Computational models in neurosciences between mechanistic and phenomenological characterizations

---

GAULD Christophe[1,2], BRUN Cédric[3], BORAUD Thomas[3,4,5], CARLU Mallory[6*], DEPANNEMAECKER Damien[6*]

[1] Department of Child Psychiatry, Université de Lyon, 59 Bd Pinel, 69 000 Lyon, France.
[2] UMR CNRS 8590 IHPST, Sorbonne University, Paris 1, France
[3] University of Bordeaux, UMR 5293, IMN, 33000, Bordeaux, France
[4] CNRS, UMR 5293, IMN, 33000, Bordeaux, France
[5] CHU de Bordeaux, IMN Clinique, 33000 Bordeaux, France
[6] Paris Saclay University, Institute of Neuroscience, CNRS, Gif-sur- Yvette, France

*Equally contributing last authors

Words: 5,813

© 2022 by the author(s). Distributed under a Creative Commons CC BY license.

Abstract (162 words)

Objective: Computational neuroscience combines mathematics, computer science models, and neurosciences for theorizing, investigating, and simulating neural systems involved in the development, structure, physiology, and cognitive abilities of the brain. Computational models constitute a major stake in translational neuroscience: the analytical understanding of these models seems fundamental to consider a translation towards clinical applications.

Method: We propose a minimal typology of computational models, which allows distinguishing between more realistic models (e.g., mechanistic models) and pragmatic models (e.g., phenomenological models).

Result: Understanding the translational aspects of computational models goes far beyond the intrinsic characteristics of models. First, we assume that a computational model is rarely uniquely mechanistic or phenomenological. Idealization seems necessary because of i) the researcher's perspectives on the phenomena and the purposes of the study (i.e., by the relativity of the model); ii) The complexity of reality across different levels and therefore the nature and number of dimensions required to consider a phenomenon. Especially, the use of models goes far beyond their function, and requires considering external characteristics rooted in path dependence, interdisciplinarity, and pluralism in neurosciences.

Conclusion: The unreasonable use of computational models, which are highly complex and subject to a shift in their initial function, could be limited by bringing to light such factors.

Keywords: Computing Methodologies; Computer Simulation; Models, Biological; Models, Theoretical.

Summations

- A detailed understanding of computational models seems fundamental to consider a translation towards clinical applications: a typology of computational models distinguishing their realistic part from their pragmatic part can be fruitful.
- The realistic part involves understanding the biological plausibility, the notion of validity of models, and the necessary internal and external approximations of computational models.
- The pragmatic part involves understanding the need to simplify computational models in order to reproduce behaviors, and to enable prediction and generalizability.
- The use of computational models brings at least three challenges that go beyond the internal understanding of the models themselves: i) concerning the historical anchoring inseparable from the models and the notion of path dependence; ii) concerning the necessary idealization of models when translating to practice; iii) concerning the necessary challenge of pluralism and sociological factors which lobbied during the translation of computational models.

Perspectives

- A simplistic conception of this typology (pragmatic / realistic) could lead to a naive view of computational neuroscience; conversely, qualifying these differences enable to support the inherent complexity of computational models.
- Highlighting external factors (path dependence, sociological factors, interdisciplinarity, etc.) influencing computational models does not necessarily improve their use and translational aspect – such findings should not undermine the inherent prudence and humility of this highly complex neuroscientific field.
- The knowledge of the various factors isolated in this article would need to be considered by the modelers as by the users of the computational models — programmatic horizon necessary to avoid a drift of uses, at the base of the theoretical stakes of translational neuroscience.

Introduction

Computational neuroscience is a relatively new but already extremely complex and dynamic field. Through its methodologies and applications mixed between dynamical systems, machine and deep learning analyzes, biological simulations, Bayesian statistics, the creation and use of computational models is one of its main challenges. Indeed, computational neuroscience developed models that can be reused for clinical applications (Varenne *et al.* 2018), and various other uses, e.g., for their descriptive power (Kording *et al.* 2020). Likewise, computational theories, for the most part resulting from the relevance of these models (Borsboom *et al.* 2020), are exported between different fields and applied in research programs that diverge in their foundations as in their objectives (Kuhn 1971). Thereby, neuroscience models are exported within ambitious projects such as the Blue Brain Project (BBP), initiatives that seek to develop infrastructure research projects exploring the human brain (Markram 2012). Despite minimum output, this project has been followed in Europe by the Human Brain Project (1.19 billion euros over ten years) (HBP now gives birth to a platform, EBRAINS, which includes a very important part of simulations, through tools such as The Virtual Brain) and in the USA by the BRAIN Initiative ($ 300 million per year for 10 years).

For methodological and ethical reasons, it seems essential to clarify the different concepts underlying these models involved in various areas of computational neuroscience, in order to increase their transparency, analyze whether they can be used with robustness outside their initial context of development, or explore their potential in terms of explanation, realism or pragmatism. Neuroscientists seem keenly aware of the social, ethical, and regulatory challenges of computational neuroscience, from potential threats to privacy until an understanding of the awareness and meaning of human and personal identity (Evers 2017). But it seems important that clarification of experiments, models, theories, and the whole scope of computational neuroscience comes from the researchers themselves involved in this scientific field – and such theoretical work has already been started in other biological fields (Blohm *et al.* 2020; Brette 2015, 2019; Haefner 2005; Pradeu *et al.* 2021).

