. Together, these features suggest non-arbitrary, medium-dependent conditions for cognitive processing, which significantly limit what types of systems can in fact realize the relevant cognitive functions. Relatedly, they might also impact our taxonomies of cognitive functions, calling for a more fine-grained classification (Chirimuuta, 2018, 2024; Godfrey-Smith, 2016; Harrison, Rorot, & Laukaityte, 2022).

We maintain that the inherent physicality of information processing systems supports the view of metabolism as constitutive of (rather than merely contributing to) cognition. *Pace* the agnostic perspective of Haueis and Colaço, the thermodynamic underpinnings of information processing significantly limit the scope of the medium independence thesis. In addition to the general point drawn from Thagard (2022, p. 84) that metabolic "energy depends on material substrates," work in active matter physics ups the ante by indicating the requisite energetic structure for information processing. In so doing, it specifies what forms of physical media are capable of supporting cognitive functions.

Rejecting the medium independence claim but accommodating a degree of multiple realizability, our proposed perspective opens up a new viewpoint for cognitive science. For example, foregrounding the energetic characteristics of cognitive systems and their role in information processing can inform research in psychiatry (e.g., Mertens et al., 2015; see Palmer, 2022), enhance understanding of glial function in the nervous system (e.g., Hasel et al., 2017), among other underexplored research avenues (Rae et al., 2024).

Acknowledgments. WR wants to thank Joanna Rączaszek-Leonardi and Marcin Miłkowski for many discussions and guidance and Daniel Piecka for conversations about the nature of information and comments on a draft of this commentary.

Financial support. WR has been supported by the Preludium grant from the National Science Center (Poland), grant number 2022/45/N/HS1/02434.

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

Competing interests. The authors declare no conflicts of interest.

References

Adamatzky, A. (2009). Reaction-diffusion computing. In R. A. Meyers (Ed.), *Encyclopedia of complexity and systems science* (pp. 7548–7565). Springer New York. https://doi.org/10.1007/978-0-387-30440-3_446.

Boussard, A., Fessel, A., Oettmeier, C., Briard, L., Döbereiner, H.-G., & Dussutour, A. (2021). Adaptive behaviour and learning in slime moulds: The role of oscillations. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *376*(1820), 20190757. https://doi.org/10.1098/rstb.2019.0757.

Čejková, J., Banno, T., Hanczyc, M. M., & Štěpánek, F. (2017). Droplets as liquid robots. *Artificial Life*, *23*(4), 528–549. https://doi.org/10.1162/ARTL_a_00243.

Čejková, J., Novák, M., Štěpánek, F., & Hanczyc, M. M. (2014). Dynamics of chemotactic droplets in salt concentration gradients. *Langmuir*, *30*(40), 11937–11944. https://doi.org/10.1021/la502624f.

Chirimuuta, M. (2018). Marr, Mayr, and MR: What functionalism should now be about. *Philosophical Psychology*, *31*(3), 403–418. https://doi.org/10.1080/09515089.2017.1381679.

Chirimuuta, M. (2024). *The brain abstracted: Simplification in the history and philosophy of neuroscience*. The MIT Press.

Collier, J. D. (1996). Information originates in symmetry breaking. *Symmetry: Science and Culture*, *7*, 247–256.

Collier, J. D. (1999). Causation is the transfer of information. In H. Sankey (Ed.), *Causation and laws of nature* (pp. 215–245). Springer Netherlands. https://doi.org/10.1007/978-94-015-9229-1_18.

Collier, J. D. (2000). Information theory as a general language for functional systems. *AIP Conference Proceedings*, *517*, 124–130. https://doi.org/10.1063/1.1291251.

Collier, J. D. (2008). Information in biological systems. In P. Adriaans, & J. van Benthem (Eds.), *Handbook of philosophy of science, Volume 8: Philosophy of information* (pp. 763–787). Elsevier. https://doi.org/10.1016/B978-0-444-51726-5.50024-8.

Collier, J. D. (2011). Information, causation and computation. In G. Dodig-Crnkovic, & M. Burgin (Eds.), *Information and computation* (pp. 89–105). World Scientific. https://doi.org/10.1142/97898142954823_0004.

Copeland, B. J. (1996). What is computation? *Synthese*, *108*(3), 335–359. https://www.jstor.org/stable/20117547.

Deacon, T. W. (2007). Shannon – Boltzmann – Darwin: Redefining information (Part I). *Cognitive Semiotics*, *1*(s1), 123–148. https://doi.org/10.1515/cogsem.2007.1.fall2007.123.

Deacon, T. W. (2008). Shannon - Boltzmann — Darwin: Redefining information (Part II). *Cognitive Semiotics*, *2*(Supplement), 169–196. https://doi.org/10.1515/cogsem.2008.2.spring2008.169.

Deacon, T. W. (2011). *Incomplete nature: How mind emerged from matter* (1st ed). W.W. Norton & Co.

Deacon, T. W. (2021). How molecules became signs. *Biosemiotics*, *14*(3), 537–559. https://doi.org/10.1007/s12304-021-09453-9.

Deacon, T. W., & García-Valdecasas, M. (2023). A thermodynamic basis for teleological causality. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *381*(2252), 20220282. https://doi.org/10.1098/rsta.2022.0282.

Deacon, T. W., García-Valdecasas, M., & Monzón-Marqués, P. (2026). Autogenesis: An alternative path to molecular information. *Biological Theory*. https://doi.org/10.1007/s13752-025-00528-1.

Drayson, Z. (2025). Defending the medium-independence of computation. *Mind & Language*, *40*(4), 458–467. https://doi.org/10.1111/mila.12536.

Fodor, J. A. (1974). Special sciences (or: The disunity of science as a working hypothesis). *Synthese*, *28*(2), 97–115. https://doi.org/10.1007/BF00485230.

Garson, J. (2003). The introduction of information into neurobiology. *Philosophy of Science*, *70*(5), 926–936. https://doi.org/10.1086/377378.

Godfrey-Smith, P. (2016). Mind, matter, and metabolism. *Journal of Philosophy*, *113*(10), 481–506. https://doi.org/10.5840/jphil20161131034.

Hanczyc, M. M. (2011). Metabolism and motility in prebiotic structures. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1580), 2885–2893. https://doi.org/10.1098/rstb.2011.0141.

Hanczyc, M. M., & Ikegami, T. (2010). Chemical basis for minimal cognition. *Artificial Life*, *16*(3), 233–243. https://doi.org/10.1162/artl_a_00002.

Hanson, A. (2021). Spontaneous electrical low-frequency oscillations: A possible role in *Hydra* and all living systems. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *376*(1820), 20190763. https://doi.org/10.1098/rstb.2019.0763.

Harrison, D. (2024). Mind the Matter: Active Matter, Basal Cognition, and the Making of Bio-Inspired Artificial Intelligence [PhD dissertation]. University of Cambridge. https://doi.org/10.17863/CAM.108493.

Harrison, D., Rorot, W., & Laukaityte, U. (2022). Mind the matter: Active matter, soft robotics, and the making of bio-inspired artificial intelligence. *Frontiers in Neurorobotics*, *16*, 880724. https://doi.org/10.3389/fnbot.2022.880724.

Harrison, D. J. (forthcoming). Rethinking the nature of matter: Active matter physics and the naturalisation of teleology. *Erkenntnis*. https://doi.org/10.1007/s10670-026-01123-x.

Hasel, P., Dando, O., Jiwaji, Z., Baxter, P., Todd, A. C., Heron, S., Márkus, N. M., McQueen, J., Hampton, D. W., Torvell, M., Tiwari, S. S., McKay, S., Eraso-Pichot, A., Zorzano, A., Masgrau, R., Galea, E., Chandran, S., Wyllie, D. J. A., Simpson, T. I., & Hardingham, G. E. (2017). Neurons and neuronal activity control gene expression in astrocytes to regulate their development and metabolism. *Nature Communications*, *8*(1), 15132. https://doi.org/10.1038/ncomms15132.

Haugeland, J. (1993). *Artificial intelligence: The very idea*. MIT Press. (Original work published 1985).

Horsman, C., Stepney, S., Wagner, R. C., & Kendon, V. (2014). When does a physical system compute? *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *470*(2169), 20140182. https://doi.org/10.1098/rspa.2014.0182.

Lagzi, I., Soh, S., Wesson, P. J., Browne, K. P., & Grzybowski, B. A. (2010). Maze solving by chemotactic droplets. *Journal of the American Chemical Society*, *132*(4), 1198–1199. https://doi.org/10.1021/ja9076793.

Lyon, P. (2020). Of what is "minimal cognition" the half-baked version? *Adaptive Behavior*, *28*(6), 407–424. https://doi.org/10.1177/1059712319871360.

Lyon, P., Keijzer, F., Arendt, D., & Levin, M. (2021). Reframing cognition: Getting down to biological basics. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *376*(1820), 20190750. https://doi.org/10.1098/rstb.2019.0750.

McGivern, P. (2020). Active materials: Minimal models of cognition? *Adaptive Behavior*, *28*(6), 441–451. https://doi.org/10.1177/1059712319891742.

Mertens, J., Wang, Q.-W., Kim, Y., Yu, D. X., Pham, S., Yang, B., Zheng, Y., Diffenderfer, K. E., Zhang, J., Soltani, S., Eames, T., Schafer, S. T., Boyer, L., Marchetto, M. C., Nurnberger, J. I., Calabrese, J. R., Oedegaard, K. J., McCarthy, M. J., Zandi, P. P., Alda, M., Nievergeldt, C. M., The Pharmacogenomics of Bipolar Disorder Study, Mi, S., Brennand, K. J., Kelsoe, J. R., Gage, F. H., & Yao, J. (2015). Differential responses to lithium in hyperexcitable neurons from patients with bipolar disorder. *Nature*, *527*(7576), 95–99. https://doi.org/10.1038/nature15526.

Miłkowski, M. (2013). *Explaining the computational mind*. The MIT Press.

Miłkowski, M. (2018). From computer metaphor to computational modeling: The evolution of computationalism. *Minds and Machines*, *28*(3), 515–541. https://doi.org/10.1007/s11023-018-9468-3.

Palmer, C. M. (2022). *Brain energy: A revolutionary breakthrough in understanding mental health and improving treatment for anxiety, depression, OCD, PTSD, and more*. BenBella Books.

Paul, C. (2006). Morphological computation. *Robotics and Autonomous Systems*, *54*(8), 619–630. https://doi.org/10.1016/j.robot.2006.03.003.

Piccinini, G. (2015). *Physical computation: A mechanistic account*. Oxford University Press.

Piccinini, G. (2021). *Neurocognitive mechanisms: Explaining biological cognition*. Oxford University Press.

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

Polger, T. W., & Shapiro, L. A. (2016). *The multiple realization book*. Oxford University Press.
Putnam, H. (1967). Psychological predicates. In W. H. Capitan, & D. D. Merrill (Eds.), *Art, mind, and religion* (pp. 37–48). University of Pittsburgh Press.
Rae, C. D., Baur, J. A., Borges, K., Dienel, G., Díaz-García, C., Douglass, S. R., Drew, K., Duarte, J., Duran, J., Kann, O., Kristian, T., Lee-Liu, D., Lindquist, B. E., McNay, E. C., Robinson, M. B., Rothman, D. L., Rowlands, B. D., Ryan, T. A., Scafidi, J., Scafidi, S., Shuttleworth, C. W., Swanson, R. A., Uruk, G., Vardjan, N., Zorec, R., & McKenna, M. C. (2024). Brain energy metabolism: A roadmap for future research. *Journal of Neurochemistry*, *168*, 910–954. https://doi.org/10.1111/jnc.16032.
Rorot, W. (2022). Counting with cilia: The role of morphological computation in basal cognition research. *Entropy*, *24*(11), 1581. https://doi.org/10.3390/e24111581.
Rorot, W. (2025). Scale-free communication? An investigation of the use of the concept 'communication' in biology and cognitive sciences [PhD Dissertation]. University of Warsaw.
Shagrir, O. (2022). *The nature of physical computation*. Oxford University Press.
Sims, M. (2025). *Slime mould and philosophy*. Cambridge University Press. https://doi.org/10.1017/9781009488648.
Thagard, P. (2022). Energy requirements undermine substrate independence and mind-body functionalism. *Philosophy of Science*, *89*(1), 70–88. https://doi.org/10.1017/psa.2021.15.
Umedachi, T., Idei, R., Ito, K., & Ishiguro, A. (2013). A fluid-filled soft robot that exhibits spontaneous switching among versatile spatiotemporal oscillatory patterns inspired by the true slime mold. *Artificial Life*, *19*(1), 67–78. https://doi.org/10.1162/ARTL_a_00081.

Rationality is metabolically bounded, not cognitively biased

Pier Luigi Sacco*

Neuroscience Imaging & Clinical Studies and BACH Center, University of Chieti-Pescara, Chieti, Italy
pierluigi.sacco@unich.it

*Corresponding author.

doi:10.1017/S0140525X26106323, e476

Abstract

Haueis and Colaço's framework can be inscribed into a positive theory of metabolic rationality where cognitive strategy selection is governed by allostatic energy management across sequential decisions under uncertainty. Heuristics emerge as solutions to a metabolic opportunity cost problem with distinct empirical predictions.

Haueis and Colaço argue that metabolic considerations should function both as evaluative constraints and as generative tools for cognitive modeling. I propose that their framework offers support for something more specific: a positive theory of *metabolic rationality* where cognitive strategy selection is itself a metabolic resource allocation problem. On this account, the brain dynamically allocates metabolic resources across a temporal horizon of anticipated cognitive demands, treating every act of deliberation as carrying a metabolic opportunity cost, i.e., the reduced capacity it creates for future cognition.

The empirical basis for this proposal is now substantial. Prefrontal and associative networks supporting deliberative reasoning have up to 67% higher energetic costs of signaling than sensorimotor regions (Castrillón et al., 2023), and brain wiring reflects optimal tradeoffs between connectivity costs and adaptive advantage (Bullmore & Sporns, 2012). The critical insight goes beyond the well-known cost of deliberation: the brain's energy budget is *finite and temporally extended*. An organism facing a continuous stream of decisions with unpredictable timing and importance cannot evaluate each in isolation. It must manage its metabolic reserves across the entire sequence. A physician in an emergency department, a forager in an uncertain landscape, and a parent navigating competing demands – all face continuous streams of decisions where exhausting metabolic resources on any single choice risks catastrophic failure on subsequent, potentially more critical ones. This transforms cognitive strategy selection from a local accuracy problem into a global resource allocation problem.

Allostatic regulation provides the physiological mechanism for this management. Sterling (2012) argues that the brain proactively anticipates future demands and adjusts resource allocation accordingly, going beyond reaction to current needs. Applied to cognition, this implies that the brain continuously forecasts the metabolic demands of upcoming cognitive tasks and calibrates the precision of current processing accordingly. The neuromodulatory architecture supporting this calibration is revealing: the brain supplements precise synaptic computation with diffuse neuromodulatory systems that modulate the gain of neural populations at low metabolic cost (Attwell & Laughlin, 2001; see also Shine et al., 2021). These systems set the *level of computational investment* that a given situation warrants, functioning as a metabolic resource allocator.

Predictive coding frameworks offer a natural computational substrate for this theory. Under predictive processing, the brain minimizes free energy through prediction error coding, a scheme that transmits only deviations from expectations, reserving metabolic expenditure for surprising information (Clark, 2013; Friston, 2010). This architecture is inherently metabolically conservative: stable environments generate small prediction errors, licensing low-precision processing and minimal metabolic expenditure, while novel or volatile environments generate large errors, triggering metabolically expensive precision-weighted updating. Crucially, predictive processing is *hierarchically organized*. Lower levels of the cortical hierarchy handle routine predictions at low metabolic cost; only when prediction errors cannot be resolved locally do they propagate upward to higher associative and prefrontal areas, that is, those regions that Castrillón et al. (2023) show to be most metabolically expensive. This means that the brain has an endogenous escalation mechanism: deliberative processing functions as a costly upper tier that the system recruits only when cheaper levels of the hierarchy fail to absorb prediction error. The apparent binary switch between System 1 and System 2 of dual-process models thus dissolves into a continuous metabolic gradient along the cortical hierarchy.

This hierarchical precision-weighting mechanism reframes several well-documented decision phenomena. Anchoring, for instance, becomes the metabolically rational reliance on existing priors when prediction errors are insufficiently large to justify the cost of updating higher-level models. Status quo preference reflects the brain's default maintenance of established generative models whose predictions remain adequate, as updating is metabolically justified only when errors accumulate beyond a threshold set by the

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

allostatic assessment of available resources. Loss aversion, in turn, can be understood through asymmetric precision weighting: negative prediction errors (worse-than-expected outcomes) receive higher precision than positive ones because they signal potential threats to metabolic homeostasis, warranting greater computational investment in updating (Palumbo et al., 2006). All three phenomena are manifestations of a single underlying principle: the brain's hierarchical allocation of metabolic resources tracks the expected informational value of further processing relative to its energetic cost.

This framework generates predictions that are distinct from both standard behavioral economics and ecological rationality accounts. First, cognitive strategy should vary with the organism's current and anticipated metabolic state and with task structure (as ecological rationality predicts). Decision-makers should become more heuristic when subsequent demands are expected to be high, independently of task simplicity, a prediction about *sequential task interaction* that neither framework makes. Second, what looks like inconsistent risk attitudes across contexts may reflect metabolic state-dependence: risk-sensitive behavior shifts with the brain's allostatic assessment of available resources relative to anticipated demands, so that risk preferences function as adaptive signals of the organism's current energetic state (Palumbo et al., 2006). Third, the framework predicts that institutional and cultural environments that reduce cognitive unpredictability should lower the metabolic cost of coordination, enabling more efficient collective decision-making, a social-level prediction that connects metabolic neuroscience to institutional design (North, 1990; Veissière et al., 2020).