Models are one of the main tools of computational neuroscience. A model is a material or formal construction used to represent a real-world object. It is an object of facilitating mediation that allows to transform reality or to anticipate it (Morgan & Morrison 1999; Varenne 2007). Minsky defines a model as an object A*, which is a model of an object A, inasmuch as an observer B can use A* to answer questions that interest her/him about A (Minsky 1965). Such a minimal definition enables to understand computational models as (mathematical) functions acting as mediators between the modeled object and a user or a modeler. Such mathematical models aim to reproduce a certain number of brain observable phenomena, functions, or dynamics. However, they cannot reproduce all the aspects of a brain function or a structure, unless being the considered brain itself. Given the significant complexity of the objects studied in computational neuroscience, this mediator value represented by the model depends on the question asked by the researcher. Indeed, as we will detail later, the choice of the model should be clearly conceived according to the question which must be solved – influencing the number of dimensions introduced in a specific model. For instance, single neuron models (*neudels*) like Izhikevich's models cannot fully account for a biological neuron (Izhikevich 2003). Indeed, they do not have a whole biophysical representation, but show apparent realistic dynamics (Izhikevich & FitzHugh 2006) by modeling behaviors similar to the time evolution of the membrane potential of a biological neuron. In other words, some of these models enable the prediction of neuronal behavior without seeking to exactly reproduce the biological parameters of the (human) brain. These examples are from a specific type of model, dynamical systems, based on differential equations. In the first section of this work, we mainly consider and refer to this category of model and open to a broader definition in the remaining sections.

Aims of the Study. A typology of models might be useful to discuss the relationship between the idealization of the model and its different applications. In this article, we will see that such a typology could explain the use, generalization and appropriation of a model by a scientific community.

Material and methods

In this perspective, we aim to propose a minimal typology of computational models. For this purpose, we use a set of inclusion criteria according to the SANRA guidelines (Scale for the Assessment of Narrative Review Articles) (Baethge et al. 2019):

- identify articles discussing the translational aspect and the use of computational models, and more particularly models of biological neurons;
- English or French language studies published in peer-review journals are used, especially by using MEDLINE (– 2021) and Cochrane (– 2021) databases;
- for all identified articles published, reference lists were also scanned to see whether we had missed any articles suitable for inclusion. These steps allowed us to reduce the possibility of non-indexed studies being missed. Single case studies, commentary, opinion articles, unpublished studies, conference posters and reviews and meta-analyzes were allowed.

Finally, the full text of the chosen literature was manually reviewed to determine inclusion into this perspective.

Results

This perspective allows to distinguish between realistic models (e.g., mechanistic models) and pragmatic models (e.g., phenomenological models).

1. A framework for computational models

Three essential characteristics of a computational model can be distinguished: a level of idealization ("modeling involves a selection of characteristic features of the modeled target"), a level of relativity ("the representation is oriented by tools and objectives"), and reversibility ("such models serve both as a representation of reality and as a support for its future modification") (Potochnik 2017). Indeed, all computational models have a level of idealization. Their relativity to technical tools can be variable, but it remains low compared to many other scientific fields (e.g., in medicine). They are highly reversible because they support research in the present time while serving as a support for its future modifications.

In accordance with the classical theorization of models (Hempel 1965), computational models have at least two main components: a realistic component and a pragmatic component. Indeed, it could be useful to distinguish the realistic part of models, used in order to facilitate mechanistic designs of a scientific object (e.g., the wooden model of an airplane) or facilitate its intelligible presentation (e.g., the double helix of DNA presented by Watson and Crick) (Varenne *et al.* 2018). On the other hand, the pragmatic part of computational models could be used to determine what type of response is expected in terms of action and behavior (e.g., the paper plane or an interdisciplinary model of pandemic management). The pragmatic part of the models aims to reproduce the target behavior for which the model was designed. Thereafter, we will detail these two components of the models, classically called the realistic and pragmatic components.

1.1. Mechanistic models (with a strong realistic component)

Models with a strong realistic component could be based on direct description of actual biological components of the brain. Most of them seek to reproduce neural mechanisms. In this approach of computational models, a behavior is prescribed by an equation through functional relations between components. This equation aims at reproducing dynamics of neural correlates at different scales.

Although the definition of a mechanism has been extensively discussed, we retain the definition of a mechanism as a network of interrelated parts, each performing its own functions, that are combined in such a way that each contributes to producing a behavior (Bechtel & Richardson 2010). For example, they can be useful in providing a causal explanation for a phenomenon. Such models are therefore mechanistic because they can be decomposable and offer a functional definition of their relations (Bechtel & Richardson 2010), allowing the model to have a biological correspondence, i.e., the relationship between variables is specified in terms of biological processes. In other words, they have a strong biological plausibility in terms of biophysics. The plausibility of a mechanistic model can be thus measured by the ability of the experimenter to identify in the physical system parts organized in a manner analogous to or conforming to the model. This realistic component and biological plausibility are intrinsically related to the notion of validity of a model, which can itself be understood in terms of validity: i) apparent, corresponding to the overall resemblance of the model to the manifest phenomenon; ii) predictive, corresponds to the power of the model to predict; iii) discriminating, corresponding to the relevance and representativity of its content; iv) structural, corresponding to the explanatory power of the model (Belzung & Lemoine 2011).

However, describing biology does not necessarily lead to good predictions, or may even limit them. In addition, the number of parameters within a computational model could be very large and cannot be modeled. And moreover, the more there are free parameters, the lower the confidence in the model is. Thus, it is often necessary to identify what are the minimum “ingredients” (biophysical variables and parameters) that are necessary to provide biological plausibility. Thus, in practice, purely mechanistic models are not possible, given the necessary consideration of a phenomenological part.

Any model that seeks to faithfully reproduce reality should, by definition, make approximations of two types. The first type of approximation is internal (Batterman & Rice 2014). For modeling, computational, temporal and tractability reasons, a computational model (e.g., a neudel) cannot model all the parameters of a real brain structure (e.g., a neuron), and therefore approximates some parameters, i.e., performs an idealization (e.g., with some ion channels) (Elgin 2017). The second type of approximation is external. Whenever a mechanistic model wishes to integrate part of the environment, it must simplify it to its limits.

These two impediments of mechanistic models (limits to prediction and internal and external approximations) lead in computational neurosciences to the development of models with a strong pragmatic component, such as the phenomenological models that we will describe.

1.2. Phenomenological models (with a strong pragmatic component)

Models with a strong pragmatic component describe simplified patterns of data, unlike mechanistic models which rather aim to grasp causes or processes (and which therefore have a higher biological plausibility due to the structural analogy with the biology of the target phenomenon). In this approach, a scale of granularity of the nervous system (e.g., a neuron) enable to describe emerging properties at the scale of physiology or behaviors.

For example, models with a strong pragmatic component can be useful for prediction (rather than explanation, as is the case with mechanistic models). These phenomenological models do not deliberately seek (or little) biological correspondence between their parameters and the biological parameters (Hilborn & Mangel 1997). Instead, they serve to facilitate biological understanding by seeking to reproduce a behavior.

For example, Leaky Integrate and Fire (LIF) models do not attempt to combine all the ion channels of a neuron. Such a model captures a global aspect of the membrane by mimicking the sub-threshold integration of the inputs, and the spiking phenomenology is a way to perform simulations by circumventing the difficulty associated with the modelisation of a large number of biological parameters. Moreover, the behavior of high-dimensional systems of nonlinear differential equations is difficult to visualize or analyze beyond a two-dimensional system which can be more easily studied by means of a phase plane analysis. Reducing the four-dimensional equation of Hodgkin Huxley (HH) (Hodgkin & Huxley 1952) to two equations is possible without reducing the biological elements considered (Depannemaecker *et al.* 2021b). This reduction leads to a loss in terms of dynamics (i.e., time evolution) but not in the number of elements considered. Another approach is to reduce it to its core dynamics. In this case, the model will be the simplest equation that keeps the properties of the phase-space to reproduce only the considered phenomenon, as the FitzHugh-Nagumo model does (Gerstner *et al.* 2002; Izhikevich & FitzHugh 2006). In these examples, the HH model considers biologically plausible underlying mechanisms to capture membrane excitability, while the FitzHugh-Nagumo model only keeps the minimal elements necessary to exhibit excitable properties comparable to neural excitability, but without giving any biophysical description. Thus, the interest of such phenomenological models is to enable prediction and generalizability within simplified models, and not to inherently imitate the physical parts of the mechanism of real biological neurons. This prediction and this generalizability, referring to the pragmatic part of the model, are commonly called the precision of the model (Massoud *et al.* 1998).

1.3. Interactions between mechanistic and phenomenological computational models

A model of a wooden plane will seek to reproduce the features of the real airplane. A wooden airplane model has a strong realistic component and is mechanistic due to the structural analogy with the shape of the target phenomenon. However, such a model flies much worse than a paper plane model, which however does not have the same shape as the real airplane, nor does it have a propeller or wheels (Brette 2012). The paper plane is an illustration of a phenomenological model, which has a strong pragmatic component. It mimics the functions and purposes for which the actual airplane was created (fly). We recognized in this distinction Marr's levels applied to modeling: the phenomenological model responds to the Level 3 (Computational), while the mechanistic model corresponds to the Level 2 (Algorithmic). Level 1 corresponds to the (organic) implementation (Bickle 2015).

In the same way, models such as LIF are very unrealistic while multicompartmental Hodgkin-Huxley (HH) models have ion channels that reproduce the real biological neuron (Brette 2012; Nelson & Rinzel 1995). However, as described in the HH model, the sodium activation variable has no biophysical equivalent in the biological sodium channel: therefore, even though a biophysical parameter is integrated into the model, there is a pragmatic dimension within the model. Various other models in computational neuroscience have both the pragmatic component of phenomenological models and the realistic component of mechanistic models. Some of them have even been designed specifically to balance this distinction between realism and pragmatism, such as AdEx models (Adaptive exponential integrate-and-fire models) which are phenomenological models whose hybridization increases their predictability and generalizability (Touboul & Brette 2008).

To go further, we claim that any mechanistic model should embed a phenomenological aspect which tie it to real life observations.

On the other hand, we also state that any phenomenological model is built on mechanistic relations between entities, but these entities do not directly relate to observables. For example, an unrealistic but very pragmatic biological neuron model, such as LIF (Brunel & van Rossum 2007), allows, depending on the input current, to make predictions on the membrane potential and the temporality of the spikes.

Indeed, in case of idealization of the model during its generalization, the comparison with the biological components allows to revise the hypotheses by adding or deleting the composition of the mechanism. The need for phenomenological models emanates from the difficulty to separate the contributions of a large number of components and mechanisms. This said, it must be stressed that commensurability between the level of description (scale and type) and the phenomenon to be explained, is necessary to give a complete explanation at this level, i.e., a necessary and sufficient relation between the elements considered and the observed phenomenon. This can be called a stabilized level of description. However, it is rarely the case in neurosciences, as we are in the presence of a very high-dimensional complex system, which makes it generally difficult to know what levels of description (i.e., nature and number of components) should be chosen to explain a given phenomenon. In other words, idealization seems necessary because: i) of the researcher's perspectives on the phenomena and the purposes of the study (i.e., by the relativity of the model); ii) of the messy nature of reality and ignorance of the number and nature of the dimensions of the phenomenon to be considered.

To sum up, any model necessarily has a mechanistic part and a phenomenological part. However, such a theoretical characterization depends on the triple relativity of the scientific field, of the researcher's question and of the temporality in which computational models are used.