Haueis and Colaço have provided cognitive science with further arguments to take metabolic constraints seriously. The natural next step is to extend such arguments to decision theory. The metabolic rationality framework sketched here suggests that human cognition is *metabolically rational*: adapted to manage a scarce energetic resource across the uncertain demands of an extended temporal horizon. This is a story about a different kind of intelligence: one that optimizes for sustained adaptive performance across a lifetime of decisions. Once we recognize this, the task for both science and policy shifts from correcting a supposedly flawed reasoner to understanding, and designing for, an exquisitely efficient one.

Financial support. This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

Competing interests. None.

References

Attwell, D., & Laughlin, S. B. (2001). An energy budget for signaling in the grey matter of the brain. *Journal of Cerebral Blood Flow and Metabolism*, *21*(10), 1133–45.

Bullmore, E., & Sporns, O. (2012). The economy of brain network organization. *Nature Reviews Neuroscience*, *13*(5), 336–49.

Castrillón, G., Epp, S., Bose, A., Fraticelli, L., Hechler, A., Belenya, R., Lio, G., Sorg, C., Straulino, E., Yakushev, I., & Riedl, V. (2023). An energy costly architecture of neuromodulators for human brain evolution and cognition. *Science Advances*, *9*(50), eadi7632.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(3), 181–204.

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, *11*(2), 127–38.

North, D. C. (1990). *Institutions, institutional change and economic performance*. Cambridge University Press.

Palumbo, R., Bortolotti, A., & Sacco, P. L. (2006). Prospect theory as active inference: A metabolic account of risk-sensitive decision-making. SocArXiv preprint https://osf.io/preprints/socarxiv/2d9v5_v1.

Shine, J. M., Müller, E. J., Bose, A., Breakspear, M., & Moran, R. J. (2021). Computational models link cellular mechanisms of neuromodulation to large-scale neural dynamics. *Nature Neuroscience*, *24*(6), 765–76.

Sterling, P. (2012). Allostasis: A model of predictive regulation. *Physiology & Behavior*, *106*(1), 5–15.

Veissière, S. P. L., Constant, A., Ramstead, M. J. D., Friston, K. J., & Kirmayer, L. J. (2020). Thinking through other minds: A variational approach to cognition and culture. *Behavioral and Brain Sciences*, *43*, e90.

Metabolic constraints in cognitive modeling: Generative or merely evaluative?

Ayoob Shahmoradi*

Ruhr University Bochum, Bochum, Germany
ayoobsh@gmail.com

*Corresponding author.

doi:10.1017/S0140525X26106049, e477

Abstract

The central distinction – between evaluative and generative metabolic considerations – appears less theoretically secure than Haueis and Colaço suggest. First, the generative examples may ultimately reduce to evaluative reasoning applied earlier in the modeling process. Second, the principal generative examples do not clearly show metabolic considerations functioning within a cognitive model.

Haueis and Colaço (H&C) make a compelling case that metabolic considerations are underutilized in cognitive modeling and should be more systematically integrated into modeling practice. However, the central distinction between evaluative and generative metabolic considerations may not be as theoretically robust as the authors suggest. I explain this through two worries. The first – *the collapse worry* – concerns the methodological role metabolism plays in the generative examples: I argue that these reduce to evaluative reasoning applied earlier in the modeling process. The second – *the parallel tracks worry* – asks whether their core generative example illustrates metabolic considerations doing cognitive-modeling work at all, or whether metabolic and cognitive reasoning simply run alongside each other without genuine integration.

The collapse worry

H&C characterize evaluative considerations as constraining the space of possible models by pruning implausible ones and partitioning the remainder by relative plausibility. Generative considerations, by contrast, are said to populate that space with new models. They present this as a genuine difference in the methodological role that metabolic knowledge plays in modeling. However, the examples offered in support of the generative role are

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

better understood as instances of evaluative reasoning applied at an earlier stage of the modeling process.

Consider their core generative example: the efficiency-trade-off hypothesis of Theriault et al. (2023). The central move here is the observation that axons thin enough to maximize information–energy efficiency are also too thin to house mitochondria, making oxidative phosphorylation structurally impossible in these fibers. Aerobic glycolysis is thus identified as the only metabolically feasible mechanism for fueling spike transmission in such axons. This is, at its core, a constraint-based argument: it rules out one computational implementation as metabolically impossible and identifies another as the only metabolically viable alternative. The predictive coding framework is then fitted to what survives this constraint. The metabolic reasoning does not so much originate the model as determine which preexisting computational frameworks are compatible with the metabolic facts. This looks like evaluation applied during model construction, not a distinct generative process.

This suggests that what H&C call generative considerations are evaluative constraint reasoning applied earlier in modeling – during construction rather than assessment. On this reading, the difference is temporal rather than logical or methodological: it concerns when metabolic constraints enter the process, not their methodological role.

The parallel tracks worry

Another problem concerns the relationship between the metabolic reasoning and the cognitive model in the Theriault et al. case. Even setting aside the collapse worry, it is unclear whether the metabolic considerations here are genuinely integrated into the cognitive model or whether they are solving a different puzzle that runs alongside it.

The metabolic puzzle the efficiency-trade-off hypothesis addresses is: why does the brain use aerobic glycolysis – a production-inefficient process – when oxidative phosphorylation is available? The answer follows entirely from structural and metabolic facts about thin axons: they cannot house mitochondria, so aerobic glycolysis is the only viable energy source during peak demand. This puzzle is resolved without any reference to prediction errors or predictive coding. The cognitive model is then appended to the metabolic solution – thin axons happen to have a firing profile compatible with prediction error transmission, so the two are connected. But this connection looks more like an observation of compatibility than genuine explanatory integration. The metabolic considerations solve a metabolic puzzle, and the cognitive model is brought in afterward to interpret what that solution means computationally.

This matters because H&C's central thesis is that metabolic considerations should be integrated into cognitive modeling. But if metabolic reasoning is doing its main work in a different explanatory register – resolving a puzzle about brain energy use rather than directly informing computational claims – then the example illustrates the co-existence of metabolic and cognitive reasoning rather than their integration. The two are linked by the common substrate of thin axons, but the metabolic considerations do not directly shape the cognitive model itself.

For metabolic considerations to play a genuinely generative role, they would need to enter into the computational story rather than merely constraining the physical implementation of independently motivated computational architectures. The Theriault et al. example stops well short of this, establishing primarily that aerobic glycolysis is compatible with prediction-error transmission, which is a considerably weaker claim.

The same analysis applies to the supporting generative example of Pache and Rossum (2023). Their exploration of energy-efficient learning strategies proceeds by imposing metabolic constraints on an already-existing neural network architecture trained with backpropagation – again, evaluative constraint satisfaction applied to an existing model space. But the parallel tracks worry applies here too, in a degraded form: unlike Theriault et al., there is no specific metabolic puzzle driving the inquiry. The starting point is the generic assumption that energy efficiency is a biological design principle – precisely the kind of generic metabolic consideration H&C themselves criticize in section 2.1.2 as insufficient for genuine cognitive modeling. Metabolic and cognitive reasoning remain only loosely connected rather than running on truly parallel tracks. Moreover, the model is implemented in artificial neural networks, making its biological relevance more indirect than the authors' generative framework might ideally require. The generative account, therefore, appears to lack a clear empirical example in which metabolic reasoning demonstrably contributes to the construction of a cognitive model, rather than constraining its implementation.

Conclusion

Neither of these criticisms undermines H&C's valuable conclusion that metabolism deserves a more prominent role in cognitive modeling. The evaluative account is well-supported and carefully developed. The generative account, however, rests on examples that either collapse into evaluation or fail to demonstrate genuine integration of metabolic considerations into cognitive models. The difficulty is, therefore, not that metabolism fails to constrain cognitive models, but that the paper's examples do not clearly show metabolism contributing to the construction of those models unless "cognitive modeling" is defined to include metabolic implementation by default. A stronger case for the generative role would require examples in which metabolic reasoning directly contributes to the computational story, rather than solving an adjacent metabolic puzzle to which a cognitive model is subsequently attached. Whether such examples exist remains an open question.

Acknowledgments. I am grateful to the Alexander von Humboldt Foundation for a postdoctoral fellowship hosted by Prof. Albert Newen at Ruhr University Bochum.

Financial support. None.

Competing interests. None.

References

Pache, A., & Van Rossum, M. C. W. (2023). Energetically efficient learning in neuronal networks. *Current Opinion in Neurobiology*, *83*, 102779. https://doi.org/10.1016/j.conb.2023.102779

Theriault, J. E., Shaffer, C., Dienel, G. A., Sander, C. Y., Hooker, J. M., Dickerson, B. C., Barrett, L. F., & Quigley, K. S. (2023). A functional account of stimulation-based aerobic glycolysis and its role in interpreting BOLD signal intensity increases in neuroimaging experiments. *Neuroscience & Biobehavioral Reviews*, *153*, 105373. https://doi.org/10.1016/j.neubiorev.2023.105373

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

Metabolism: A constraint and driver of neurodevelopment

Maheen Siddiqui[a]* and Emily J. H. Jones[a,b]

[a]Centre for Brain and Cognitive Development, Birkbeck, University of London, London, UK and [b]Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK
m.siddiqui@bbk.ac.uk
e.jones@bbk.ac.uk

*Corresponding author.

doi:10.1017/S0140525X26106244, e478

Abstract

This commentary extends recent accounts of metabolic constraints in cognitive modelling by considering their role in brain development. We argue that shifting energy availability both constrains and drives the emergence of functional specialisation and network organisation. Integrating interactive specialisation, we show how metabolic factors shape competitive and cooperative interactions, and how atypical energy metabolism may contribute to neurodivergent developmental trajectories.

Haueis and Colaço argue for incorporating metabolic constraints into cognitive modelling, suggesting that this may improve the biological plausibility of models of cognitive function. The authors propose evaluative and generative metabolic constraints that can refine existing cognitive models and guide the development of new ones. In this commentary, we build on the proposed framework by taking a developmental perspective, where metabolic constraints may play an even more critical role. At birth, the brain is reported to consume approximately 50–60% of total body metabolism (Kuzawa, 1998), decreasing over the first six months before rising alongside rapid synaptic and network growth (Kuzawa et al., 2014). By age 5, brain energy use peaks at levels 2–3 times higher than in adults, utilising approximately 66% of resting metabolic rate and 43% of daily energy expenditure (Kuzawa & Blair, 2019). These shifts coincide with rapid structural development, network organisation, and the emergence of functional specialisation. A metabolism-informed developmental perspective, therefore, can extend the authors' framework by demonstrating how emerging energetic constraints may shape the trajectory of cognitive development in both typical and neurodivergent developmental pathways. Within the proposed framework, *evaluative* approaches may determine whether a given model is metabolically feasible at a particular age, whereas *generative* approaches may use changes in energy availability and expenditure to explain how variation in developmental trajectories occurs.

In adult models, metabolic constraints are treated as relatively stable. However, in development, energy availability/expenditure changes over time. This makes *evaluative* metabolic considerations particularly important for assessing the plausibility of developmental cognitive models and helps rule out models that assume capacities, such as sustained neural activity or complex long-range connectivity, that may not be feasible given an infant brain's limited energy budget. A clear example comes from early neural activity patterns observed in premature infants. Electroencephalography studies show characteristic "delta brush" activity, consisting of brief bursts of oscillatory activity followed by periods of relative quiescence (Whitehead, Pressler, & Fabrizi, 2017). One interpretation is that these discontinuous patterns may reflect underlying metabolic constraints; sustained activity may be energetically too costly for the immature brain, necessitating intermittent bursts separated by recovery periods. From an evaluative metabolic perspective, this pattern suggests that models assuming continuous or sustained neural firing in early development may be implausible, as they do not account for the limited and fluctuating energy availability at this stage.

Using a *generative* approach, metabolism can also help explain developmental changes. Increasing energetic capacity may support the emergence of more complex processing and the gradual specialisation of brain networks, for example, in the development of executive functions or social cognition. Over the first few years of life, the brain becomes increasingly specialised and efficient, and this shift may be partly driven by changes in metabolic capacity. Thereby suggesting that metabolism may play an active role in shaping developmental changes and simultaneously act as a constraint that ultimately determines which neural pathways are strengthened or pruned and how cognitive functions emerge over time. Computational modelling work supports this view by demonstrating that brain network organisation emerges from trade-offs between energy cost, communication efficiency, and functional demands (Fakhar et al., 2025).

Drawing on the interactive specialisation framework (Johnson, 2011), the developing brain can be seen as a self-organising network in which cortical regions cooperate and compete to acquire functional specialisation. Metabolic efficiency and energetic costs may shape these competitive and cooperative interactions. That is, regions that require lower energy expenditure to perform computations are more likely to be repeatedly recruited, thereby strengthening connections. Similarly, cooperative interactions may emerge when regions coordinate activity in ways that minimise total energy expenditure, stabilising networks that consistently activate together. Changes in energetic capacity across development, owing to factors such as vascular development (Catalina Camacho et al., 2020; Chiron et al., 1992; Takahashi et al., 1999; Yu et al., 2023), neural circuitry development (Bianchi et al., 2013; Huttenlocher, 1990; Huttenlocher & Dabholkar, 1997; Le Magueresse & Monyer, 2013), and the metabolic environment (Chugani, Phelps, & Mazziotta, 1987; Takahashi et al., 1999), may constrain which neural computations are feasible at each developmental stage, biasing the formation of short-range neural connections before more costly long-range connections can be supported. Resting-state functional MRI studies (Zhang, Shen, & Lin, 2019) illustrate the developmental emergence of long-range functional connectivity over the first year. This provides evidence for the brain favouring local, low-cost connections, with integration across distant regions becoming possible as energetic capacity increases.

Lastly, a metabolism-informed account of development can potentially provide insight into neurodivergent development more broadly, where early differences in neural structure, circuitry, and functional networks may interact with metabolic constraints to

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

shape alternative trajectories of cortical specialisation. Evidence suggests that mitochondrial dysfunction, region-specific differences in metabolic activity, and atypical oxygen and glucose metabolism may limit the energy available to support neural computations across brain regions (Anitha et al., 2012; Ginsberg et al., 2012). These alterations in energy metabolism have been implicated across a range of neurodevelopmental and neurological conditions, including autism (Chauhan & others, 2011; Gu et al., 2013), attention deficit hyperactivity disorder (Pinto Payares et al., 2024), Alzheimer's disease and Parkinson's disease (Zong et al., 2024), and are likely to influence which regions can sustain neural activity and support computations at different developmental stages, possibly guiding competitive and collaborative interactions. Such interactions between metabolic constraints and neural circuitry may contribute to the diverse cognitive, sensory, and behavioural profiles observed across neurodevelopmental and neurological conditions, and conversely, the fact that pronounced cognitive impairments are reliably observed in mitochondrial disease suggests that energy metabolism plays a fundamental role in supporting normal brain function.

To conclude, a metabolism-informed developmental perspective extends Haueis and Colaço's framework. Across typical development, changes in metabolic capacity both constrain and guide the emergence of functional specialisation, shaping which neural pathways are strengthened or pruned and supporting the transition from specialised and integrated networks. By including metabolic considerations into developmental models, we improve our understanding of the emergence of typical mental functions and of the patterns of cognitive, sensory, and behavioural profiles observed in neurodevelopmental conditions.

Financial support. MS is funded Wellcome Trust (225099/Z/22/Z). EJHJ is supported by the Economic and Social Research Council (ES/R009368/1). EJHJ is also supported by the AIMS-2-TRIALS programmes funded by the Innovative Medicines Initiative Joint Undertaking Grant No. 777394. This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation programme, with in-kind contributions from the European Federation of Pharmaceutical Industries and Associations companies and funding from Autism Speaks, Autistica, and SFARI.

Competing interests. The author(s) declare none.

References

Anitha, A., Nakamura, K., Thanseem, I., Yamada, K., Iwayama, Y., Toyota, T., Matsuzaki, H., Miyachi, T., Yamada, S., Tsujii, M., Tsuchiya, K. J., Matsumoto, K., Iwata, Y., Suzuki, K., Ichikawa, H., Sugiyama, T., Yoshikawa, T., & Mori, N. (2012). Brain region-specific altered expression and association of mitochondria-related genes in autism. *Molecular Autism*, *3*(1), 12. https://doi.org/10.1186/2040-2392-3-12

Bianchi, S., Stimpson, C. D., Duka, T., Larsen, M. D., Janssen, W. G. M., Collins, Z., Bauernfeind, A. L., Schapiro, S. J., Baze, W. B., McArthur, M. J., Hopkins, W. D., Wildman, D. E., Lipovich, L., Kuzawa, C. W., Jacobs, B., Hof, P. R., & Sherwood, C. C. (2013). Synaptogenesis and development of pyramidal neuron dendritic morphology in the chimpanzee neocortex resembles humans. *Proceedings of the National Academy of Sciences*, *110*(Supplement_2), 10395–10401. https://doi.org/10.1073/pnas.1301224110

Catalina Camacho, M., King, L. S., Ojha, A., Garcia, C. M., Sisk, L. M., Cichocki, A. C., Humphreys, K. L., & Gotlib, I. H. (2020). Cerebral blood flow in 5- to 8-month-olds: Regional tissue maturity is associated with infant affect. *Developmental Science*, *23*(5), e12928. https://doi.org/10.1111/desc.12928

Chauhan, A., Gu, F., Essa, M. M., Wegiel, J., Kaur, K., Brown, W. T., & Chauhan, V. (2011). Brain region-specific deficit in mitochondrial electron transport chain complexes in children with autism. *Journal of neurochemistry*, *117*(2), 209–220. https://doi.org/10.1111/j.1471-4159.2011.07189.x

Chiron, C., Raynaud, C., Mazière, B., Zilbovicius, M., Laflamme, L., Masure, M. C., Dulac, O., Bourguignon, M., & Syrota, A. (1992). Changes in regional cerebral blood flow during brain maturation in children and adolescents. *Journal of Nuclear Medicine*, *33*(5), 696–703.