Discussion

1. Computational models and path dependence

Based on the identification of the two types of models widely confused, we broaden the understanding of these computational models to embed them in their research environment, in relation to their historical path dependence. Path dependence explains the continued use of a model based on historical preference or current use: a model may persist even if newer and more efficient alternatives are available, for instance because it is easier or more cost-effective to continue along an already set path than to create an entirely new one (Pierson 2000).

We propose to identify two possible path dependencies: the polysemy of the term “computation” and a precocious occurrence of theoretical questions about computational neuroscience in the history of cognitive sciences.

First, the term of computation has evolved over time. In the scientific literature of cognitive science, we track down a confusion between the term “computational” and “computationalism”. The former is applied to a particular science (such as brain science) and seems more restricted than the term of computationalism. Indeed, the term “computational” refers only to the processing of information allowing the performance of the function (hence its association with the term “functionalism”) (Marr 1982). However, there may be different mechanisms underlying this realization of a function (e.g., the flight of a bird depends on the flapping of wings, but the bird can also hover). For Marr (1982/2010), the level of the computation is the level of the function, while the level of the mechanism is both the level of the rules which allows this function (i.e., algorithm level) and the level of the structural characterization of the system (i.e., implementation level), which are well separate. However, computational sciences bear a slightly different sense of “computational”, by sitting in between Marr’s levels, as they generally aim at explaining functions by algorithms.

In a completely different field, computationalism is a form of cognitivism: it refers to the processing of information, via different methods, which can be especially connectionist. Such terminological confusion has led to a hermeticity of the fields of study of computational neuroscience (Aizawa 2010).

Secondly, theoretical questions about computational aspects of neurosciences emerged very early in the history of cognitive sciences. Computational neuroscience was born with models of biological neurons, with a seminal article published in 1907 (Brunel & van Rossum 2007; Lapicque & Lapicque 1907), introducing dynamical systems theory within neurosciences, on the assumption that neuronal spiking was the support of in the peripheral nervous system. We can trace back to the first Macy Conference in 1941 the need to promote communication between scientific disciplines in an attempt to "restore the unity of science", a main objective of scientific positivism. This interest in so-called cybernetics and the presentation by Turing of a model of cognition as an embedded mechanism of treatment of information constrained by a hardware (Turing 1950) has led to limit the modeling of complexity, whether at the neuronal or mechanical level (Hevern 2012; Nagel & Newman 2008).

The distinctions between the computational models currently used in different fields of neurosciences must therefore be understood through their history, which itself influences the different fields currently involved in the use of these models.

2. Idealization and pluralisms in computational neuroscience

Computational models derive their robustness from coherent interactions between the three fields of physics, mathematics and biology. However, as in any science, there are limitations in the use of such models which do not come from the intrinsic model, but from its use and extrapolation, and therefore involve notions like those of idealization and abstraction (Bickle 2006; Sober 1999).

In many areas of the brain, neurons seem organized in populations of units which share some similar properties. Prominent examples are columns in the somatosensory and visual cortex (Hubel & Wiesel 1962; Mountcastle 1957) and pools of motor neurons (Frotscher 1996). Given the large number of neurons within such a column or pool, one possible approach is to describe the mean activity of the neuronal population, rather than the spiking of each individual neurons (Abbott & van Vreeswijk 1993; Amit & Brunel 1997; Brunel & Hakim 1999; Gerstner *et al.* 1993). The population activity equations allow study signal transmission, neural coding, oscillations and synchronization as well as the formation of activity patterns in spatially structured populations (Kähne *et al.* 2017). Moreover, density equations allow integration of different internal states (Knight 1972), with the stochastic spike triggering (Abbott & van Vreeswijk 1993; Brunel & Hakim 1999) or the state of refractoriness (Wilson & Cowan 1972). For instance, for the Wilson–Cowan model (describing the dynamics of interactions between populations of simple excitatory and inhibitory neuron models), the validity of these population equations relies on stringent assumptions such as homogeneous populations, absence of finite size effects and adaptation. In this type of model, all the dynamical complexity that may exist at the lower scale is eliminated. Of course, these limits can be smoothed. For example, the variability of a parameter between one neuron and the next is often replaced by slow noise in the parameters. In the same vein, adaptation could be phenomenologically corrected by introducing a slow variable that integrates over the population activity in the past (Knight 1972). However, some deeper questions resist any such correction.

Thus, in the case of neural mass and neural field models, the various tools mentioned above, specific to averages and populations of neurons (e.g., a transfer function which designates which output is produced as a function of the input), are introduced in the computational model. But such an introduction needs to question if the effects observed within the model are observed for the same reasons as the effects observed within the real biological neuron.

The causes that govern the behavior of the model are not necessarily the causes that govern the behavior of the biological neuron. In other words: is the causality introduced into the model similar to the biological causality, similarity necessary for extrapolation? For instance, doesn’t the airplane metaphor raise strong doubts that the causes that make the paper plane fly have anything to do with the causes that make the real plane fly?

This deep question calls the legitimacy of translating a concept to build a model: does the model equation contain the same set of concepts as the object being studied? Moreover, in the case where the validation of the models is performed through simulations, the observed effect looks similar to the real effect – but it cannot be proven that the causes of these effects are similar. The same phenomenon is reproduced, without any intuitive or mathematical argument allowing to claim that the cause of this observed behavior is similar to the real cause. It is true that when a pure mechanistic single biological neuron is studied, the equations provide causal transparency. When scaling up, i.e., during the passage from one scale of description to another, it is not clear that causality is maintained.

This observation of the same phenomenon related to multiple underlying causal mechanisms is called “multirealizability” (or multiple realizability). This could be understood as a generalization to models of “degeneracy” observed in biology. Multiple realizability concerns the difficulty in relating the behavioral function of a model with structural or organizational properties (e.g., flying can be achieved in multiple ways, by flapping its wings or by soaring) (Polger & Shapiro 2016).

2.1. Multirealizability and scales of explanations

Multiple realizability may appear in particular in the very specific case of scaling up. To continue with the example of biological neuron models, multirealizability appears when the model is translated from a scale (which was used to build the model) to another, e.g., from neudels to neural network. A macroscopic phenomenon (e.g., epilepsy) can be explained by different causal mechanisms evolving at several scales (from genetic to interconnectivity network) having separate temporal scales (temporal dissection) (Depannemaecker *et al.* 2021a). Therefore, multirealizability requires exploring (or even integrating) different scales (from microscopic to macroscopic scales). Thus, multirealizability necessarily appears when biological neurons models are integrated in a process of scaling. However, knowledge of the structure of a phenomenon is not necessary to highlight, because proxies may be sufficient to draw relevant inferences (e.g., Newton did not need to know the structure of water to describe the phenomenon of tides, but only its coefficient of viscosity and its density).

We said that the injection of phenomenology into a computational model (e.g., AdEx) allows to compensate for what the model does not intrinsically consider. Likewise, the scaling process within the framework of biological neural networks requires reintroducing an approximation, by the necessary addition of a part of phenomenology in the model. Take the example of hyper-dissipation in meteorology: because researchers do not have infinite precision, they do not have access to what is happening under a minimum size scale. As such minimal scales cannot be resolved, it is necessary to add a parameter to the equation. This phenomenological injection corresponds to the “edges” of the model, defined as the impossibility of integrating realistic parameters during the complexification of the model. Scaling up takes the model away from biological plausibility by increasing its phenomenological part (i.e., scaling up necessarily increases the phenomenology part of the model since there are more external parameters to consider). The risk of such a scaling up is therefore to neglect this phenomenological injection, which could lead to interpretation bias.

In this context, we could however cite the existence of causal modeling, which seems to offer a way out of the problems posed by multiple realizability. Such models provide probabilistic or deterministic versions to overcome the effects of multiple realizability from an interventionist perspective. In practice, such models lose in explanatory power what they gain in intervention capacity (Pearl 2009, 2010).

2.2. ***Idealization and pluralism***

The multiplicity of physical and algorithmic implementations of a lower level at a higher level challenges the connection and validity between levels. Epilepsy or depression, i.e., macroscopic phenomena, could not be accurately modeled in the context of computational neuroscience: thus, even if scientific progress is possible thanks to computational models, such models do not refer only to a single higher level. However, the problem of multirealizability can be approached in a more fruitful way for computational science than by raising the problem of reductionism.

Indeed, debates about levels of explanation in neurosciences demonstrate that the field is increasingly departing from the classical hierarchical models in which a fundamental physical or algorithmic level is deemed the only truly explanatory level, and which allows all higher levels of a complex system's organization have to be reduced to it. Thereby, explanatory pluralism is the thesis according to which the explanation of a phenomenon can especially refer to constructs or mechanisms belonging to several different scales (Cartwright 1979; Hacking 2002; Suppes 1977). There are two forms of explanatory pluralism: integrative pluralism (Mitchell 2009) and tolerant or non-integrative pluralism (Kellert *et al.* 2006). By studying these two forms of pluralisms, we will see that the question of integration between scales leads to the question of the necessary interdisciplinarity of neurosciences.

3. Integrative and tolerant pluralism for computational neuroscience

Integrative pluralism attempts to establish small local integrations between levels of analysis, without seeking to build a large theoretical structure. In the case of biological neural networks modeling a seizure disorder, there are many factors acting at different times and levels of the simulation of the system (or pathology) – such as nonlinearity or noise. Neither the model itself, independently of these factors, nor the environment external to the model are sufficient on their own to explain its behavior. There is not a single composition of causes involving different levels that will do the job in all cases (O'Malley *et al.* 2014).

In other words, certain factors and certain levels are more important than others. These factors at certain levels are considered more important than others are called producers of differences (Kendler 2012). For example, epilepsy can be caused by several producers of differences on at least three different levels: genetic (genetic variants influence the onset of epilepsy and / or the onset of a seizure), neuronal (hyper-synchrony as it can be modeled by biological neural networks) and environmental (trauma, dietary factors, disruption of circadian rhythms, etc.). All of these factors are producers of differences because they are involved in the development and maintenance of an epileptic disorder. But depending on the interindividual heterogeneity or different types of epilepsy, causal signatures are different, i.e., differences makers seem to be more or less concentrated at certain levels in some cases while in other cases they could be distributed on all levels (Woodward 2003).

Therefore, causal signatures designate certain sets of concentrated levels which best explain a pathology. For example, a West syndrome or an epileptic encephalopathy would have more producers of differences in the genetic domain, and therefore has a genetic causal signature. A generalized tonic-clonic epilepsy would be better explained at the neuronal level and a temporal epilepsy has a causal signature based on developmental mechanisms, where neuronal plasticity plays major roles in creating a state of underlying hyperexcitability (Steinlein 2004). These findings are even more pronounced in psychiatric disorders, where environmental factors, for example, represent an important causal signature in depression while they are much less important in schizophrenia. However, this kind of integrative pluralism does not seem relevant to understanding the interest of models of biological neurons, an interest which would be restricted to certain macroscopic phenomena. On the contrary, tolerant (or non-integrative) pluralism seems much more fruitful for understanding biological neural networks.