Chugani, H. T., Phelps, M. E., & Mazziotta, J. C. (1987). Positron emission tomography study of human brain functional development. *Annals of Neurology*, *22*(4), 487–497. https://doi.org/10.1002/ana.410220408

Fakhar, K., Akarca, D., Luppi, A. I., Oldham, S., Hadaeghi, F., Vértes, P. E., Bullmore, E., Hilgetag, C., & Astle, D. (2025). Human cortical networks trade communication efficiency for computational reliability. *BioRxiv*, 2025.12.11.693716. https://doi.org/10.64898/2025.12.11.693716

Ginsberg, M. R., Rubin, R. A., Falcone, T., Ting, A. H., & Natowicz, M. R. (2012). Brain transcriptional and epigenetic associations with autism. *PLoS ONE*, *7*(9), e44736. https://doi.org/10.1371/journal.pone.0044736

Gu, F., Chauhan, V., Kaur, K., Brown, W. T., LaFauci, G., Wegiel, J., & Chauhan, A. (2013). Alterations in mitochondrial DNA copy number and the activities of electron transport chain complexes and pyruvate dehydrogenase in the frontal cortex from subjects with autism. *Translational Psychiatry*, *3*(February), e299. https://doi.org/10.1038/tp.2013.68

Huttenlocher, P. R. (1990). Morphometric study of human cerebral cortex development. *Neuropsychologia*, *28*(6), 517–527. https://www.ncbi.nlm.nih.gov/pubmed/2203993

Huttenlocher, P. R., & Dabholkar, A. S. (1997). Regional differences in synaptogenesis in human cerebral cortex. *The Journal of Comparative Neurology*, *387*(2), 167–178. https://www.ncbi.nlm.nih.gov/pubmed/9336221

Johnson, M. H. (2011). Interactive Specialization: A domain-general framework for human functional brain development? *Developmental Cognitive Neuroscience*, *1*(1), 7–21. ISSN 1878-9293. https://doi.org/10.1016/j.dcn.2010.07.003

Kuzawa, C. W. (1998). Adipose tissue in human infancy and childhood: An evolutionary perspective. *American Journal of Physical Anthropology*, *107*(S27), 177–209. https://doi.org/10.1002/(SICI)1096-8644(1998)107:27+<177:AID-AJPA7>3.0.CO;2-B

Kuzawa, C. W., & Blair, C. (2019). A hypothesis linking the energy demand of the brain to obesity risk. *Proceedings of the National Academy of Sciences*, *116*(27), 13266–13275. https://doi.org/10.1073/pnas.1816908116

Kuzawa, C. W., Chugani, H. T., Grossman, L. I., Lipovich, L., Muzik, O., Hof, P. R., Wildman, D. E., Sherwood, C. C., Leonard, W. R., & Lange, N. (2014). Metabolic costs and evolutionary implications of human brain development. *Proceedings of the National Academy of Sciences*, *111*(36), 13010–13015. https://doi.org/10.1073/pnas.1323099111

Le Magueresse, C., & Monyer, H. (2013). GABAergic interneurons shape the functional maturation of the cortex. *Neuron*, *77*, 388–405. https://doi.org/10.1016/j.neuron.2013.01.011

Pinto Payares, D. V, Spooner, L., Vosters, J., Dominguez, S., Patrick, L., Harris, A., & Kanungo, S. (2024). A systematic review on the role of mitochondrial dysfunction/disorders in neurodevelopmental disorders and psychiatric/behavioral disorders. *Frontiers in Psychiatry*, *15*, 1389093. https://doi.org/10.3389/fpsyt.2024.1389093

Takahashi, T., Shirane, R., Sato, S., & Yoshimoto, T. (1999). Developmental changes of cerebral blood flow and oxygen metabolism in children. *American Journal of Neuroradiology*, *20*, 917–922. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.513.4751&rep=rep1&type=pdf

Whitehead, K., Pressler, R., & Fabrizi, L. (2017). Characteristics and clinical significance of delta brushes in the EEG of premature infants. *Clinical Neurophysiology Practice*, *2*, 12–18. https://doi.org/10.1016/j.cnp.2016.11.002

Yu, Q., Ouyang, M., Detre, J., Kang, H., Hu, D., Hong, B., Fang, F., Peng, Y., & Huang, H. (2023). Infant brain regional cerebral blood flow increases supporting emergence of the default-mode network. *ELife*, *12*, e78397. https://doi.org/10.7554/eLife.78397

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

Zhang, H., Shen, D., & Lin, W. (2019). Resting-state functional MRI studies on infant brains: A decade of gap-filling efforts. *NeuroImage*, *185*, 664–684. https://doi.org/10.1016/j.neuroimage.2018.07.004

Zong, Y., Li, H., Liao, P., Chen, L., Pan, Y., Zheng, Y., Zhang, C., Liu, D., Zheng, M., & Gao, J. (2024). Mitochondrial dysfunction: Mechanisms and advances in therapy. *Signal Transduction and Targeted Therapy*, *9*(1), 124. https://doi.org/10.1038/s41392-024-01839-8

Beyond neurocentrism: situating cognition within the metabolic demands of the whole organism

Francesca Starita*, Giorgio Li Pira and Giuseppe Di Pellegrino

Department of Psychology "Renzo Canestrari", University of Bologna, Bologna, Italy
francesca.starita2@unibo.it
giorgio.lipira2@unibo.it
g.dipellegrino@unibo.it

*Corresponding author.

doi:10.1017/S0140525X26106219, e479

Abstract

We reorient Haueis and Colaço's metabolic framework of cognitive modeling by drawing on neuroscientific evidence from animal motor control, reinforcement learning, and decision-making to show that cognition emerged to serve the metabolic demands of the whole organism, rather than those of the brain alone. We also illustrate the generative value of our reframing to model cognition in artificial systems.

Haueis and Colaço argue that neural metabolism should be treated as a constitutive constraint on cognitive models to assess their biological plausibility and, drawing on extensive neuroscientific evidence, propose ways to evaluate and generate such models. We agree that metabolism shapes neural computations and that ignoring it risks producing biologically implausible models of cognition. However, we identify two limitations in their framework. First, although the authors acknowledge interactions between the brain and the body, their account remains largely neurocentric, treating the brain in relative isolation from the organism in which it is embedded. Second, metabolism is viewed primarily as a constraint that limits neural computations underlying cognition. We argue instead that cognition is not simply limited by the metabolic constraints of the brain, but that cognition actually emerged to serve the metabolic demands of the whole organism. Through such extension and a reorientation of Haueis and Colaço's framework, we reframe the problem of cognitive modeling from how metabolism constrains neural computation and cognition to how cognition functions to meet the metabolic demands of the whole organism in the service of its survival imperative.

In animals, the nervous system is not metabolically autonomous: its energy supply is regulated at the level of the entire organism, where multiple physiological systems compete for limited resources (Barrett & Simmons, 2015). Thus, the architecture of the nervous system, and that of cognition and behavior, has emerged as a solution to the problem of controlling a specific body in interaction with its environment (Chiel & Beer, 1997; Cisek, 2019, 2022). The organization of the octopus' nervous system provides an illustration of this. With an invertebrate body, maintaining a point-to-point (somatotopic) brain representation of different body parts would be metabolically prohibitive. Instead, the octopus relies on a highly decentralized nervous system, with approximately two-thirds of its neurons located in its arms. This organization enables the use of peripheral sensory tagging for motor control: when an arm contacts a stimulus, the sensory input does not merely signal its presence but specifies the location for action, effectively serving as an address for motor commands. The brain just sends an *en passant* motor signal along the arm, which is only executed where it coincides with local sensory input. In this way, sensory input functions as a temporary, action-guiding map without the need for central somatotopic representations (Hochner et al., 2023). Thus, in the octopus, the implementation of motor control reflects a solution to organism-level metabolic constraints, rather than solely neural ones.

Crucially, the same metabolic principles extend to higher-level cognitive functions such as reinforcement learning and decision-making, where metabolism does not merely constrain but actively generates behavioral policies, shaping which actions are selected and how they are executed. Organisms continuously allocate limited energy across competing needs, such as whether to forage or rest, explore or exploit. These trade-offs are governed by the current metabolic state of the body, which cannot be inferred from the cost of neural activity alone. Hunger or satiation systematically shape value computation, action selection, and execution (Juechems & Summerfield, 2019; Keramati & Gutkin, 2014; Yoshida et al., 2024). For example, animals that have learned to avoid a salt-associated cue can immediately reverse this preference when sodium-depleted, approaching the stimulus without any new learning (Robinson & Berridge, 2013). While the stimulus remains unchanged, the organism's internal state has entirely restructured its value and the cost–benefit logic of the learned associations. Similarly, metabolism dynamically regulates the energetic investment of action itself. Animals invest additional energy to move faster only when the anticipated benefit justifies the metabolic cost (Shadmehr & Ahmed, 2020; Starita et al., 2022; Summerside, Shadmehr, & Ahmed, 2018). This calculation shifts dynamically with the organism's physiological state, such that, in conditions of metabolic depletion or threat, the neural and physical cost of a vigorous action is discounted by the gain of physiological restoration (Garofalo et al., 2024; Keramati & Gutkin, 2014). These examples suggest that the computations underlying decision-making are organized in the service of the metabolic needs of the body rather than being merely constrained by it. Considering only the energetic costs of neural computations risks mischaracterizing the function and organization of cognition, as it may overlook how the fluctuating metabolic needs of the body define the very purpose and objective of the neural computation.

Our framework also reshapes the comparison between biological and artificial systems. The relevant question is no longer whether artificial systems match the energy constraint of the brain, but whether their architectures incorporate analogous principles of

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

whole-body metabolic regulation. Standard computational models of reinforcement learning generally treat goals and values as externally fixed, with rewards and costs specified in advance. However, the above empirical evidence shows that this is not how learning and decision-making operate in biological agents. Homeostatic reinforcement learning provides a formal computational model that more closely mimic biological cognition, replacing abstract reward maximization with (metabolic) drive reduction (Keramati & Gutkin, 2014). In such models, the value of an action is dynamically scaled by the urgency of a physiological deficit. These systems show that *intelligent*, adaptive behavior can emerge without externally specified rewards, for example, by learning to forage for nutrients solely by monitoring and optimizing the maintenance of internal metabolic set points (Yoshida et al., 2024). Embodied artificial systems thus show that metabolism can function as a genuinely generative architectural foundation also for artificial cognition.

In conclusion, Haueis and Colaço have made a compelling case that metabolism belongs at the center of cognitive modeling. We have argued that this center cannot hold at the level of the neuron or the brain alone; it must expand to encompass the organism as a whole, the body whose survival imperatives give neural computation and cognition both its constraints and its meaning.

Financial support. This work was partially supported by the European Union under the Italian National Recovery and Resilience Plan (NRRP) of NextGenerationEU, partnership (Mission 4 – Component 2 – Investment 1.1) PRIN 2022 (No. 104, 2/2/2022, CUP J53C24002790006).

Competing interests. Competing interests: The authors declare none.

References

Barrett, L. F., & Simmons, W. K. (2015). Interoceptive predictions in the brain. *Nature Reviews Neuroscience*, *16*(7), 419–429. https://doi.org/10.1038/nrn3950.

Chiel, H. J., & Beer, R. D. (1997). The brain has a body: Adaptive behavior emerges from interactions of nervous system, body and environment. *Trends in Neurosciences*, *20*(12), 553–557. https://doi.org/10.1016/S0166-2236(97)01149-1.

Cisek, P. (2019). Resynthesizing behavior through phylogenetic refinement. *Attention, Perception, and Psychophysics*, *81*(7), 2265–2287. https://doi.org/10.3758/s13414-019-01760-1.

Cisek, P. (2022). Evolution of behavioural control from chordates to primates. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *377*(1844), Article no. 20200522. https://doi.org/10.1098/rstb.2020.0522.

Garofalo, S., Finotti, G., Starita, F., Bouchard, A., & Fecteau, S. (2024). Effort-based decision making. In G. J. Boyle, G. Northoff, A. K. Barbey, F. Fregni, M. Jahanshahi, A. Pascual-Leone, & B. J. Sahakian (Eds.), *The SAGE handbook of cognitive and systems neuroscience*. SAGE Publications Ltd. https://doi.org/10.4135/9781529616651.n13.

Hochner, B., Zullo, L., Shomrat, T., Levy, G., & Nesher, N. (2023). Embodied mechanisms of motor control in the octopus. *Current Biology*, *33*(20), R1119–R1125. https://doi.org/10.1016/J.CUB.2023.09.008.

Juechems, K., & Summerfield, C. (2019). Where does value come from? *Trends in Cognitive Sciences*, *23*(10), 836–850. https://doi.org/10.1016/j.tics.2019.07.012.

Keramati, M., & Gutkin, B. (2014). Homeostatic reinforcement learning for integrating reward collection and physiological stability. *ELife*, *3*, Article no. e04811. https://doi.org/10.7554/ELIFE.04811.

Robinson, M. J. F., & Berridge, K. C. (2013). Instant transformation of learned repulsion into motivational "wanting". *Current Biology*, *23*(4), 282–289. https://doi.org/10.1016/J.CUB.2013.01.016.

Shadmehr, R., & Ahmed, A. A. (2020). *Vigor*. The MIT Press. https://doi.org/10.7551/mitpress/12940.001.0001.

Starita, F., Garofalo, S., Dalbagno, D., Degni, L. A. E., & di Pellegrino, G. (2022). Pavlovian threat learning shapes the kinematics of action. *Frontiers in Psychology*, *13*, 6336. https://doi.org/10.3389/fpsyg.2022.1005656.

Summerside, E. M., Shadmehr, R., & Ahmed, A. A. (2018). Vigor of reaching movements: Reward discounts the cost of effort. *Journal of Neurophysiology*, *119*(6), 2347–2357. https://doi.org/10.1152/JN.00872.2017.

Yoshida, N., Daikoku, T., Nagai, Y., & Kuniyoshi, Y. (2024). Emergence of integrated behaviors through direct optimization for homeostasis. *Neural Networks*, *177*, 106379. https://doi.org/10.1016/J.NEUNET.2024.106379.

Metabolism and the mind–body problem

Paul Thagard*

University of Waterloo, Canada
pthagard@uwaterloo.ca

*Corresponding author.

doi:10.1017/S0140525X26106116, e480

Abstract

The authors argue convincingly that metabolic considerations are highly relevant to evaluating and generating models of cognition. However, they understate their relevance to the mind–body problem. Metabolic reasoning about energy consumption should be combined with other constraints of speed, space, and history, providing an alternative to multiple realizability and medium independence: coherent materialism.

Philipp Haueis and David Colaço make an excellent case for their main conclusion that metabolic considerations are highly relevant to evaluating and generating models of cognition. However, they have understated the relevance of their analysis for the central philosophical question of the relation between mind and matter. I will describe how metabolic reasoning about energy consumption should be combined with other constraints of speed, space, and history. The result provides an alternative to multiple realizability and medium independence: coherent materialism.

The key question concerns the relation between mental processes and physical processes. The mind–brain identity theory says that all mental processes, such as problem solving, are brain processes, but multiple realizability claims that the mind could be realized by a wide range of physical processes, such as digital computers and space aliens. Medium independence narrows the range by allowing that some physical media may be relevant to restrictions on realization but is vague about what these restrictions might be. Haueis and Colaço astutely notice that metabolism is highly relevant to restricting the range of possible realizations, because mental systems operating in the real world must acquire and use energy economically and efficiently to enable them to operate in their environments.

Metabolism, however, is not the only constraint on cognitive effectiveness in real-world situations. One advantage of neural networks is their massive parallelism, which allows them to perform many computations simultaneously. The brains of humans and other organisms use this parallelism to work rapidly in the

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

world in pursuits where speed matters, in finding food, avoiding predators, and acquiring mates. Speed is a property of physical systems, as we see in the contrast between relatively slow neuronal firing and incredibly fast (trillions of times per second) operation of silicon chips.

The constraint of speed interacts with the metabolic constraint of energy consumption. If greater speed comes at the cost of energy expenditure, then it may not improve an entity's ability to survive and reproduce in a relatively non-threatening world.

A third restriction on the physical realization of mental processes is space, because thinking must be an appropriate size to operate in their environments. Human brains could not be much bigger because they initially grow in and are birthed from mothers. Realizability and independence are therefore limited by the size of the thinking substrate. The data centers that run current artificial intelligence models are enormous, covering multiple square miles, but they cannot be unlimited because of the availability of land. The constraints of metabolism and space interact because large systems require more energy to enable them to operate.

The fourth constraint on physical realization is history, which operates because both biological evolution and technical design cannot pull new arrangements out of the air but must work with prior physical properties. Biology never managed to produce brains capable of extrasensory perception, and computer design also has to build on top of currently available technologies. Metabolism interacts with history because any new entity, evolved or designed, is limited by the energy capabilities of its predecessors.

These four constraints show the implausibility of multiple realizability and make even medium independence suspect. But we should not swing back to mind–brain identity that assumes zero independence, not because of fanciful thought experiments, but because generative AI has brought computers capable of many intelligent operations, including sophisticated language use, problem solving, learning, and many kinds of reasoning. The rise of powerful AI beginning with ChatGPT in 2022 shows that minds are realizable in special computers as well as brains.

Consideration of these four interacting constraints might be described as scant realizability or minimal independence, but I call the resulting view of mental-physical arrangements *coherent materialism* (Thagard, 2025, in press). Here coherence is not just the vague idea of fitting together, but rather the precise idea, computable by neural networks and other algorithms, that coherence is satisfaction of multiple conflicting constraints (Thagard, 2000). According to coherent materialism, only rare physical processes will sufficiently satisfy constraints of energy, time, space, and history in their environments to achieve mental accomplishments such as intelligent problem solving. At the moment, we have only two realizations of such minds that we know about, using biological brains and engineered computer chips using GPU and similar processors. Others may turn up through technological innovation or biological evolution on other planets, but constraints that include metabolism show that we should not expect many ways of physical production of minds.