The tolerant pluralism postulates that certain phenomena at different levels may be required to answer different questions, without these levels necessarily being integrated into a coherent whole (Van Bouwel 2011). The choice of a level of explanation depends on the epistemic and pragmatic interests of the researcher (Kendler & Parnas 2017). Some levels are more important in terms of explanatory strength, generalizability to several phenotypes, specificity or manipulability in terms of therapeutic action. For example, biological neural networks are undoubtedly the best model in the current state of science to represent neural synchronization in terms of strength, generalizability, specificity, even manipulability, because despite all their approximations (necessary in any model), they allow to obtain with confidence a representation of neurons behavior.

With a view to translational research and according to the logic of tolerant pluralism, the transfer of these models to clinical practice (i.e., diagnosis based on semiology and therapeutic management) should be carried out carefully. The best level of explanation for epilepsy (or any other neurological or psychiatric pathology) will depend on the question of the researcher (or the research and medical community), and there is no evidence to intuitively say that the questions answered by biological neural networks (including hyper-synchrony) will be relevant, for example, to the clinician. At least, it will be up to the researcher building models of biological neurons to convince other research communities of the relevance and applicability of his/her findings to the fields of application to which she/he is addressing.

4. Sociological factors

Finally, the adoption of a tolerant pluralism, which accepts to conceive of multiple explanatory levels but retains only the most relevant for a given research question, leads to the question of interdisciplinarity. Often, a scientific result (e.g., biological neural networks) in a given context have to be translated to another level (e.g., to a pathology or to a larger integrative project such as the Human Brain Project).

It seems important to consider sociological constraints that follow these scientific results. Each result engages the conviction of a researcher that his/her hypothesis, which has been proven by experimentation, can become a hypothesis transferable to another discipline or another project (Callon 1986). In this way, science studies and social epistemology has developed frameworks, e.g., theories of translation and regimes of promise (Joly 2015). For instance, for a result to be accepted by the research community and reach a larger audience, it is not so much the proof of its effectiveness for new discoveries in this field that is important (necessary, but not sufficient), but its credibility and legitimacy. Three types of legitimacy are described: charismatic, tradition-based, and rational (Kim 2020). Credibility depends on factors such as the plausibility of the model in the scientific landscape, the verification procedures or its pragmatic aspect (Blok & Jensen 2011). Note that this transferability hypothesis can also be misused, as it is not impossible that a researcher, relying on his/her legitimacy and credibility, extrapolates non-transferable results in other disciplines or projects.

Computational neuroscience could be thus analyzed as a network of social and political relationships. Scientific controversy constitutes an important stage in the validation of a scientific object (Pestre 2006): a model will not necessarily prevail because it is close to reality, but because it is embedded in a context and a scientific community (Latour 2005). Obtaining an experimental result would not be enough to end scientific debates. On the contrary, such a result will be developed in a scientific community during an acculturation to practices and a concrete use of this result, through the exchange of arguments in the midst of controversies (Raynaud 2017). Consensus on a result would therefore be more the result of controversy than of a formal explanation isolated from any context. Therefore, in the scientific landscape, the contingency of discoveries depends on the acceptance of a computational model in a given context, at a given time and for a given community.

5. Conclusion

Computational neurosciences are concerned with analyzing models, providing a hierarchy of scientific concepts embedded in a scientific dynamic. A minimal typology of computational models allows distinguishing between more realistic models (i.e., mechanistic models) and pragmatic models (i.e., phenomenological models). A model is rarely (if ever) uniquely mechanistic or phenomenological but has a part of each. When designing the model, or when using it, the cursor should be placed to best answer a particular question. More precisely, the model must be optimized for a specific function towards addressing this question.

Questions about idealization and scientific pluralism are related hot topics. Indeed, the passage from one scale of description to another for a computational model requires discussing the interaction between scales, a discussion that can be conducted through scientific pluralism. Pluralism proposes to consider interdisciplinarity in computational neuroscience. Adopting a tolerant version of pluralism leads to conceiving a hierarchy of the epistemic relevance of a discovery. This prioritization necessarily involves unscientific factors, leads to a discussion of interdisciplinarity and involves sociological aspects which condition the success of some computational models among many others. Computational neurosciences, based on mathematics, seem to offer a common language to different sciences and thus facilitate their empirical, conceptual and sociological interactions. However, the presence of a theoretical framework guiding the construction of the models entails a path dependence of these models, at the risk of biasing the statement necessarily oriented by this framework.

Authors contributions:

GAULD Christophe: Writing, Original draft preparation, Conceptualization.

BRUN Cédric: Conceptualization, Methodology, Supervision.

BORAUD Thomas: Supervision, Reviewing, Validation.

CARLU Mallory: Supervision, Methodology, Reviewing, Resources, Validation.

DEPANNEMAECKER Damien: Supervision, Methodology, Reviewing, Resources, Validation.

Acknowledgements: Any

Financial support: Any

Statement of interest: No competing interests

Abbreviations

AdEx: Adaptive Exponential integrate-and-fire

HH: Hodgkin-Huxley

LIF: Leaky Integrate and Fire

BBP: Blue Brain Project

TVB: The Virtual Brain

HBP: Human Brain Project

Reference list

Abbott LF and van Vreeswijk C (1993). Asynchronous states in networks of pulse-coupled oscillators. *Physical Review E*, 48(2), 1483–1490.

Aizawa K (2010). Computation in cognitive science: it is not all about Turing-equivalent computation. *Studies in History and Philosophy of Science Part A*, 41(3), 227–236.

Amit DJ and Brunel N (1997). Dynamics of a recurrent network of spiking neurons before and following learning. *Network: Computation in Neural Systems*, 8(4), 373–404.

Baethge C, Goldbeck-Wood S, Mertens S (2019). SANRA—a scale for the quality assessment of narrative review articles. *Research Integrity and Peer Review*, 4:5.