The new AI systems illustrate the interactions of the four constraints of energy, time, space, and history, sharply in contrast with how humans satisfy these constraints. Compared to humans, AI data centers are massive energy hogs, requiring hundreds of megawatts to operate compared to the 20 watts of human brains. On the other hand, they are capable of enormous speed, because brains are limited by the relative slowness of neural firing and neurotransmitter propagation. AI data centers are becoming ridiculously large, but they can be built that way without having to worry about being born or carried around. AI models are free of the historical constraints of having to evolve slowly through genetic inheritance but are becoming capable of major leaps in part because they can help to design themselves, as happens when programming models made by OpenAI and Anthropic contribute to their next generation.

The new AI and considerations of the metabolism of brains together point the way to a new solution to the mind–body problem, coherent materialism. The energy constraint of metabolism should be viewed as interacting with the additional constraints of time, space, and history to greatly restrict realizability and independence. Hence what Haueis and Colaço call the "metabolic turn" in philosophy is even more consequential than they surmise.

Acknowledgements. None.

Financial support. This project has no funding.

Competing interests. The author has no competing interests.

References

Thagard, P. (2000). *Coherence in thought and action*. MIT Press.
Thagard, P. (2025). *Dreams, jokes, and songs: How brains build consciousness*. Oxford University Press.
Thagard, P. (in press). *Boom or doom? Philosophy and psychology of the new artificial intelligence*. MIT Press.

Relating metabolic constraints with human-level cognitive architecture

Frank van der Velde*

Department of Learning, Data-Analytics and Technology, Section Cognition, Data and Education, Faculty of Behavioral, Management and Social Sciences, University of Twente, Netherlands
veldefvander@outlook.com

*Corresponding author.

doi:10.1017/S0140525X26106335, e481

Abstract

Metabolic constraints are important for brain modeling. But to be useful, they have to be integrated with an understanding of the architecture of human-level cognition. Otherwise, they could even lead modeling in the wrong direction.

Haueis and Colaço correctly address the importance of metabolic constraints on brain processing and brain modeling, and illustrate this with a number of valuable examples.

However, their target article does not appropriately address the influence of metabolic constraints on the nature of human cognition, even though that would be one of its most important topics. In the evolution of human cognition, there would have been an interaction between metabolic constraints and the development of cognitive abilities as required by human-level cognition.

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

A well-known view is that human-level cognition depends on compositional processing, and the computational architecture that could achieve this is the Von Neumann Architecture (VNA) or variants thereof (Fodor & Pylyshyn, 1988; Newell, 1990). But the VNA is highly energy inefficient. The target article indeed refers to this and suggests that analog AI chips could solve this energy problem. But this solution throws away the baby with the bathwater. To see why, we have to look a bit at the structure of the VNA, both from an energy and a cognitive perspective.

Basically, the VNA consists of a central processing unit (CPU) connected to a data-bus, which is connected to memory, e.g., given by registers. These are shielded off from the data-bus; otherwise, they would interfere with processing. In the execution of a program, they can be opened by the CPU using "address" codes. In this way, data from memory can be transported to the CPU for processing, or data from the CPU can be transported to memory. This creates the well-known "bottleneck" of the VNA, because at any given time only one of these operations can be carried out (Backus, 1978).

The structure of the VNA is important because it relates directly to its energy consumption. Computations in the CPU do not consume most of its energy. Instead, the main energy consumption derives from data transfer over the data-bus (Hess, 2025). This begs the question of whether this data transfer is actually needed. The VNA uses it because it is designed as a universally programmable computer. Hence, it also uses it for numerical computations as found in vector and matrix operations. Here, indeed, energy will be reduced by using, e.g., analog AI chips (Mehonic & Kenyon, 2022).

But these processors do not (yet) provide the compositional processing underlying human cognition, as exemplified in language. For example, a sentence is composed out of words. The sentence structure expresses specific relations between these words, which can be retrieved with relational or "who does what to whom" questions. As noted, a standard view on cognition is that we represent and process sentences in this way as well. An important argument for this was given by Newell (1990), based on the notion that the amount of information stored at any given location in any architecture is limited due to physical reasons.

Hence, information needed for more complex processing will be distributed over an architecture. An example is given by the observation that information related to words is distributed over large parts of the cortex (Huth et al., 2016). Yet, words can be used on the fly to compose and process arbitrary sentences. According to Newell, this requires an architecture that has access to distant and distributed information. Access then allows the transfer of distant information to integrate it with local processing. Much in the same way as the VNA transfers data to process it in the CPU. Yet, the access and data transfer underlying compositional processing would increase energy consumption, as it does in the VNA.

But perhaps compositional processing would not be needed for higher-level cognitive processing. An alternative could be given by large language models (LLMs). They do not represent and process sentences in a compositional way but are trained directly on a huge set of sentences. However, they need to be trained on more sentences than humans will see in their lifetime (Linzen, 2020). The energy costs of this are even more staggering (Ji & Jiang, 2026). Perhaps this could be reduced with other kinds of processors (Mehonic & Kenyon, 2022). But given the number of sentences required for training, it is questionable whether this would result in significant energy reductions.

Moreover, LLMs exhibit a number of problems from a cognitive perspective. For example, they do not reliably answer relational questions, addressing the structure of a sentence, even for sentences that they have learned (Berglund et al., 2024). In contrast, children learn language with a minute subset of sentences compared to LLMs (Saxton, 2010), and they can answer relational questions because they learn and process language in a compositional manner (Marcus et al., 1999).

Hence, on the one hand, compositional processing reduces energy consumption because it does not require learning, e.g., massive numbers of sentences. But on the other hand, the transfer of information needed for compositional processing increases energy consumption in architectures such as the VNA.

This contrast requires the development of cognitive architectures that are capable of compositional processing in an energy-reduced manner. Promising examples are based on forms of neurosymbolic processing (Hitzler, Sarker, & Eberhart, 2023), or processing akin to that. They could achieve compositional processing in a neural manner. But the energy consumption of information transfer needs to be reduced as well.

A solution for this problem is access to distant information without having to transfer it for processing, as with distributed information that is interconnected in a small-world network structure (Watts & Strogatz, 1998). The connection paths in this structure allow access to distant information and can be used to interconnect local processing at different sites, e.g., by shifting attention from one site to another (van der Velde, 2023). This would reduce energy consumption because it does not require information transfer.

To summarize, the focus on metabolic constraints is important. But on its own it is not sufficient and could even lead modeling in the wrong direction. To be useful, it has to be combined with a clear understanding of its effects on human-level cognitive architecture.

Financial support. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Open access funding provided by University of Twente.

Competing interests. Author declares no competing interests.

References

Backus, J. (1978). Can programming be liberated from the von neumann style? A functional style and its algebra of programs. 1977 ACM turing award lecture. *Communications of the ACM*, *2*, 613–641. https://dl.acm.org/doi/pdf/10.1145/359576.359579

Berglund, L., Tong, M., Kaufmann, M., Balesni, M., Stickland, A., Korbak, T., & Evans, O. (2024). The Reversal Curse: LLMs trained on "A is B" fail to learn "B is A". In *Proceedings of the Twelfth International Conference on Learning Representations* (ICLR. 2024).

Fodor, J. A., & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, *28*, 3–71. https://doi.org/10.1016/0010-0277(88)90031-5

Hess, P. (2025). Why a decades old architecture decision is impeding the power of AI computing. *IBM Research*. Accessed 5 April, 2026. https://research.ibm.com/blog/why-von-neumann-architecture-is-impeding-the-power-of-ai-computing

Hitzler, P., Sarker, M. K., & Eberhart, A. (2023). *Compendium of neurosymbolic artificial intelligence*. IOS Press, pp. 249–271.

Huth, A. G., de Heer, W. A., Griffiths, T. L., Theunissen, F. E., & Gallant, J. L. (2016). Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature*, *532* (7600), 453–458. https://doi.org/10.1038/nature17637

Ji, Z., & Jiang, M. (2026). A systematic review of electricity demand for large language models: Evaluations, challenges, and solutions. *Renewable and*

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

Sustainable Energy Reviews, *225*, 116159. https://doi.org/10.1016/j.rser.2025.116159
Linzen, T. (2020). How can we accelerate progress towards human-like linguistic generalization? In D. Jurafsky, J. Chai, N. Schluter, & J. Tetreault, (Eds.), *Proceedings of the 58th annual meeting of the association for computational linguistics* (pp. 5210–5217). Association for Computational Linguistics. https://doi.org/10.18653/v1/2020.acl-main.465
Marcus, G. F., Vijayan, S., Bandi Rao, S., & Vishton, P. M. (1999). Rule learning in seven-month- old infants. *Science*, *283*, 77–80. https://doi.org/10.1126/science.283.5398.77
Mehonic, A., & Kenyon, A. J. (2022). Brain-inspired computing needs a master plan. *Nature*, *604*, 255–260. https://doi.org/10.1038/s41586-021-04362-w
Newell, A. (1990). *Unified theories of cognition*. Harvard University Press.
Saxton, M. (2010). *Child language*. Sage.
Van der Velde, F. (2023). The neural blackboard theory of neuro-symbolic processing: Logistics of access, connection paths and intrinsic structures. In P. Hitzler, M. K. Sarker, & A. Eberhart (Eds.), *Compendium of neurosymbolic artificial intelligence* (pp. 249–271). IOS Press.
Watts D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, *393*, 440–442. https://doi.org/10.1038/30918

Not every model in cognitive science is (or should be) mechanistic

Danielle J. Williams*

Philosophy, Purdue University, West Lafayette, USA
danielle.williams@purdue.edu

*Corresponding author.

doi:10.1017/S0140525X26106104, e482

Abstract

Haueis and Colaço (H&C) label all models as cognitive. However, just because cognitive science is so-called doesn't mean that all models are. When type differences are ignored, an argument runs the risk of prescribing a monist modeling strategy. In their case: mechanistic. Such a prescription is empirically inadequate and overly restrictive, as modeling in cognitive science relies on many modeling types.

H&C label all models in cognitive science as "cognitive," as indicated by their blanket use of the term "cognitive model." Because of this labeling scheme, they are free to explore and rely on any modeling strategy to make their case. However, just because cognitive science is so-called does not mean that all models are cognitive models in particular and that every model should be subject to their requirement in general. This is more than a semantic complaint. Distinguishing model types matters because any view that categorizes all models as a single type will fail to capture the complexity of explanation in the mind sciences (and how those models may work toegether). Moreover, when type differences are ignored, the argument runs the risk of prescribing a monist modeling strategy that is tethered to a single explanatory framework. In H&C's case: mechanistic. Such an approach is empirically inadequate and overly restrictive as cognitive science relies on a diverse set of modeling types to triangluate on the phenomenon of interest.

Cognitive models is a particular model type that takes a cognitive system as their intended target where the features included purport to characterize the system in a way that captures its cognitive *functioning*. These models describe a system in terms of representations, processes, and resources (Samuels, 2018; Weiskopf, 2017). Cognitive models can be contrasted with mechanistic models, which represent a system in terms of its organization, demonstrating how component activities produce the phenomena of interest (Kaplan & Craver, 2011). Canonical mechanistic models explain in terms of organized activities of and interactions among components, i.e., its "parts" (Craver, 2006), while cognitive models focus on a functional explanation in terms of the process the system performs and/or what the mechanism may represent.

An important difference between cognitive models and mechanistic models is that the mechanistic strategy depends on the model making determinate claims about the localization of components (Weiskopf, 2017). Cognitive models don't necessarily do this. Instead, they often lack commitments to particular spatial organization, even if the models include diagrams or texts that rely on spatial relations (Weiskopf, 2011). This difference is critical when it comes to H&C's proposal that metabolic considerations should be integrated into *all* models. The idea that all models must incorporate metabolism inevitably requires a particular type of description and focus that shapes the model into a form of mechanistic explanation. This consequence is demonstrated by their critique of Kurzban et al., as we will see.

H&C introduce metabolic considerations as a normative constraint on modeling where explanatory adequacy comes, in part, from whether it includes those details. To demonstrate the reasonableness of their proposal, the authors provide two examples of successful inclusion of metabolic considerations and then argue that the strategy can be generalized to *all* models.

The first example is the cyclical model of synaptic memory consolidation (Karbowski, 2019). The cyclical nature of this model accounts for the energy exchange between the system and its environment. Karbowski is interested in the molecular events in dendritic spines and is explicitly characterizing the model in terms of a chemical mechanism. The second example (Jaras et al., 2021) describes the leaky integrate-and-fire (LIF) model of the action potential, but with a reference to metabolic features. The resulting energy-dependent LIF model (EDLIF) captures the differences that ATP makes to the reset value of the membrane potential.

Both the cyclical model and the EDLIF are clear-cut examples of a mechanistic modeling strategy that provides a mechanistic explanation. Given this, it is no surprise that the models easily integrate metabolic aspects. After all, mechanistic models can lie anywhere on a continuum between a mechanism sketch and an ideally complete description (Craver, 2006). What H&C show is that you can *always* fill in more details.

Contrast these examples with H&C's critique of Kurzban et al., who are interested in understanding the *phenomenology* of mental effort. Their model consists of a narrative of the problem

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

space along with an abstract description of how cognitive demand requires a cost/benefit analysis. This model (given as a description) is meant to explain the *experiences* of boredom and mental effort associated with vigilance tasks. As a bonafide cognitive model, it explains mental effort in terms of representational processes *without* localizing the posited mechanism within the brain – it provides a *functional* description. In Marrian terms, their strategy is deliberately silent on implementation details. However, this does not mean that they ignore them entirely.

In their artictle, H&C specifically take issue with a statement about the model given by the authors in another BBS target article. In a reply to H&C's commentary, Kurzban et al. state that they agree that the brain is constrained by biophysical thermodynamic laws but disagree that such factors should be incorporated into *their* explanation (Kurzban et al., 2013, p. 714). H&C argue that this is a problem because *all* models should explicitly reference metabolic features (their normative constraint). Because Kurzban et al. refuse to acquiesce, their model must be rejected. But notice what's really going on. H&C actually reject the explanatory *framework* of cognitive modeling in favor of a mechanistic-style explanation. Indeed, incorporating metabolic details would transform the Kurzban model into a kind of cognitive–mechanistic hybrid model, if not mechanistic entirely. If this is the strategy that H&C want to take for all models, then cognitive models will be slowly picked off and either replaced or transformed into mechanistic models.

Properly distinguishing model types reveal that H&C's metabolic constraint fails to capture the diverse modeling strategies operating within cognitive science. Once modeling types are made explicit, it becomes difficult to justify a metabolic constraint on *all* models. If, however, one insists on restricting model types based on the inclusion of metabolic constraints, it's worth considering whether all there is to cognitive science is mechanistic explanation. My take is that there's more to it than that (Williams, 2023; Williams, 2025).

Financial support. This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

Competing interests. The author declares no competing interests.

References

Craver, C. F. (2006). When mechanistic models explain. *Synthese 153*, 355–376.

Jaras, I., Harada, T., Orchard, M. E., Maldonado, P. E., & Vergara, R. C. (2021). Extending the integrate-and-fire model to account for metabolic dependencies. *European Journal of Neuroscience*, *54*(4), 5249–5260.

Kaplan, D. M., & Craver, C. F. (2011). The explanatory force of dynamical and mathematical models in neuroscience: A mechanistic perspective. *Philosophy of Science*, *78*(4), 601–627.

Karbowski, J. (2019). Metabolic constraints on synaptic learning and memory. *Journal of Neurophysiology*, *122*(4), 1473–1490.

Kurzban, R., Duckworth, A., Kable, J. W., & Myers, J. (2013). An opportunity cost model of subjective effort and task performance. *Behavioral Brain Sciences*, *36*(6), 661–679.

Samuels, R. (2018). Classical computational models. In Sprevak & Columbo (Eds.), *The routledge handbook of the computational mind* (pp. 103-119). Routledge.

Weiskopf, D. A. (2011). Models and mechanisms in psychological explanation. *Synthese*, *183*, 313–338.

Weiskopf, D. A. (2017). The explanatory autonomy of cognitive models. In D. M. Kaplan (Ed.), *Explanation and integration in mind and brain sciences* (pp. 44–69). Oxford University Press.

Williams, D. J. (2023). Implementation and Interpretation: A Unified Account of Physical Computation. Available from ProQuest Dissertations & Theses Global. (2865991732).

Williams, D. J. (2025). How to be a realist about computational neuroscience. *Synthese*, *205*(102), 1–27.

Extending metabolic considerations to the default mode network

Yi-Sheng Wong*

Institute for Human Development and Potential, Agency for Science Technology and Research, Singapore
yishengwong.ysw@gmail.com

*Corresponding author.

doi:10.1017/S0140525X26106074, e483

Abstract

Haueis and Colaço distinguish evaluative and generative roles for metabolic considerations in cognitive modeling. I extend this framework to the default mode network, arguing that this network's elevated aerobic glycolysis constitutes a metabolic constraint on spontaneous thought. This constraint shapes the richness of spontaneous thought across varying task demands and predicts load-dependent phenomenological shifts through inter-network resource competition.

Haueis and Colaço argue that metabolic considerations should play both evaluative and generative roles in cognitive modeling. While their framework is compelling, their examples are drawn almost exclusively from the neuronal and synaptic levels. This leaves open whether the framework extends to large-scale brain networks that support complex cognition.

Here, I make a strong case that it does. Focusing on the default mode network and spontaneous thought, I show that metabolic specializations at the network level are capable of both pruning existing models and generating new, testable predictions. This extends the evaluative (Section 3.1) and generative (Section 3.2) framework into a domain where metabolic constraints are especially pronounced but remain largely untheorized.

The default mode network's metabolic specialization

Core regions of the default mode network rank among the highest in aerobic glycolysis across the entire human brain (Vaishnavi et al., 2010). Despite producing less ATP per glucose than oxidative phosphorylation, aerobic glycolysis affords rapid and flexible energy availability (Magistretti & Allaman, 2015). Within the efficiency taxonomy that Haueis and Colaço outline (Section 2.1.2), this metabolic profile reflects a trade-off: the network sacrifices thermodynamic efficiency for temporal responsiveness. Notably, the same regions also show neotenous gene expressions linked to synaptic growth and plasticity (Goyal et al., 2014), implying a state kept metabolically pliable well beyond typical developmental windows. This pliability is well-suited to the dynamic recombination of information that characterizes

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

spontaneous thought: autobiographical memory, future thinking, social cognition, and creative thought (Christoff et al., 2016).

Evaluative application: Pruning models of mind wandering

The executive failure hypothesis (Kane & McVay, 2012; McVay & Kane, 2010), for instance, proposes that mind wandering (i.e., a member of a family of spontaneous-thought phenomena; Christoff et al., 2016) arises primarily from failures of top-down control. Under this account, the energetic cost of thought content plays little explicit role. Yet empirical evidence suggests that patterns of default mode network activity during working-memory maintenance track the phenomenological richness of ongoing experience (Sormaz et al., 2018). In other words, this network continues to support spontaneous thought even when attentional resources are committed elsewhere. This finding is hard to square with an account that explains mind wandering solely in terms of failures of attentional control.

Metabolic data help explain this pattern. Using simultaneous PET-fMRI, Stiernman et al. (2021) demonstrated that glucose metabolism in the default mode network remains elevated even when blood-oxygen-level-dependent signals decrease – a dissociation that conventional neuroimaging alone would miss. Godbersen et al. (2023) revealed further that the default mode network's metabolic activity varies depending on which task-positive networks are active. Collectively, these results indicate that what determines the character of spontaneous thought is competition for bioenergetic resources across networks, not just the strength of executive suppression. In the terms of Haueis and Colaço, metabolism here functions as an explanatory constraint: it structures which cognitive forms can be sustained under a given task load, without itself initiating those processes. The executive failure account, which ignores this dimension, is thereby prunable.

Generative application: Predicting phenomenology

The default mode network's metabolic profile also generates specific, testable predictions. First, as external task demands increase, spontaneous thought should become progressively less vivid, shifting from richly detailed, multisensory simulations toward more schematic mentation. This should correlate with reductions in aerobic glycolysis within core default mode network regions specifically, over and above changes in frontoparietal activation (Godbersen et al., 2023; Sormaz et al., 2018). Second, individuals with higher resting aerobic glycolysis in the default mode network should sustain more elaborate spontaneous thought under high cognitive load. More broadly, stable differences between people in the richness, sensory specificity, and temporal extension of their spontaneous thought should partly reflect metabolic architecture rather than attentional control alone. Recent work linking resting-state thought contents to resting-state functional connectivity (Yu, Wong, & Billaud, 2025) offers early support for this point, suggesting that thought patterns may partly mirror underlying metabolic organization. Taken together, these predictions illustrate how metabolic knowledge, joined with network-level neuroscience, can populate the model space with novel, testable hypotheses.

Conclusion

The default mode network case shows that Haueis and Colaço's framework scales to large-scale cognitive systems. Elevated aerobic glycolysis in this network functions as a domain-specific constraint, shaping the phenomenological texture of thought (Raichle, 2015; Vaishnavi et al., 2010). Rather than treating mind wandering as executive failure (Kane & McVay, 2012), a metabolically grounded view reframes spontaneous thought as bioenergetically constrained cognition, sculpted by competition for finite metabolic resources across interacting large-scale networks (Godbersen et al., 2023; Stiernman et al., 2021). This perspective complements, rather than replaces, executive-control (Kane & McVay, 2012) and switching-based theories (Wong, Willoughby, & Machado, 2023), suggesting that spontaneous thought is shaped jointly by attentional control, cognitive flexibility, and metabolic supply. Integrating such constraints into cognitive models broadens the explanatory scope of metabolic approaches and reinforces the value of integrating metabolic considerations across every level of analysis, from individual synapses up to whole-brain networks.

Acknowledgements. During the preparation of this work, the author used ChatGPT to assist with drafting. The author reviewed, edited, and take full responsibility for the final content.

Financial support. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Competing interests. None.

References

Christoff, K., Irving, Z. C., Fox, K. C. R., Spreng, R. N., & Andrews-Hanna, J. R. (2016). Mind-wandering as spontaneous thought: A dynamic framework. *Nature Reviews Neuroscience*, *17*(11), 718–731. https://doi.org/10.1038/nrn.2016.113

Godbersen, G. M., Klug, S., Wadsak, W., Pichler, V., Raitanen, J., Rieckmann, A., Stiernman, L., Cocchi, L., Breakspear, M., Hacker, M., Lanzenberger, R., & Hahn, A. (2023). Task-evoked metabolic demands of the posteromedial default mode network are shaped by dorsal attention and frontoparietal control networks. *eLife*, *12*, e84683. https://doi.org/10.7554/eLife.84683

Goyal, M. S., Hawrylycz, M., Miller, J. A., Snyder, A. Z., & Raichle, M. E. (2014). Aerobic glycolysis in the human brain is associated with development and neotenous gene expression. *Cell Metabolism*, *19*(1), 49–57. https://doi.org/10.1016/j.cmet.2013.11.020

Kane, M. J., & McVay, J. C. (2012). What mind wandering reveals about executive-control abilities and failures. *Current Directions in Psychological Science*, *21*(5), 348–354. https://doi.org/10.1177/0963721412454875

Magistretti, P. J., & Allaman, I. (2015). A cellular perspective on brain energy metabolism and functional imaging. *Neuron*, *86*(4), 883–901. https://doi.org/10.1016/j.neuron.2015.03.035

McVay, J. C., & Kane, M. J. (2010). Does mind wandering reflect executive function or executive failure? Comment on Smallwood and Schooler (2006) and Watkins (2008). *Psychological Bulletin*, *136*(2), 188–197. https://doi.org/10.1037/a0018298

Raichle, M. E. (2015). The brain's default mode network. *Annual Review of Neuroscience*, *38*(1), 433–447. https://doi.org/10.1146/annurev-neuro-071013-014030

Sormaz, M., Murphy, C., Wang, H.-T., Hymers, M., Karapanagiotidis, T., Poerio, G., Margulies, D. S., Jefferies, E., & Smallwood, J. (2018). Default mode network can support the level of detail in experience during active task states. *Proceedings of the National Academy of Sciences*, *115*(37), 9318–9323. https://doi.org/10.1073/pnas.1721259115

Stiernman, L. J., Grill, F., Hahn, A., Rischka, L., Lanzenberger, R., Panes Lundmark, V., Riklund, K., Axelsson, J., & Rieckmann, A. (2021). Dissociations between glucose metabolism and blood oxygenation in the human default mode network revealed by simultaneous PET-fMRI. *Proceedings of the National Academy of Sciences*, *118*(27), e2021913118. https://doi.org/10.1073/pnas.2021913118

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

Vaishnavi, S. N., Vlassenko, A. G., Rundle, M. M., Snyder, A. Z., Mintun, M. A., & Raichle, M. E. (2010). Regional aerobic glycolysis in the human brain. *Proceedings of the National Academy of Sciences*, *107*(41), 17757–17762. https://doi.org/10.1073/pnas.1010459107

Wong, Y.-S., Willoughby, A. R., & Machado, L. (2023). Reconceptualizing mind wandering from a switching perspective. *Psychological Research*, *87*, 357–372. https://doi.org/10.1007/s00426-022-01676-w

Yu, J., Wong, Y.-S., & Billaud, C. H. A. (2025). Thought contents during rest account for functional connectivity-behavior associations. *Imaging Neuroscience*, *3*, IMAG.a.55. https://doi.org/10.1162/IMAG.a.55

Metabolic constraints on neural fatigue: Evidence from passive induction paradigms

Alexandre Zénon[a]* and Oleg Solopchuk[b]

[a]Univ. Bordeaux, CNRS, INCIA, UMR 5287, F-33000 Bordeaux, France and [b]Department of Systems Neuroscience, Universitätsklinikum Hamburg-Eppendorf: Universitatsklinikum Hamburg-Eppendorf, Hamburg, Germany
alexandre.zenon@u-bordeaux.fr
o.solopchuk@uke.de

*Corresponding author.

doi:10.1017/S0140525X26106086, e484

Abstract

Haueis and Colaço argue that metabolic considerations should constrain and generate cognitive models. Research in cognitive fatigue provides a test case. Passive fatigue induction reveals specific representational degradation that points to local rather than global metabolic mechanisms, pruning potential models. We argue that local aerobic glycolysis offers a coherent, testable generative account of these findings.

Haueis and Colaço make a compelling case that metabolic considerations remain systematically underused in cognitive modeling, functioning as background assumptions rather than active constraints or generative tools. We complement their treatment of mental effort in section 4.1 with considerations from the cognitive fatigue literature that illustrate both the evaluative and generative roles their framework identifies.

A persistent methodological obstacle in this literature is that standard fatigue paradigms confound fatigue mechanisms with motivation, sleepiness, boredom, and practice effects. Most studies use sustained effortful performance to induce fatigue in participants, making it difficult to isolate the contribution of local metabolic alterations from other behavioral and computational factors. This confound has probably contributed to the stalemate between resource and computational accounts, mentioned in the review.

Ioannucci et al. (2023) addressed this issue by inducing fatigue passively through prolonged repetitive visual stimulation of a single quadrant of the visual field. Performance on a subsequent texture discrimination task was impaired specifically for stimuli in the stimulated quadrant, and multivariate pattern analysis of fMRI data revealed degraded neural representations of that region in the visual cortex. Adjacent, unstimulated regions were unaffected. This spatial specificity instantiates the evaluative role of metabolic considerations in the sense of guidelines E4 and E5: it prunes the space of possible models by ruling out any account based on global resource depletion or prefrontal-specific mechanisms, and it constrains the locus of the relevant metabolic process to the activation history of a circumscribed neural assembly. The result does not identify a mechanism but narrows the space of possible hypotheses considerably.

Despite this constraint, there is currently no consensus on the local metabolic mechanism underlying such representational degradation. Recent data suggest glutamate accumulation could be a prominent candidate. Wiehler et al. (2022) reported elevated glutamate concentrations in the lateral prefrontal cortex following a day of cognitively demanding work. They proposed that glutamate accumulation signals metabolic saturation and down-regulates neural processing. However, the reported effect was driven primarily by concentration decreases in the resting control group rather than increases in the fatigued group, which is the opposite of what an accumulation account would predict. Moreover, it was restricted to frontal regions, providing no generalizable mechanism for assembly-level representational failure. Therefore, this metabolic account does not satisfy guideline E2 (specifying the relevant process with sufficient precision) in this case.

We propose that the glycolytic account of Theriault et al. (2023), the core generative example the authors discuss in the review, offers an alternative starting point. Thin, information-energy-efficient axons, present in large proportion in the visual cortex, do not contain mitochondria and thus depend on aerobic glycolysis. Since sustained capacity of aerobic glycolysis is constrained by the renewal rate of NAD^+ and lactate clearance (Dienel, 2019), prolonged or repeated recruitment of the same neuronal assembly could progressively deplete glycolytic throughput below the level required to maintain precise signaling. This mechanism is not prefrontal-specific: it applies in principle to any assembly whose axonal composition favors thin fibers, making it consistent with the spatial specificity of the Ioannucci et al. findings. This constitutes a generative metabolic consideration in the sense of guidelines G2 and G4: it populates the space of possible models with a candidate mechanism grounded in known enzymatic constraints. More importantly, it generates testable, quantitative predictions: the degree of representational degradation (indexed by MVPA classifier accuracy) should scale with the local proportion of thin axons or with the proportion of aerobic glycolysis in the stimulated region. Closing this loop, by combining passive induction designs with spectroscopic measurement of local lactate dynamics as a proxy for glycolytic activity, would instantiate the virtuous cycle of generative and evaluative modeling that Haueis and Colaço depict in their Figure 7.

Financial support. AZ received funding from the ANR grant COURRIER ANR-24-CE10-5680. O.S. was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – RU5389, project number 461947532.

Competing interests. None.

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

References

Dienel, G. A. (2019). Brain glucose metabolism: Integration of energetics with function. *Physiological Reviews*, *99*(1), 949–1045. https://doi.org/10.1152/physrev.00062.2017

Ioannucci, S., Chirokoff, V., Dilharreguy, B., Ozenne, V., Chanraud, S., & Zénon, A. (2023). Neural fatigue by passive induction: Repeated stimulus exposure results in cognitive fatigue and altered representations in task-relevant networks. *Communications Biology*, *6*(1), 142. https://doi.org/10.1038/s42003-023-04527-5

Theriault, J. E., Shaffer, C., Dienel, G. A., Sander, C. Y., Hooker, J. M., Dickerson, B. C., Barrett, L. F., & Quigley, K. S. (2023). A functional account of stimulation-based aerobic glycolysis and its role in interpreting BOLD signal intensity increases in neuroimaging experiments. *Neuroscience & Biobehavioral Reviews*, *153*, 105373. https://doi.org/10.1016/j.neubiorev.2023.105373

Wiehler, A., Branzoli, F., Adanyeguh, I., Mochel, F., & Pessiglione, M. (2022). A neuro-metabolic account of why daylong cognitive work alters the control of economic decisions. *Current Biology*, *32*(16), 3564–3575. https://doi.org/10.1016/j.cub.2022.07.010

Metabolic reality of cognitive load

Liye Zou[a], Peng Wang[b], Jun Yin[c], Kim Ouwehand[d] and Fred Paas[d]*

[a]Shenzhen University, China; [b]Vrije Universiteit Amsterdam, Netherlands; [c]Ningbo University, China and [d]Erasmus University Rotterdam, Netherlands
liyezou123@gmail.com
p.wang@vu.nl
yinjun1@nbu.edu.cn
ouwehand@essb.eur.nl
paas@essb.eur.nl

*Corresponding author.

doi:10.1017/S0140525X26106050, e485

Abstract

Haueis and Colaço argue that metabolic constraints should prune and generate cognitive models. Cognitive load theory offers an applied framework for testing these constraints. Working memory load can be viewed as the behavioral expression of local ATP limits. Primary–secondary knowledge, embodied offloading, and instructional efficiency map onto bioenergetic trade-offs. Extending CLT to human–AI symbiosis further shows how shared systems can support dynamic load balancing and testable metabolic predictions.

We welcome Haueis and Colaço's argument that metabolic realities should both constrain and generate cognitive models. Their distinction between evaluative constraints that rule out implausible accounts and generative principles that guide plausible ones is especially useful. Their rejection of global glucose-depletion models of mental effort and their emphasis on local metabolic demands tied to specific neural substrates challenge any strong version of medium independence. We agree that the brain's energy budget is not background scenery but an active constraint on sustainable information processing. Our claim is that cognitive load theory (CLT) provides a mature applied framework for testing and operationalizing these metabolic constraints in learning, instruction, and, increasingly, human–AI symbiosis. In that sense, CLT can help move metabolic realism from a valuable philosophical demand to a cumulative research program.

The key link is straightforward. The working-memory limits central to CLT are not just functional metaphors; they are the behavioral expression of the local bioenergetic constraints Haueis and Colaço describe (Sweller, van Merriënboer, & Paas, 2019). Working memory is the limited workspace for processing novel, biologically secondary information. When that workspace is overloaded, performance declines, and effort rises. CLT has long explained this pattern in terms of capacity limits and allocation trade-offs. A metabolic perspective helps explain why such limits exist. Local ATP availability, synaptic maintenance costs, and metabolic by-products place hard bounds on how many novel elements can be processed and integrated at once. Decades of CLT findings on mental effort, instructional design, and learning outcomes therefore offer a natural empirical test bed for metabolic models of cognition. Our recent synthesis extends this logic to AI context windows, treating them as bounded workspaces analogous to human working memory and showing that similar load-management principles may scale to hybrid systems (Wang et al., 2026; see Figure 1). The parallel is not identity, but it is theoretically productive because both systems face finite processing budgets and must balance storage, selection, and coordination under those budgets.

A first implication concerns secondary knowledge. Haueis and Colaço note that evolution has shaped neural efficiency. CLT formalizes this point through the distinction between biologically primary and secondary knowledge (Paas & Sweller, 2012). Primary knowledge, such as face recognition or basic language, is acquired with little conscious effort because evolution has already tuned the system for it. Secondary knowledge, such as reading, algebra, and scientific concepts, requires sustained high-cost processing in systems not optimized for these culturally recent tasks. Instructional designs that reduce extraneous load provide direct evidence of this metabolic burden. Replacing conventional problems with worked examples, for instance, reliably lowers reported mental effort while improving retention and transfer. These effects can be interpreted as metabolic optimization: fewer resources are wasted on redundant search, leaving more available for schema construction and automation. The primary–secondary distinction is therefore not merely psychological. It also marks differences in the bioenergetic demands of learning and in the kinds of instructional support culturally recent knowledge requires.