Batterman RW and Rice CC (2014). Minimal Model Explanations. *Philosophy of Science*, 81(3), 349–376.

Bechtel W and Richardson RC (2010). *Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research*, MIT Press.

Belzung C and Lemoine M. (2011). Criteria of validity for animal models of psychiatric disorders: focus on anxiety disorders and depression. *Biology of Mood & Anxiety Disorders*, 1(1), 9.

Bickle J (2006). Multiple Realizability. In *Encyclopedia of Cognitive Science*, American Cancer Society. doi:10.1002/0470018860.s00116

Bickle J (2015). Marr and Reductionism. *Topics in Cognitive Science*, 7(2), 299–311.

Blohm G, Kording KP and Schrater PR (2020). A How-to-Model Guide for Neuroscience. *ENeuro*, 7(1). doi:10.1523/ENEURO.0352-19.2019

Blok A and Jensen TE (2011). *Bruno Latour: Hybrid Thoughts in a Hybrid World*, Routledge.

Borsboom D, Maas H, Dalege J, Kievit R and Haig B (2020). *Theory Construction Methodology: A practical framework for theory formation in psychology*, PsyArXiv. Retrieved from https://psyarxiv.com/w5tp8/

Brette R (2012). What is computational neuroscience? Retrieved from http://romainbrette.fr/what-is-computational-neuroscience-i-definitions-and-the-data-driven-approach/

Brette, R (2015). Philosophy of the Spike: Rate-Based vs. Spike-Based Theories of the Brain. *Frontiers in Systems Neuroscience*, 9. doi:10.3389/fnsys.2015.00151

Brette, R (2019). Is coding a relevant metaphor for the brain? *Behavioral and Brain Sciences*, 42. doi:10.1017/S0140525X19000049

Brunel N and Hakim V (1999). Fast Global Oscillations in Networks of Integrate-and-Fire Neurons with Low Firing Rates. *Neural Computation*, 11(7), 1621–1671.

Brunel N and van Rossum M (2007). Quantitative investigations of electrical nerve excitation treated as polarization: Louis Lapicque 1907 · Translated by: *Biological Cybernetics*, 97, 341–349.

Callon M (1986). Éléments pour une sociologie de la traduction : La domestication des coquilles Saint-Jacques et des marins-pêcheurs dans la baie de Saint-Brieuc. *L'Année Sociologique (1940/1948-)*, 36. Retrieved from https://www.jstor.org/stable/27889913

Cartwright N (1979). Causal Laws and Effective Strategies. *Noûs*, 13(4), 419–437.

Depannemaecker D, Destexhe A, Jirsa V and Bernard C (2021a). Modeling seizures: From single neurons to networks. *Seizure - European Journal of Epilepsy*, 90, 4–8.

Depannemaecker, D., Ivanov, A., Lillo, D., Spek, L., Bernard, C., & Jirsa, V. (2021b). *A unified physiological framework of transitions between seizures, sustained ictal activity and depolarization block at the single neuron level.* Retrieved from https://www.biorxiv.org/content/10.1101/2020.10.23.352021v2

Elgin CZ (2017). *True Enough*, 1st edition, Cambridge, MA: The MIT Press.

Evers K (2017). The contribution of neuroethics to international brain research initiatives. *Nature Reviews Neuroscience*, 18(1), 1–2.

Frotscher M (1996). Synaptic Transmission. In R. Greger & U. Windhorst, eds., *Comprehensive Human Physiology: From Cellular Mechanisms to Integration*, Berlin, Heidelberg: Springer, pp. 321–334.

Gerstner W, Ritz R and van Hemmen JL (1993). Why spikes? Hebbian learning and retrieval of time-resolved excitation patterns. *Biological Cybernetics*, 69(5), 503–515.

Gerstner W, Wulfram Kistler and M.W (2002). *Spiking Neuron Models: Single Neurons, Populations, Plasticity*. *Spiking Neuron Models: Single Neurons, Populations, Plasticity*. doi:10.1017/CBO9780511815706

Hacking I (2002). Historical Ontology. In P. Gärdenfors, J. Woleński, & K. Kijania-Placek, eds., *In the Scope of Logic, Methodology and Philosophy of Science: Volume Two of the 11th International Congress of Logic, Methodology and Philosophy of Science, Cracow, August 1999*, Dordrecht: Springer Netherlands, pp. 583–600.

Haefner JW (2005). *Modeling Biological Systems: Principles and Applications*, 2nd edn, Springer US. doi:10.1007/b106568

Hempel G (1965). *Aspects of Scientific Explanation*, Free Press.

Hevern VW (2012). Uexküll, J. von. In R. W. Rieber, ed., *Encyclopedia of the History of Psychological Theories*, New York, NY: Springer US, pp. 1143–1144.

Hilborn R and Mangel M (1997). *The Ecological Detective: Confronting Models with Data*, 1st edition, Princeton, NJ: Princeton University Press.

Hodgkin AL and Huxley AF (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of Physiology*, 117(4), 500–544.

Hubel DH and Wiesel TN (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology*, 160(1), 106-154.2.

Izhikevich EM (2003). Simple model of spiking neurons. *IEEE Transactions on Neural Networks*, 14(6), 1569–1572.

Izhikevich EM and FitzHugh R (2006). FitzHugh-Nagumo Model. *Scholarpedia*, 1(9), 1349.

Joly PB (2015). Le régime des promesses technoscientifique. In *Sciences et technologies émergentes: pourquoi tant de promesses?*, Hermann, Editeurs des Sciences et des Arts, p. np.