A second implication concerns embodied offloading. Haueis and Colaço argue that metabolic demands tie information processing to particular substrates. Recent CLT work on embodied cognition shows how learners can exploit this constraint rather than merely suffer from it (Zou et al., 2025). Gestures, object manipulation, and other environmental interactions can offload computationally demanding operations onto motor and perceptual systems. When learners trace diagrams or manipulate concrete models, some of the burden otherwise borne by frontoparietal networks can be redistributed. Experimental work shows that such strategies often reduce mental-effort ratings while improving immediate performance, retention, or transfer. On this view, embodied actions are not peripheral aids but metabolically efficient ways of reorganizing computation across the brain, body, and environment. We have argued that shared workspaces and retrieval-augmented generation play a comparable offloading role

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

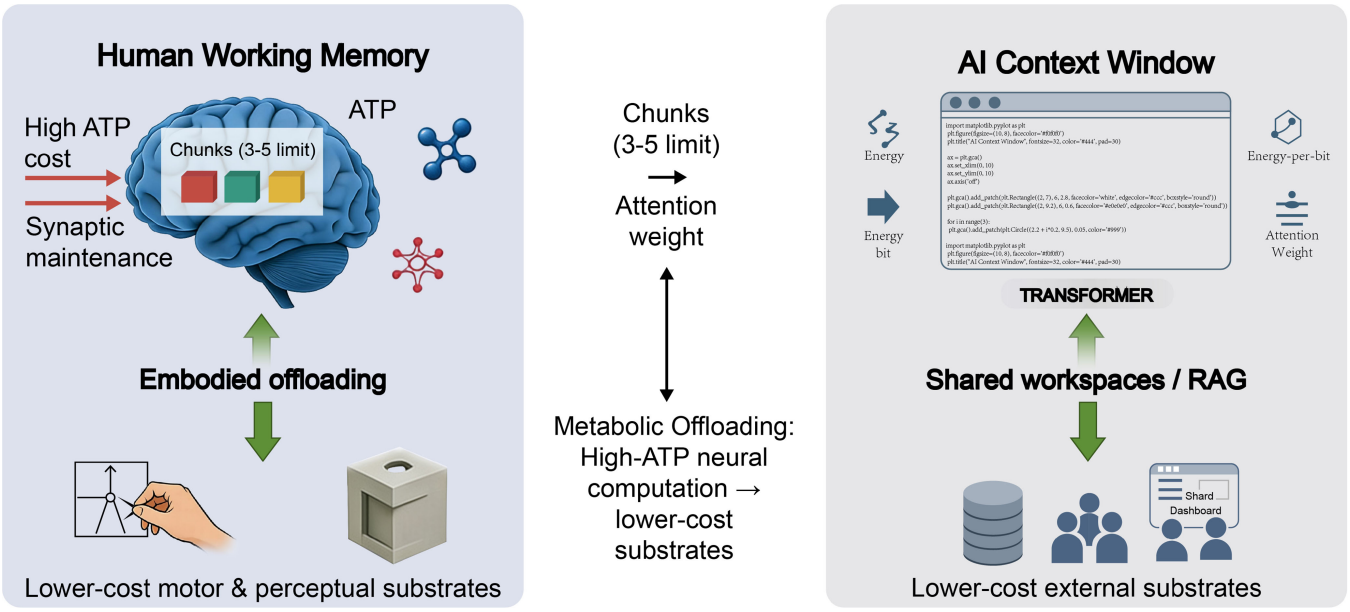


Figure 1. Working memory as the functional manifestation of local metabolic constraints. Left panel: human working memory (3–5 chunks) limited by ATP availability and synaptic maintenance costs. Right panel: AI context window (thousands of tokens) limited by energy per bit and attention dilution. Arrows show metabolic offloading via embodied actions (human) and shared workspaces/retrieval-augmented generation (human–AI symbiosis), converting high-ATP neural computation to lower-cost substrates.

in human–AI systems by reducing avoidable load on both the person and the model while preserving performance.

A third implication is methodological. CLT's instructional-efficiency metric can be interpreted in the ATP/bit terms proposed by Haueis and Colaço (Paas & van Merriënboer, 1993). For decades, CLT research has combined performance with mental-effort ratings to estimate how efficiently instruction turns invested effort into learning gains. Reframed biophysically, this becomes an approximation of energy spent per unit of durable, retrievable information. That translation is indirect, but it generates clear predictions: designs that minimize unnecessary processing should yield better efficiency profiles, and hybrid human–AI pipelines should benefit from dynamic load balancing when either partner approaches overload. CLT intervention studies already provide many opportunities to evaluate such predictions across formats, sequencing principles, and coordination structures.

This also suggests concrete tests of the evaluative standards Haueis and Colaço advocate. Randomized studies that manipulate element interactivity, format, or modality while measuring effort, performance, and physiological proxies such as pupil dilation or EEG could assess whether a cognitive model respects local bio-energetic constraints. When an instructional manipulation reduces effort without impairing learning, it offers indirect evidence for a more favorable energy-information trade-off. Future work could tighten this bridge by combining CLT paradigms with direct metabolic imaging, including in hybrid systems. In short, CLT does not merely illustrate the target article's proposal. It offers an empirical and interventional program for turning metabolic considerations into testable predictions about learning, instruction, and coordinated human–AI cognition.

Acknowledgments. The authors have nothing to report.

Financial support. This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

Open access funding provided by Erasmus University Rotterdam.

Competing interests. The authors declare no competing interests.

References

Paas, F., & Sweller, J. (2012). An evolutionary upgrade of cognitive load theory. *Educational Psychology Review*, *24*(1), 27–45. https://doi.org/10.1007/s10648-011-9179-2

Paas, F., & van Merriënboer, J. J. (1993). The efficiency of instructional conditions: An approach to combine mental-effort and performance measures. *Human Factors*, *35*(4), 737–743. https://doi.org/10.1177/001872089303500412

Sweller, J., van Merriënboer, J. J., & Paas, F. (2019). Cognitive architecture and instructional design: 20 Years later. *Educational Psychology Review*, *31*(2), 261–292. https://doi.org/10.1007/s10648-019-09465-5

Wang, P., Liu, H., Zou, L., & Paas, F. (2026). Overloaded minds and machines: A cognitive load framework for human-AI symbiosis. *Artificial Intelligence Review*, *59*, 109. https://doi.org/10.1007/s10462-026-11510-z

Zou, L., Zhang, Z., Mavilidi, M., Chen, Y., Herold, F., Ouwehand, K., & Paas, F. (2025). The synergy of embodied cognition and cognitive load theory for optimized learning. *Nature Human Behaviour*, *9*(5), 877–885. https://doi.org/10.1038/s41562-025-02152-2

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

Authors' Response
Giving metabolism its due

Philipp Haueis[a]* and David Colaço[b]

[a]Department of Philosophy, Institute for Studies of Science, Universität Bielefeld, Germany and [b]Center for Mathematical Philosophy, Ludwig-Maximilians-Universität München, Germany
philipp.haueis@uni-bielefeld.de
david.colaco@lmu.de

*Corresponding author.

doi: 10.1017/S0140525X26107298, e486

Abstract

The commentators point out elaborations and challenges to incorporating metabolic considerations into cognitive modeling. We agree with many of them and highlight that these issues can only be addressed when we actively try to include metabolism while evaluating and generating models. The main message of our response is thus that metabolic considerations are epistemically valuable, regardless of the particular modeling approach, conceptualization, or specific ontological commitments regarding cognition. Yet, the commentator's contributions also reveal that metabolism alone is never the decisive factor in cognitive modeling, and that its value for particular projects needs to be demonstrated in the modeling process, rather than assumed from the outset.

R1. Introduction

The diverse scholarship represented in the commentaries proves that metabolism should be discussed more in the mind and brain sciences. The commentaries highlight that metabolism plays a role for research on multiscale dynamics, brain networks, functional specialization, cognitive load theory, predictive processing, 4E cognition, neurodevelopment, resource rationality, cognitive ontology, computation, artificial intelligence, and more. They also reveal that beyond these topics, metabolism raises questions about how we conceptualize cognition and ontological commitments regarding the mind and brain.

The above sketch crystallizes two main points of our response. First, the concerns the commentators present can only be answered, and perhaps can only be posed, by trying to incorporate metabolic considerations into cognitive modeling via our account. Whether these attempts will succeed depends on how the results bear out. Second, we stress the *cross-cutting epistemic value* of metabolism for cognitive modeling; there is no one modeling approach, conceptualization of cognition, or ontological commitment that is decisively favored over another. Instead, we invite all parties to consider how explicitly considering metabolism could improve their modeling approaches and theorizing about cognition and the mind. Some commentators take us (or want us) to say something stronger, and we clarify where we agree and disagree with these more specific claims. Nonetheless, they speak to a main theme of the target article; modelers should determine how their models fit into metabolic limits or at least whether their models can be understood in terms of these limits.

Our response includes five core topics: whether, how, and under what circumstances metabolic considerations can be incorporated into cognitive models (R2), whether incorporating these considerations favors some modeling approaches over others (R3), whether metabolism is constitutive of cognition (R4), how our work reappraises the mind–brain relation (R5), and how our claims connect to efficiency and energy costs of information processing in brains and in artificial intelligence (R6).

R2. Incorporating metabolism into models

The first question is under what conditions we can incorporate metabolism into cognitive modeling. We reiterate our desiderata (R2.1) before clarifying misunderstandings of what we mean by "evaluation" and "generation" (R2.2). We also point out that holding onto implausible models enriches our account by connecting it to the context of pursuit (R2.3).

R2.1. Our desiderata

We include three desiderata for metabolic knowledge to act as full-blown evaluative or generative considerations in cognitive models. First, the knowledge must be *specific* to a computation, a cognitive capacity, or a neural system. Specificity in this sense is not restricted to localization in brain areas or systems, as some commentators inferred from our core examples (Williams, Mace et al.). We do not think that including metabolism always entails localizing a cognitive capacity (see R3.1). Second, metabolic knowledge must be *relevant* to modeling the target, which takes the form of explanatory constraints. Again, some commentators (mistakenly) infer from the mental effort case that metabolism must have *direct explanatory relevance* (Hendrickx), which is not our view (see R3). Third, metabolic knowledge must be *legible*, which may require that the model itself is re-formatted, e.g., to include specific information-processing costs (Khoudhary et al.).

Beyond our examples, the desiderata are corroborated by the examples commentators have provided. Zou et al. specify implications of metabolic considerations for cognitive load theory. Adenosine triphosphate (ATP) availability and synaptic maintenance costs are relevant to human working memory, as they constrain "how many novel elements can be processed and integrated at once." Metabolic knowledge is also legible to experimental measures like an instruction-efficiency metric, although further work is needed to translate it into modeling parameters. Zenon & Solopchuk specify metabolic considerations for cognitive fatigue, arguing that passive fatigue paradigms rule out global resource or prefrontal-only models of this capacity. Their generative hypothesis shows that aerobic glycolysis is legible to cognitive fatigue models, resulting in testable quantitative predictions about cognitive parameters like degree of representational degradation. The commentaries on metabolic considerations for models of mind-wandering and the default mode network (Wong), alternatives to stimulus–response views of neural function (Alvarez-Hernandez & Theriault), and neurotypical and neurodivergent development (Siddiqui & Jones) are also consistent with our desiderata.

Yet, our desiderata are not trivially met in all cases. Historically, metabolism at the microscale of ATP use was hard to measure

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

(Kondrakiewicz & Gonçalves), leading to an absence of data to explicitly fit cognitive models (Khoudary et al.). Yet there are "work-arounds," such as using coding efficiency schemes (Pohl), indirect constraint by matching models to neural dynamics (Kondrakiewicz & Gonçalves), calculating ATP using basic molecular biological knowledge (Sengupta, Stemmler, & Friston, 2013), and measuring metabolism at different scales (Deery et al.). Thus, despite challenges, the examples and strategies suggest that satisfying the desiderata is possible and, in fact, a substantial achievement when developing cognitive models. They also highlight that a model in which the challenge is not met is deficient, even if it is understandable why it has not been met.

R2.2. Realities of evaluating and generating scientific models

With these desiderata reiterated, we defend their use in evaluation and generation. Shahmoradi argues that our distinction collapses because generation is simply evaluation at an earlier modeling stage. This criticism rightly points out that there is no sharp distinction between the context of discovery and of justification (Bechtel & Richardson, 2010; Kästner & Haueis, 2021), so generating new models may require evaluation. However, it would be wrong to conclude that we cannot articulate the difference between evaluating and generating models, nor that there are no clear cases of the two. In clear cases of generating a model, we are not testing a model's adequacy, explanatory power, or predictive success regarding the cognitive capacity or neural system that is targeted. In clear cases of evaluation, we explicitly assess two (or more) models along these dimensions. We think that our core examples in Sections 3.1.2 and 4.1.2 provide clear cases. Hendrickx picks up predictive success as an evaluation criterion to argue that in hybrid mental effort models (Section 4.2), "the metabolic component explains the same behavioral change that motivational/computational models already predict." At least currently, metabolic considerations do not seem to uniquely favor hybrid models. We agree with Hendrickx's conclusion, but for different reasons. First, evaluation is a method for assessing the relative strength of models. It need not be a success condition, where a model is shown to be strongest qua the evidence in question. Second, we did not present hybrid models as evaluative metabolic considerations, but to argue that computational and resource models are not in opposition. In fact, hybrid models like Holroyd's (2025) better match generative metabolic considerations. They provide a proof-of-concept ("controllosphere"), potential explanations (fatigue may signal neurotoxic waste buildup) and help us reassess the target system (link between computational mechanisms and metabolic processes). Third, pace Hendrickx, metabolism can be used in evaluation even when it is not directly explanatory of features of mental effort but instead acts as an explanatory constraint (Section 3.1.1). Metabolic considerations, when they meet the desiderata, can be used to evaluate *any* model of mental effort, including resource models like the ego depletion model. Hendrickx juxtaposes computational and motivational approaches to modeling with metabolic ones, whereas we challenge this juxtaposition. We think that metabolism can be used to decide which of the many motivational or computational model versions are the most feasible. But we do not think that simply adding metabolic details to models makes them better in all cases. Rather, adding such details needs to fulfill the functions that we highlight in our desiderata and guidelines.

R2.3. Holding onto "implausible" models

The converse of the previous worry is that evaluative metabolic considerations are too stringent. Embrey worries that following guidelines (E1)–(E5) would lead us to prematurely discard potentially productive models. There may be no way to currently incorporate metabolic considerations into opportunity cost models of mental effort, for example, which we assume worries Embrey for evaluation against other, more developed models.

We think that Embrey helpfully points toward the context of pursuit (Laudan, 1978), which exists alongside the context of discovery (akin to generative considerations) and of justification (akin to evaluative considerations). New models initially may be less explanatory or predictive than older alternatives but nevertheless be pursuitworthy because of their potential to solve outstanding problems (Haueis & Kästner, 2022; Šešelja & Straβer, 2014). That a model is metabolically implausible initially does not mean it must be discarded immediately. Rather, it is a limitation that must be addressed in its development. We embrace Embrey's point to pursue nascent cognitive models but insist that it does not remove concerns about their plausibility. Having no idea how something is implemented is still a deficiency, but it alone need not utterly undermine pursuit of an algorithmic model, for instance.

This appeal to pursuit cuts both ways. The opportunity models Embrey discusses are now developed, and pursuit should support linking them with metabolic considerations. We can (and should) ask questions about their metabolic feasibility. This, we think, aligns with the heart of Marrian analysis (Marr, 1982). If Marr's "levels" are a mode of analyzing systems from the top down, then evaluation of models in terms of how well they fit metabolic constraints seems consistent with it. We do not think anything in Marr's analysis precludes researchers from using implementational-level details in assessing a computation that is likely implemented in this system, much less in using implementational constraints in this way. Marrian analysis presupposes *some* implementational story as part of this analysis. Put in slogan form, if a top-down analysis gets stymied before it reaches the bottom, there is a problem.

R3. Does metabolism favor particular modeling approaches?

Although we argue for incorporating metabolic considerations into cognitive modeling, we do not argue that a particular approach to modeling, such as computational, mechanistic, dynamical, or predictive processing, best or solely satisfies our account. Although we are generous while counting models as cognitive, this section addresses how particular modeling choices relate to metabolism.

Before addressing these responses, we stress that a cognitive model can be easily confused with its target, its explanatory role, or its exploratory origin. For example: we take Fig. 6 to be the novel cognitive model proposed by Theriault et al. (2023). The concern from Shamoradi mistakes the target of modeling, a neural system with thin axons computing prediction errors that can underlie cognitive activities, with a consideration during the model's development, the question of why aerobic glycolysis occurs in the brain. This question is answered by the efficiency-tradeoff hypothesis, which is part of the model, not the model itself. That the researchers include established components (e.g., prediction error) speaks not against combining them into a new model.

Correspondingly, when thinking in terms of explanation, including explanatory constraints, it is important that we keep in

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

mind the target of explanation (Colaço, 2020). Karbowski's cyclic model has synaptic memory consolidation as its target. Najenson claims that in this model, metabolic constraints are explanatory background conditions needed to rule out impossible models "before the explanatory work begins." We disagree with this narrow view of explanatory work. Models include factors besides direct explanatory factors, such as ATP hydrolysis to power synaptic consolidation (Fig. 5). According to Klein (2018), resources share only some properties with background conditions, and others with direct explanatory factors. Consider ATP. Like background conditions, it is available in bulk, causally contributes to many processes, and has a specific chemical form. But like explanatorily relevant factors in (mechanistic) explanations – and pace Najenson – ATP can have specific effects on memory; it can enhance or inhibit the amount of long term potentiation and memory consolidation in a gradual fashion (Hu et al., 2025; Wang et al., 2004). Finally, Karbowski (2024) thinks of energy and information as physical states. So contrary to Najenson, energy can structure and guide the flow of information (e.g., memory lifetime) through a system (e.g., the synapse), although other interpretations of information flow are of course possible. This reinforces that the relevance of metabolism is specific to the cognitive capacity or neural system, which is why we do not think one can simply transfer our interpretation of the mental effort case to the synaptic memory case. However, these concerns should not distract from our main point. We think of metabolic constraints as explanatory. Yet, what contributes to the explanation is the information from the *entire model* (plus background knowledge), and not the constraints simpliciter.

R3.1. Connections between mechanistic and computational modeling

Our response to Najenson may suggest that we favor a mechanistic approach, in which systems are decomposed and localized to entities and activities, over other modeling approaches. Williams contends that nonlocalizationist modeling speaks against a putatively "monist" mechanistic modeling approach.

For our desiderata, specificity need not entail localization in the classic mechanistic sense (Bechtel & Richardson, 2010). Our guidelines do not only apply to mechanistic modeling, and many of the commentator's examples show how metabolism plays a role in other approaches (Kong et al., Jamalabadi & Hahn, Starita et al., Penacchio & Clemente, Deery et al.). Contra Williams, we do not think of cognitive models as a specific type of (nonmechanistic) model, nor of all cognitive models as (ideally) mechanistic. Rather, cognitive models convey information *about* particular cognitive capacities or neural systems relevant to these capacities. Likewise, we think that different modeling approaches can contribute to explanatory understanding of a given target (Hochstein, 2017). Our account thus provides no defense of monist over pluralist stances to modeling.

We agree with Williams that our core examples fit the mechanistic approach. But we reject that including metabolic considerations in any manner that satisfy our desiderata automatically makes a model mechanistic. We worry that this reading implies that any metabolic details count as localization. Likewise, our guidelines (E1)–(E5) at no point require explicitly representing metabolic considerations, much less representing them as system components; explicitly represented considerations are neither necessary nor sufficient for successful evaluation. Mere appeal to metabolism counting as localization or componential is too low of a bar for distinguishing cognition or computation from mechanism, especially when other mechanistic criteria are not satisfied. Further, it underestimates the helpful methods that are employed in cognitive modeling. For instance, these systems can be approached via appeal to Simon's near-decomposability of complex systems (Simon, 1962; Weinberger, 2020), allowing us to assess how a system's operations relate to an ascribed computational activity. Modeling methods like this supply a defense for our position without limiting us to any one framework, such as neurocomputationalism (Piccinini, 2020), from which some concerns about the relation between mechanism and computation may stem. Nonetheless, any approach must answer when a model contributes to explanatory understanding, and metabolism is part of the evidential constraints by which we can prune, partition, and populate the space of possible models.

Although computational modeling is often associated with Marrian analysis, controversies over what "computational level" or "computation" means abound (Egan, 2025; Shargrir, 2010). On our ecumenical approach, computational modeling provides mathematical descriptions of operations, which process information in the brain, and there is a specific relation (mapping, narrative, or other), which relates this attributed computational function to features of cognitive and neural systems. Our view of evaluative metabolic constraints also follows mainstream computational neuroscience, which takes metabolism to be external to the operations or mechanisms that implement the computational function (see Section 3.2.1).

For some commentators the abstract nature of computation is in tension with metabolic considerations (Dengsø), whereas others argue that it makes metabolism irrelevant for this modeling approach (Embrey, Mace et al.). There are too many presuppositions for us to definitively address this issue here. What is often overlooked in these discussions, however, is that in complex biological and artificial neural networks, the computational function that modelers attribute *approximates* the operations of the actual system (cf. Cao & Yamins, 2024, p. 3). Like other constraints (e.g., tractability van Roij, 2008), metabolic considerations can help us adjudicate which candidate best approximates an actual system operation. We agree with Williams and Embrey that not every model will currently allow such comparison directly. They may have to be re-expressed before they can be connected to metabolic considerations (Khoudary et al.).

R3.2. Dynamical and multiscale modeling

Some commentators emphasize the importance of metabolism for modeling scale and dynamics of cognitive systems. Kong et al., for instance, highlight that greater appreciation of meso- and macroscale neural networks can offer a better venue for integrating metabolic considerations in cognitive modeling. We agree that the mesoscale is important (Colaço, 2021; Haueis, 2014, 2021), yet what counts as the "mesoscale" can vary amongst models, and these models can also span multiple levels of neural organization. This speaks to Penacchio & Clemente, where system-level considerations are recommended in addition to research at the neuronal level, and Deery et al., in which a scaling gap is overcome through imaging methods to relate metabolic activities to larger-scale networks. Our desiderata and other commentaries highlight that many cognitive phenomena are on one reading or another "multilevel" or "multiscale" (Haueis, 2022; Mok & Love, 2023).

Combining scale and dynamics, Favela also encourages us to incorporate multiscale dynamics into our framework. We agree

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

that incorporating metabolism can also accommodate and benefit from bringing in temporal dynamics, as metabolic processes, including some from our examples, highlight temporal dynamics, sometimes as an implication of their multiscale spatial character. This is not something one gets for free from including metabolism; it must be addressed. However, which dynamics matter will be as variable as which considerations are incorporated. There is not a one-size-fits-all option for incorporating time. Likewise, we should distinguish incorporating brain dynamics from dynamic approaches to cognitive modeling. The latter often speak of "energy functions" or "energy landscapes" without being explicitly constrained by metabolic considerations (Beer, 2003; Hopfield, 1982; Rabinovich, Huerta, & Laurent, 2008). De Wit makes a similar point about ecological approaches, which are theoretically sympathetic to, but background, metabolism. These examples illustrate why we do not think that dynamical approaches are the favored, let alone only, options available for incorporating metabolic considerations. Other options are live, and whether they can incorporate metabolic considerations depends on the hard currency of the descriptive, explanatory, and predictive success of the models.

R3.3. Predictive processing approaches

Predictive processing (PP) and free energy principle (FEP) combine tenets of the previously discussed approaches. We read PP and FEP as presupposing that (a) information is physical, (b) information processing is subject to efficiency constraints, and (c) cognitive functions are subject to metabolic constraints because they minimize prediction error or free energy (Friston, 2010; Sengupta, Stemmler, & Friston, 2013; Sprevak & Smith, 2023) These assumptions make PP and FEP models of cognitive capacities suitable candidates for metabolic considerations because any neural system that regulates metabolism "must embody a model of the variables it regulates" (Jamalbadi & Hahn). Likewise, minimizing prediction error reduces redundant signals, which could reduce "metabolic costs within a global computing scheme" (Pennacchio & Clemente, and see Hechler, de Lange, & Riedl, 2023 for experimental evidence).

Our target article points out that PP and FEP approaches can generate novel cognitive models (Section 3.3.2). Yet, metabolic considerations do not uniquely favor these approaches, as metabolism can also be integrated into other types of modeling. There is also a further question that merits careful examination. Because PP and FEP incorporate metabolism and energy-information efficiency as theoretical axioms, models derived from these axioms are already metabolically constrained (e.g., Ali et al., 2022; Zhang et al., 2025). Can evaluative metabolic considerations also be used to select which of multiple PP schemes is the most biologically plausible for a given cognitive capacity or neural system?

R4. Is metabolism constitutive of cognition?

Garcia-Cazorla & Mochel, Rorot and Starita et al. argue that metabolism may be constitutive of cognition itself. This claim raises issues of how we *conceptualize* cognition. Conceptualization includes but goes well beyond concerns about modeling (Colaço, 2022; Haueis, 2018). A concept of cognition requires positing characteristics or properties that make a system cognitive, such that the extension of COGNITION is delineated. Yet cognitive scientists seem to build models that succeed, albeit imperfectly, to describe, classify, predict, and explain how (uncontroversial) cognitive systems behave. The existence of successful cognitive models in the absence of an agreed-upon concept of cognition shows that conceptualization and modeling are different yet interdependent scientific activities. Thus, one can incorporate metabolism into cognitive modeling, regardless of whether one adopts a narrow or permissive concept of cognition. The constitutive perspective is thus compatible with, but not necessary for our account.

One reason for our neutrality is as follows. We accept for the sake of argument that metabolism is *necessary* for (at least biological) cognition (Godfrey-Smith, 2016a). Yet whether cognition is also constitutive is presently *underdetermined* by evidence. If metabolism is only necessary, evidence of a metabolic factor influencing a cognitive capacity might at best be seen as enabling cognition, and at worst as something which must be ruled out as a confounder when attempting to identify other conditions necessary for cognition (Colaço, 2024; Haueis, 2018). If metabolism is constitutive, by contrast, the very same evidence is a positive feature of cognition.

Our neutrality extends to claims about the essential function of cognition, e.g., allostasis (Sacco). Our account can be integrated with but does not prove such a theory. Other approaches with different guiding principles can also benefit from evaluative and generative metabolic considerations. Likewise, the underdetermination worry equally applies to functional ascription. If we treat functional ascriptions as hypotheses to be tested via empirical insights, there is a productive and operative manner in which these tests can inform us of functions and of cognition more generally.

A similar logic plays out for viewing metabolism as necessary versus constitutive in cognitive modeling. Garcia-Cazorla & Mochel argue that "metabolic states are internally represented and used to drive processing," but metabolism being internally represented is not the same as metabolism itself being representational. If one conceptualizes cognition as essentially representational (Bermudez, 2010, but see Ramsey, 2017), then one could model how metabolism drives processing of representations without metabolism itself being part of those representations.

Conversely, Rorot argues that the physicality of information processing speaks for ruling in unconventional cases of cognition (e.g., oil droplets, slime molds). PP also assumes that information is physical (R3.5), but a modeler who applies this modeling approach to conventional systems (e.g., the hippocampus or visual system) is not required to count unconventional systems as cognitive. This, again, is a bigger conceptual issue (see Colaço, 2023).

Finally, Starita et al. argue for an embodied or enactive concept of cognition because nervous systems, cognition, and behavior evolved "as a solution to the problem of controlling a specific body in interaction with its environment," but a modeler might investigate this solution while also assuming – contrary to enactivism – that nervous systems *regulate and control* the metabolism of the rest of the body (Theriault et al., 2025). Starita et al.'s octopus example is not a good case for this difference in outlook because "the octopus is suffused with nervousness, [its] body is not a separate thing that is controlled by the brain or nervous system." (Godfrey-Smith, 2016b, p. 75).

These perspectives reinforce our claim about the cross-cutting epistemic value of metabolic considerations. Yet, issues of conceptualization caution us against taking evidence as decisive for drawing the boundaries of cognition narrowly or broadly. Although metabolism can and should be an important ingredient for delineating the mark of cognition if indeed there is one (Colaço,

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

2022), no one party gets to exclusively claim it in boundary disputes.

R5. Realization, medium, and ontological commitments

Despite our neutrality, our account invariably relates to issues of multiple realization, medium independence, and other ontological commitments. We respond to commentaries by clarifying that the implications of our thesis depend both on the imputed *content* of and *epistemic stance* toward these positions.

R5.1. Functionalism and multiple realization/realizability

In Section 4.1 we argue that there is a substantive reading of multiple realization that is compatible with incorporating metabolic considerations into cognitive modeling. Metabolic factors themselves might be multiply realized across the biological realm, as their life-sustaining function is something they have in common. Guest et al. support this reading because they argue that the same functional capacity is multiply realized if there are multiple mechanisms that underlie that capacity.

Although not necessarily unclear, scholars do not always mean the same thing by "multiple realizability." For example, Thagard argues against strong multiple realizability but favors a scant version, where brains and certain AI systems realize intelligence because they mutually satisfy energy, time, space, and history constraints. Is this in tension with Guest et al.? The answer not only depends on the content, but also our epistemic stance toward multiple realization (or realizability). If Guest et al. and Thagard's positions are read as working hypotheses, for instance, then they are not at odds with one another. Both are revisable means of exploring relations between cognition and its physical basis. However, if they are core metaphysical commitments, then facts about metabolism challenge them (Cao, 2022), and both are in tension with general worries about multiple realization (Polger & Shapiro, 2016). Nonetheless, neither Guest et al. nor Thagard are in principle at odds with our thesis. It is just that incorporating metabolic considerations into cognitive modeling does not require any one metaphysical outlook, nor does it alone prove any one outlook correct.

R5.2. Medium (in)dependence of computation

Several commentaries shed light on how "medium independence" should be understood as connecting with our account. According to Williams (see also Williams, 2025) there are two senses of medium independence: as a formal system that is realizable in many physical systems (Haugeland, 1985), and as an abstract, less detailed description of a physical system that has enough degrees of freedom to execute a computational rule (Piccinini, 2015). Mace et al. criticize that our view is too focused on Piccinini's account. We take seriously that it might be idiosyncratic qua medium independence positions and perhaps different throughout his existing works. What Piccinini gets right, however, is that medium independence specifies a relation between a physical system and a cognitive activity that permits variability between systems that underwrite (or realize) the activities. When we engage in cognitive modeling, we have expectations about how the model relates to the brain (or the brain plus, for 4E proponents), including expectations about the media that realize computational operations. Those media may vary without corresponding variance in the operations.

Taking medium independence as an epistemic stance toward the system (see R5.1), we wonder whether implications of the two senses manifest in the context of cognitive modeling. We do not read Haugeland's (1985, p. 58) sense of medium independence as entailing that *any* arbitrary medium can materialize the same formal system. And we leave open the idea that metabolic considerations can materialize in different media yet have a unity to be captured in our modeling practices. This parallels our points about Marrian levels (R2.3). As cognitive modelers we operate with a working but revisable delineation of media of interest, e.g., spiking time differences for neural computation (Piccinini, 2015). These media can share constraints, such as availability of metabolic resources. It seems entirely consistent with the spirit of medium independence that we can use these constraints to assess models of how these systems operate. Thagard's position, we think, is a litmus test: we take it as consistent with medium independence as a working hypothesis, but we imagine other commentators would not. Generally, how metabolism relates to medium independence requires us to explicate the position, our epistemic stance toward it and careful reflection upon scholarship opposed to medium independence.

R5.3. Ontological perspectives

Our discussions of realization and medium, paired with Section 4, return us to an exploration of ontological perspectives when incorporating metabolic considerations in cognitive modeling. No doubt, ontological positions can influence modeling. As Dengsø illustrates, if you are ontologically committed to metabolism being constitutive of cognition – which we are not – then you build different models. More minimally, if one commits to metabolism being relevant to cognition, which is closer to what we hold, then this commitment can influence what system characteristics are deemed plausible during evaluation. This is shown in commentaries that address architectural commitments of cognitive systems and, by consequence, how they should be modeled. Van der Velde argues that metabolic considerations favor architectures that allow compositional processing and access to long-distance information (e.g., small-world networks). Percy uses energy efficiency to hypothesize that evolution favors anatomically specialized structures, whereas Siddiqui & Jones show how such specialization can arise via connection strengthening on a developmental timescale. Moreau & Wiebels alternatively argue that because cognitive systems draw on common resource pools, "sharp and stable boundaries between cognitive components are unlikely to hold across all contexts."

We think these takes on the ontological implications of metabolism are all worthy of further investigation. Yet they are not all the same. The position of Moreau & Wiebels includes metabolism in the specification of cognitive modules as local optima, making them medium-dependent. This aligns with criticisms of medium independence which argue that metabolism challenges the distinction between computationally relevant and irrelevant properties (Cao, 2022; Chirimuuta, 2022). By contrast, Percy and Siddiqui & Jones posit energy-saving mechanisms on evolutionary or developmental timescales that could be implemented in different media, akin to energy-saving learning rules (cf. Section 3.2.3). Looking at metabolism through an ontological lens, we are skeptical that there is a singular perspective on cognitive systems and how we model them.

Shifting to a potentially different sense of "ontology," Mace et al. challenge our account by highlighting how questions of *cognitive ontology* affect the role metabolic considerations may play in cognitive modeling. Having hopefully allayed worries about

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

monism, we do not commit to any idea that metabolic properties of a physical system are needed for identifying, much less individuating, a cognitive capacity or process. We think this matter lies in the domain of conceptualization (R4). Identification is a complex topic, both when identifying target phenomena (Colaço, 2020) and domains for concepts (Colaço, 2022; Haueis, 2024, Haueis & Margulies, 2026). In these works, we address general and discipline-specific processes by which categorization of the constructs of a so-called "cognitive ontology" might operate, provisionally, iteratively, and non-monistically. This means that reconceptualizing cognitive capacities will happen alongside, and can inform (or be informed by), metabolically inspired cognitive modeling.

The idea, however, that models cannot be compared or generated, even provisionally, without *resolving* conceptual issues first is too strong for productive, realistic science. If guidelines E1 and G1 cannot be satisfied, then metabolic incorporation does not get off the ground. For this reason, Mace et al.'s discussion of vehicle indeterminacy points to issues in cognitive modeling more generally. If modelers cannot specify the capacity that the model targets, a whole swath of comparisons is undercut, not just evaluative metabolic considerations. If it is not possible to compare models on any dimension, then metabolism will not inform evaluation between the two. Metabolic considerations are no "silver bullet." Indeed, we think that no one factor, be it about brains, the body, or the environment, deserves this label in any realistic depiction of the sciences of the mind.

R6. Efficiency and energy costs of information processing

A final theme of the commentaries addresses appeals to efficiency and energy costs, including their relation to and differences between biological cognition and artificial intelligence. These commentaries reinforce the need to differentiate senses of "efficiency," which invariably must be contextualized to both the target of the models that are evaluated or generated and the metabolic considerations germane to these models.

R6.1. Efficiency

Section 2.1.2 discusses several senses of "efficiency," which are not equivalent but nonetheless can be relevant to cognitive modeling. Comparing systems in terms of "efficiency" risks conflation and requires us to specify the target and scale at which we compare it. Karbowski suggests that at the scale of neurons and synapses, the brain uses more energy per operation than digital and neuromorphic supercomputers, even though the brain is more efficient at other spatial and temporal scales. This highlights how efficiency comparisons become misleading when one ignores scale. Fleig-Goldstein similarly points out how comparing brains and Large Language Models (LLMs) like Chat-GPT must consider time and task specificity. Both commentaries point to our first tidbit in the target article. The brain uses a lot of energy, and we are not surprised that on some calculations its operations are more expensive than those in silico. This is why principles of neural design (Section 2.1.1) are valuable; they delimit how energy costs can be minimized because otherwise these costs add up quickly.

However, we think sympathetic to Karbowski's general outlook, and one must ask what counts as an "operation" in both biological and artificial systems while calculating energy per operation. Computational power in artificial systems like El Capitan or Hala Point is measured in floating point operations, where an operation is an arithmetic computation (addition, subtraction, multiplication, division). If one takes seriously Karbowski's claim that "an action potential corresponds to a single neural operation," it is questionable that one neural operation equates to one arithmetic operation. Likewise, "operation" at the neuronal or synaptic scale is not "task" in the vein of Fleig-Goldstein; if a brain operating at 17W in a human body and AI system operating at 35MW perform the same task, there is a potential time difference in task performance where AI systems will operate more efficiently. It is also worth noting that these wattage estimates do not tell the whole story. On the one hand, the human brain simultaneously performs multiple cognitive functions, so comparing LLM inference to the brain's whole energy budget (Fleig-Goldstein) is misleading. On the other hand, the human brain's 17W budget also goes to repair and maintenance, with far less dedicated to cognitive activities (Levy & Calvert 2021). Maintenance costs for supercomputers, such as humans who maintain and repair data centers, are typically not factored into energy budget estimates of AI systems. These quibbles with calculation reinforce our and Karbowski's point that precision and specificity are key for informative brain–AI comparisons. Failing to do either leads to a misleading calculation.