Kähne M, Sokolov IM and Rüdiger S (2017). Population equations for degree-heterogenous neural networks. *Physical Review E*, 96(5), 052306.

Kellert SH, Longino HE and Waters CK (2006). *Scientific Pluralism*, U of Minnesota Press.

Kendler K (2012). The dappled nature of causes of psychiatric illness: replacing the organic–functional/hardware–software dichotomy with empirically based pluralism. *Molecular Psychiatry*, 17(4), 377–388.

Kendler K and Parnas J (2017). *Philosophical Issues in Psychiatry IV: Psychiatric Nosology*, Oxford University Press.

Kim SH (2020). Max Weber. In E. N. Zalta, ed., *The Stanford Encyclopedia of Philosophy*, Winter 2020, Metaphysics Research Lab, Stanford University. Retrieved from https://plato.stanford.edu/archives/win2020/entries/weber/

Knight BW (1972). Dynamics of Encoding in a Population of Neurons. *Journal of General Physiology*, 59(6), 734–766.

Kording KP, Blohm G, Schrater P and Kay K (2020). Appreciating the variety of goals in computational neuroscience. *ArXiv:2002.03211 [q-Bio]*. Retrieved from http://arxiv.org/abs/2002.03211

Kuhn TS (1971). Notes on Lakatos. In R. C. Buck & R. S. Cohen, eds., *PSA 1970: In Memory of Rudolf Carnap Proceedings of the 1970 Biennial Meeting Philosophy of Science Association*, Dordrecht: Springer Netherlands, pp. 137–146.

Lapicque L and Lapicque M (1907). Recherches sur la loi d'excitation électrique. *J Physiol Path Gén*.

Latour B (2005). *Reassembling the Social: An Introduction to Actor-Network-Theory*, OUP Oxford.

Markram H (2012). The Human Brain Project. *Scientific American*, 306(6), 50–55.

Marr D (1982). *Vision: a computational investigation into the human representation and processing of visual information*, Cambridge, Mass: MIT Press.

Massoud TF, Hademenos GJ, Young WL, Gao E, Pile-Spellman J and Viñuela F (1998). Principles and philosophy of modeling in biomedical research. *The FASEB Journal*, 12(3), 275–285.

Minsky M (1965). Matter, Mind and Models. Retrieved from https://dspace.mit.edu/handle/1721.1/6119

Mitchell. M (2009). *Unsimple Truths: Science, Complexity, and Policy*, University of Chicago Press.

Morgan MS and Morrison M (Eds.). (1999). *Models as Mediators: Perspectives on Natural and Social Science*, 1st edn, Cambridge University Press. doi:10.1017/CBO9780511660108

Mountcastle VB (1957). Modality and topographic properties of single neurons of cat's somatic sensory cortex. *Journal of Neurophysiology*, 20(4), 408–434.

Nagel E and Newman J (2008). *Gödel's Proof*. (D. R. Hofstadter, Ed.), Revised ed. edition, New York: NYU Press.

Nelson M and Rinzel J (1995). The Hodgkin-Huxley model. *The Book of GENESIS: Exploring Realistic Neural Models with the GEneral NEural SImulation System*, 29–51.

O'Malley MA, Brigandt I, Love AC … Rohwer F (2014). Multilevel Research Strategies and Biological Systems. *Philosophy of Science*, 81(5), 811–828.

Pearl J (2009). *Causality*, Cambridge University Press.

Pearl J (2010). An Introduction to Causal Inference. *The International Journal of Biostatistics*, 6(2). doi:10.2202/1557-4679.1203

Pestre D (2006). *Introduction aux Science Studies*. Retrieved from https://www.cairn.info/introduction-aux-science-studies--9782707145963.htm

Pierson K (2000). Increasing Returns, Path Dependence, and the Study of Politics. *The American Political Science Review*, 94(2), 251–267.

Polger TW and Shapiro LA (2016). *The Multiple Realization Book*, Oxford, New York: Oxford University Press.

Potochnik A (2017). *Idealization and the Aims of Science*, University of Chicago Press.

Pradeu, T, Lemoine M, Khelfaoui M and Gingras Y (2021). *Philosophy in Science: Can philosophers of science permeate through science and produce scientific knowledge?*

Raynaud D (2017). *Scientific Controversies: A Socio-Historical Perspective on the Advancement of Science*, Routledge.

Sober E (1999). The Multiple Realizability Argument against Reductionism. *Philosophy of Science*, 66(4), 542–564.

Steinlein OK (2004). Genetic mechanisms that underlie epilepsy. *Nature Reviews Neuroscience*, 5(5), 400–408.

Suppes P (1977). Is Visual Space Euclidean? *Synthese*, 35(4), 397–421.

Touboul, J and Brette R (2008). Dynamics and bifurcations of the adaptive exponential integrate-and-fire model. *Biological Cybernetics*, 99(4), 319.

Turing AM (1950). Computing machinery and intelligence. *Mind*, LIX(236), 433–460.

Van Bouwel J (2011). What can democratic theory teach us about scientific pluralism, objectivity and consensus? In *Three rivers philosophy conference 2011 : science, knowledge, and democracy, Abstracts*. Retrieved from http://hdl.handle.net/1854/LU-1853227

Varenne F (2007). *Du modèle à la simulation informatique*, Vrin.

Varenne F, Silberstein, M and Dutreuil S (2018). *Modéliser et simuler: Epistémologies et pratiques de la modélisation et de la simulation*, Éditions Matériologiques.

Wilson HR and Cowan JD (1972). Excitatory and Inhibitory Interactions in Localized Populations of Model Neurons. *Biophysical Journal*, 12(1), 1–24.

Woodward J (2003). *Making Things Happen: A Theory of Causal Explanation*, Oxford University Press.