This nuance regarding "efficiency" also speaks to Pohl, who argues that abstract coding efficiency considerations can play nongeneric roles in computational neuroscience. His examples succeed precisely because they meet our desiderata; coding efficiency considerations are specific (e.g., bullfrog inner ear), relevant (e.g., ecological relevance, explaining variance), and legible (information theory applied to neural systems). The only difference is that Pohl discusses hypotheses rather than models, so our evaluative guidelines may not apply directly. However, Pohl recognizes that efficient coding need not be linked (via representation or explanation) to metabolic considerations. We thus find abstract considerations promising, yet we emphasize that not all "efficiency" is the same. This also makes precise why "evolution by natural selection need not always produce perfectly optimal solutions" (de Wit; see also Nave, 2026). As our core generative example highlights, sometimes different forms of efficiency cannot be mutually maximized, but tradeoff against each other. We also see no issue with expanding efficiency considerations beyond the brain as de Wit proposes, as long as such modeling efforts make metabolism explicit and satisfy our desiderata and guidelines.

R6.2. Systematizing thoughts on energy use

Our account provides a systematic way to incorporate metabolic considerations and to think about energy use. Yet Fleig-Goldstein worries that without a resource-rational analysis (RRA), there is no rationale for uniquely focusing on metabolic considerations over other physical or biological detail of the target system. We repeat that metabolism can be conceptualized in multiple ways, none of which makes it unique or a "silver bullet" for cognitive modeling. Besides, Section 3.1.1 highlights that we are not utterly opposed to RRA, even though we think it is open how RRA's unifies versus how informative it is for specific cases. Fleig-Goldstein is right that RRA is generative in a top-down manner, but we present systematic guidelines for a bottom-up, exploratory use of metabolic considerations to rigorously generate new cognitive models (Haueis, 2022). In this sense, we could suggest RRA is too top-

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

down just as Fleig-Goldstein suggests we are too bottom-up. No doubt, both together are more beneficial than either alone.

Yet we worry that Fleig-Goldstein mischaracterizes generative considerations for reasons that parallel our reply to Shahmoradi's collapse worry (R2.2). As our guidelines indicate, model generation need not be *ex nihilo*. Rather, parameters are constructed or adapted from existing models (G3), and our core example from Theriault shows both construction and adaptation of a model using tenets of predictive processing, which are not monolithic. Choosing and combining existing model components in new ways can result in a new model, even if it fits an existing modeling approach, such as those we discuss in *R3.* What matters is that the process begins (G1) by specifying the cognitive capacity or neural system that one wants to model, which may well include specifying the task the capacity must solve. Here, we agree with Fleig-Goldstein: determining what one wants to model and why is critical. RRA can achieve this, but we also insist that not every analysis will move from the computational to algorithmic level. We do not always have a computational description of the target (say in G1). Our core generative example shows that researchers equally move from "implementation" to algorithmic or computational descriptions. In these cases, metabolic considerations can serve as starting points for plausible candidates ultimately to be tested. These approaches can synergize, rather than oppose each other.

R7. Conclusion

The commentaries and our responses speak to the diversity of implications of metabolic considerations for cognitive modeling. Despite this diversity, our responses highlight the unity of our analysis of the role of metabolism in cognition. Metabolic considerations are of cross-cutting epistemic value because they can be used to evaluate and generate cognitive models regardless of the modeling approach that is used. For mechanistic models, metabolic resource considerations can "invalidate (or at least cast suspicion upon)" certain models of causal organization (Klein, 2018, p. 8). For computational models, metabolism or energy use can help us select the best mathematical approximation or generate novel mathematical descriptions of how systems process information. Metabolic considerations at different scales can be linked in multiscale cognitive models and help build and constrain models of brain dynamics. Although it will not and cannot be expected to solve all the issues associated with these modeling approaches, our framework is of value to all of them.

Yet, our responses equally sketch another aspect of our outlook, which returns us to why metabolism has not been given its due. One reason, included in the target article, is no doubt knowledge limitations. Incorporation requires knowledge of metabolism that fits our desiderata. More research must be and is currently being conducted into the metabolic dimensions of the mind and brain, including by some commentators (e.g., Bo et al., 2025; Chiel et al., 2026, Deery et al., 2026; Guichet et al., 2026; Jamadar et al., 2025; Karbowski, 2024; Peters et al., 2026; Sacco, 2026; Theriault et al., 2025; Xue et al., 2026). Many commentaries, however, also suggest that including metabolism requires that one narrowly and decisively commits to a specific modeling approach, conceptualization of cognition or ontology. Our response casts doubt on this presupposition, and no commentary sways our neutrality on this matter. Rather, there is much more modeling freedom granted to incorporating metabolism, and further research is needed to determine where the bounds of this freedom lie. We expect that these bounds are far reaching, but minimally they remain to be determined.

References

Ali, A., Ahmad, N., de Groot, E., van Gerven, M. A. J., & Kietzmann, T. C. (2022) Predictive coding is a consequence of energy efficiency in recurrent neural networks. *Patterns, 3*(12), 1–12. https://doi.org/10.1101/2025.06.25.661482

Bechtel, W., & Richardson, R. C. (2010). *Discovering complexity: Decomposition and localization as strategies in scientific research*. MIT Press.

Beer, R. D. (2003). The dynamics of active categorical perception in an evolved model agent. *Adaptive Behavior*, *11*(4), 209–243. https://doi.org/10.1177/1059712303114001

Bermudez, J. (2010). *Cognitive science: An introduction to the science of the mind*. Cambridge University Press.

Bo, K., Kelly, C., Trumpff, C., Hirano, M., Thiebaut de Schotten, M., Picard, M., & Wager, T. D. (2025). Mitochondrial energy transformation capacity influences brain activation during sensory, affective, and cognitive tasks. bioRxiv. https://doi.org/10.1101/2025.06.25.661482

Cao, R. (2022). Multiple realizability and the spirit of functionalism. *Synthese*, *200*(6), 506. https://doi.org/10.1007/s11229-022-03524-1

Cao, R., & Yamins, D. (2024). Explanatory models in neuroscience, Part 1: Taking mechanistic abstraction seriously. *Cognitive Systems Research*, *87*, 101244. https://doi.org/10.1016/j.cogsys.2024.101244

Chiel, H. J., Coggan, J. S., Datta, G., Fellous, J.-M., Nourse, W. R. P., Quinn, R. D., & Thomas, P. J. (2026). Brain-inspired energy efficient technologies for next-generation artificial intelligence. *Biological Cybernetics*, *120*(2), 5. https://doi.org/10.1007/s00422-026-01038-4

Chirimuuta, M. (2022). The case for medium dependence: Comment on neurocognitive mechanisms by Gualtiero Piccinini. *Journal of Consciousness Studies*, *29*(7–8), 185–194. https://doi.org/10.53765/20512201.29.7.185

Colaço, D. (2020). Recharacterizing scientific phenomena. *European Journal for Philosophy of Science*, *10*(2), 14. https://doi.org/10.1007/s13194-020-0279-z

Colaço, D. (2022). Why studying plant cognition is valuable, even if plants aren't cognitive. *Synthese*, *200*(6), 453. https://doi.org/10.1007/s11229-022-03869-7

Colaço, D. (2023). Connecting unconventional cognition to humans: Unification and generativity. *JoLMA*, *3*(9), 163–178. https://doi.org/10.30687/Jolma/2723-9640/2023/02/002

Colaço, D. (2024). When remediating one artifact results in another: control, confounders and correction. *History and Philosophy of the Life Sciences*, *46*(1), 5. https://doi.org/10.1007/s40656-023-00606-2

Colaço, D. (2026). The problem of explaining shifting targets. *European Journal for Philosophy of Science*, *16*(2), 29. https://doi.org/10.1007/s40656-023-00606-2

Colaço, D. J. (2021). How do tools obstruct (and facilitate) integration in neuroscience? In J. Bickle, C. F. Craver, & A. S. Barwich (Eds.), *The Tools of neuroscience experiment* (pp. 221–238). Routledge.

Deery, H. A., Liang, E., Moran, C., Egan, G. F., & Jamadar, S. D. (2026). Brain glucodynamic variability is an essential feature of the metabolism–cognition relationship. *Proceedings Of The National Academy of Sciences*, *123*(3), e2510850123. https://doi.org/10.1073/pnas.2510850123

Egan, F. (2025). *Deflating mental representation*. MIT Press.

Friston, K. (2010). The free-energy principle: A unified brain theory?. *Nature Reviews Neuroscience*, *11*(2), 127–138.

Godfrey-Smith, P. (2016a). Mind, matter, and metabolism. *The Journal of Philosophy*, *113*(10), 481–506.

Godfrey-Smith, P. (2016b). *Other minds: The octopus, the sea, and the deep origins of consciousness*. Farrar, Straus and Giroux.

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press

Guichet, C., Achard, S., Mermillod, M., & Baciu, M. (2026). Cognitive flexibility and brain network energy in healthy aging: An allostatic perspective from the SENECA model. *Imaging Neuroscience*, 4. https://doi.org/10.1162/IMAG.a.1091

Haueis, P. (2014). Meeting the brain on its own terms. *Frontiers in Human Neuroscience*, *8*, 815. https://doi.org/10.3389/fnhum.2014.00815

Haueis, P. (2018). Beyond cognitive myopia: A patchwork approach to the concept of neural function. *Synthese*, *195*(12), 5373–5402.

Haueis, P. (2021). Multiscale modeling of cortical gradients: The role of mesoscale circuits for linking macro- and microscale gradients of cortical organization and hierarchical information processing. *NeuroImage*, *232*, 1–11. https://doi.org/10.1016/j.neuroimage.2021.117846

Haueis, P. (2022). Descriptive multiscale modeling in data-driven neuroscience. *Synthese*, *200*(2), 129. https://doi.org/10.1007/s11229-022-03551-y

Haueis, P. (2024). A generalized patchwork approach to scientific concepts. *The British Journal for the Philosophy of Science*, *75*(3), 741–768. https://doi.org/10.1086/716179 (originally published 2021).

Haueis, P., & Kästner, L. (2022). Mechanistic inquiry and scientific pursuit: The case of visual processing. *Studies in History and Philosophy of Science*, *93*, 123–135. https://doi.org/10.1016/j.shpsa.2022.03.007

Haueis, P., & Margulies, D. S. (2026). Developing concepts for neuroscience: A philosophical toolkit. *European Journal of Neuroscience*, *63*(2), e70403. https://doi.org/10.1111/ejn.70403

Haugeland, J. (1985). *Artificial intelligence: The very idea*. MIT Press.

Hechler, A., de Lange, F. P., & Riedl, V. (2023). Energy savings across the cortex for confidently predicted visual input. *bioRxiv*, 1–39. https://doi.org/10.1101/2023.12.08.570804

Hochstein, E. (2017). Why one model is never enough: A defense of explanatory holism. *Biology & Philosophy*, *32*(6), 1105–1125. https://doi.org/10.1007/s10539-017-9595-x

Holroyd, C. B. (2025). The controllosphere: The neural origin of cognitive effort. *Psychological Review*, *132*(3), 603. https://doi.org/10.1037/rev0000467

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences*, *79*(8), 2554–2558. https://doi.org/10.1073/pnas.79.8.2554

Hu, Y., Wang, X., Liu, K., Xiong, S., Cai, B., Guan, J. S., Gu, M., & Xie, H. (2025). Polarized ATP synthase in synaptic mitochondria induced by learning and plasticity signals. *Communications Biology*, *8*(1), 1618. https://doi.org/10.1038/s42003-025-08963-3

Jamadar, S. D., Behler, A., Deery, H., & Breakspear, M. (2025). The metabolic costs of cognition. *Trends in Cognitive Sciences*, *29*(6), 541–555. https://doi.org/10.1016/j.tics.2024.11.010

Karbowski, J. (2024). Information thermodynamics: From physics to neuroscience. *Entropy*, *26*(9), 779. https://doi.org/10.3390/e26090779

Kästner, L., & Haueis, P. (2021). Discovering patterns: On the norms of mechanistic inquiry. *Erkenntnis*, *86*, 1635–1660. https://doi.org/10.1007/s10670-019-00174-7 (originally published 2019).

Klein, C. (2018). Mechanisms, resources, and background conditions. *Biology & Philosophy*, *33*(5), 36. https://doi.org/10.1007/s10539-018-9646-y

Laudan, L. (1978). *Progress and its problems: Towards a theory of scientific growth*. University of California Press.

Levy, W. B., & Calvert, V. G. (2021). Communication consumes 35 times more energy than computation in the human cortex, but both costs are needed to predict synapse number. *Proceedings of the National Academy of Sciences*, *118*(18), e2008173118. https://doi.org/10.1073/pnas.2008173118

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. MIT Press.

Mok, R. M., & Love, B. C. (2023). A multilevel account of hippocampal function in spatial and concept learning: Bridging models of behavior and neural assemblies. *Science Advances*, *9*(29), eade6903. https://doi.org/10.1126/sciadv.ade6903

Nave, K. (2026). Beyond control: Finding the purpose of enactive cognitive science. *Adaptive Behavior*, *0*(0), 1–21. https://doi.org/10.1177/10597123261435796

Peters, M. A. K., Michael Yoo, S. B., Klincewicz, M., Toyoizumi, T., Richards, B., Webb, T., & Lau, H. (2026). Looking to the brain to improve energy efficiency of AI. *Current Biology*, *36*(14), R821–R832. https://doi.org/10.1016/j.cub.2026.06.020

Piccinini, G. (2015). *Physical computation: A mechanistic account*. Oxford University Press.

Piccinini, G. (2020). *Neurocognitive mechanisms: explaining biological cognition*. Oxford University Press.

Polger, T.W., & Shapiro, L.W. (2016). *The multiple realization book*. Oxford University Press.

Rabinovich, M., Huerta, R., & Laurent, G. (2008). Transient dynamics for neural processing. *Science*, *321*(5885), 48–50. https://doi.org/10.1126/science.1155564

Ramsey, W. (2017). Must cognition be representational? *Synthese 194*(11), 4197–4214. https://doi.org/10.1007/s11229-014-0644-6

Sacco, P. L. (2026). The metabolic logic of cognitive biases. *Trends Open*, *00035*, 1–12. https://doi.org/10.1016/j.treopn.2026.05.012

Sengupta, B., Stemmler, M. B., & Friston, K. J. (2013). Information and efficiency in the nervous system—a synthesis. *PLoS Computational Biology*, *9*(7), e1003157. https://doi.org/10.1371/journal.pcbi.1003157

Šešelja, D., & Straßer, C. (2014). Epistemic justification in the context of pursuit: A coherentist approach. *Synthese*, *191*(13), 3111–3141. https://doi.org/10.1007/s11229-014-0476-4

Shagrir, O. (2010). Marr on computational-level theories. *Philosophy of Science*, *77*(4), 477–500. https://doi.org/10.1086/656005

Simon, H. (1962). The architecture of complexity. *Proceedings of the American Philosophical Society*, *106*(6), 467–482. http://links.jstor.org/sici?sici=0003-049X%2819621212%29106%3A6%3C467%3ATAOC%3E2.0.CO%3B2-1

Sprevak, M., & Smith, R. (2023). An introduction to predictive processing models of perception and decision-making. *Topics in Cognitive Science*, *00*(2023), 1–28.

Theriault, J. E., Katsumi, Y., Reimann, H. M., Zhang, J., Deming, P., Dickerson, B. C., Quigley, K. S., & Barrett, L. F. (2025). It's not the thought that counts: Allostasis at the core of brain function. *Neuron*, *113*(24), 4107–4133. https://doi.org/10.1016/j.neuron.2025.09.028

Theriault, J. E., Shaffer, C., Dienel, G. A., Sander, C. Y., Hooker, J. M., Dickerson, B. C., Barrett, L. F., & Quigley, K. S. (2023). A functional account of stimulation-based aerobic glycolysis and its role in interpreting BOLD signal intensity increases in neuroimaging experiments. *Neuroscience & Biobehavioral Reviews*, *153*, 105373. https://doi.org/10.1016/j.neubiorev.2023.105373

Van Rooij, I. (2008). The tractable cognition thesis. *Cognitive Science*, *32*(6), 939–984. https://doi.org/10.1080/03640210801897856.

Wang, Y., Haughey, N. J., Mattson, M. P., & Furukawa, K. (2004). Dual effects of ATP on rat hippocampal synaptic plasticity. *Neuroreport*, *15*(4), 633–636. https://doi.org/10.1097/00001756-200403220-00012

Weinberger, N. (2020). Near-decomposability and the timescale relativity of causal representations. *Philosophy of Science*, *87*(5), 841–856. https://doi.org/10.1086/710760

Williams, D. J. (2025). Two senses of medium independence. *Mind & Language*, *40*(4), 437–447.

Xue, K., Rezayat, F., Qi, T., Shen, L., Zhao, B., Huang, X., Marvin, J. S., & Ye, L. (2026). Learning shapes the energy cost of neural tasks. bioRxiv. https://doi.org/10.64898/2026.07.01.735889

Zhang, M., Chitic, R., & Bohté, S. M. (2025). Energy optimization induces predictive-coding properties in a multi-compartment spiking neural network model. *PLoS Computational Biology*, *21*(6), e1013112. https://doi.org/10.1371/journal.pcbi.1013112

https://doi.org/10.1017/S0140525X25103956 Published online by Cambridge University Press