



# Structures in Physics and Neuroscience

## Structural Realism and the Unity of Science

Majid Davoody Beni<sup>1,2</sup> · Georg Northoff<sup>3,4,5,6,7,8,9</sup>

Received: 6 February 2020 / Accepted: 25 May 2020  
© Springer Nature B.V. 2020

### Abstract

We offer to extend structural realism into the field of mind and brain studies. The naturalised metaphysics of structural realism has been defined in terms of unification of sciences. The unification program has been carried out nicely in fields of Quantum Mechanics and General Relativity. But for the structural realist metaphysics to receive a recommendation, the unification program needs to be extended to the fields of especial sciences. Our aim in the paper is twofold. On the one hand, we present a structural realist theory of the brain that is well-supported by the best theories of computational neurosciences. On the other hand, and at the same time, we provide further reason for being optimistic about the structural realism metaphysical claim which is based on the idea of unification in sciences.

**Keywords** Structural realism · Unification · Difference-based coding · Neuronal structures · Sparse coding

## 1 Introduction

Structural realism (SR) is a respectable theory of contemporary philosophy of science. More precisely, SR is a philosophical thesis about the structural foundations of reality, and it has been developed into a number of different branches. There is Epistemic SR (or ESR), which holds that what we know about the world is its fine structure. Then there is Ontic SR (OSR), which holds that the structure is all that there is. We offer to extend SR into the field of mind and brain

---

Majid Davoody Beni and Georg Northoff have contributed equally to this paper

✉ Majid Davoody Beni  
davoody1980@gmail.com

Georg Northoff  
Georg.Northoff@theroyal.ca

Extended author information available on the last page of the article

studies. Although we avoid making a straightforward distinction between ontic and epistemic formulations in this paper, we see SR as conveying important ontological (metaphysical) theses. A restricted, scientifically informed, and naturalised sense of metaphysics is at issue here. Without agreeing to every aspect of Ladyman and Ross's (2007) defence of OSR, our conception of metaphysics is in line with their philosophical suggestion, according to which the chief metaphysical thesis of SR asserts the (possible) unification of the sciences. Indeed, there are good reasons for optimism as regards the viability of this claim. SR has been successful then and again at accounting for the unity of science across the history of classical physics and electromagnetic as well as quantum mechanics (Worrall 1989; Ladyman 1998; French and Ladyman 2003). More recently, SRists also went further to show how a philosophical construal of General Relativity (GR) could be integrated into this unified picture (Lam and Esfeld 2012a; Esfeld and Lam 2010). However, the agenda still suffers from certain limitations. For example, although the structuralist program unifies theories across different domains of scientific activity, they practically succeeded at establishing such unity mostly (if not only) within the borders of physics. Taking physics as the centrepiece of unification indicates that SR is committed to physicalism and reductionism, which means that SR cannot respect the autonomy of special sciences—such as biology, psychology, etc.—in accordance with its primary promises (see Ladyman 2008). Recently, there has been some unorthodox forms of SR that steer clear of the physicalist tendency of SR. Beni's (2019) *Cognitive Structural Realism* (CSR), for example, endeavours to rely on neuro-computational structures, rather than physical ones, for the purpose of regimentation of scientific representations. This is complemented by showing the explicitly neuro-ecological dimension of the brain as recently formulated in the concept of the world-brain relation (Northoff 2018). Specifically, the concept of world-brain relation means that, empirically, the brain aligns to the temporo-spatial structure of its environmental context. For instance, when we listen to music, our brain aligns its activity, i.e., its phases, to the phases, i.e., tones of the music; we entrain to the music rhythm and are in tune with it. The brain aligns to the world resulting in world-brain relation for which there is plenty of empirical evidence (Northoff 2018; Northoff and Huang 2017). While ontologically, world-brain relation entailing a neuro-ecological view of the brain requires relation-based ontology rather than property-based ontology (see Northoff 2018 for details). Such relation-based ontology can be well found in SR which is the main topic of this paper.

Our enterprise in this paper is in line with an unconventional statement of SR (as developed by Beni and Northoff on previous occasion), but orthodox versions of SR lean towards physicalism and reductionism all the same. For the unification program to be fulfilled, the theories of different domains of biology, neuroscience, and psychology have to be attached to the infrastructure that is allegedly lying at the foundation of reality. Ideally, the unification should take place without endangering the autonomy of special sciences. However, despite paying lip service to the autonomy of special sciences, orthodox versions of SR conceive of the foundation of reality in physicalist terms. Our enterprise in this paper aspires to fulfil an egalitarian version of unification, in the sense that it alleviates the reductionist

tendency of orthodox forms of SR. And although there have been some attempts at extending SR into the domain of the special sciences—such as biology (Lloyd 1994, 2015; French 2014), psychology (Hasselmann et al. 2010; Beni 2016, 2018; Northoff 2018), and economics (Ross 2008)—the fulfilment of the unification program leaves a lot to be achieved yet. We aim to respond to this call by taking structural realism beyond physics. The paper aims to map the road of the extension of SR into neuroscience, and in this way carry out the unificationist agenda of SR. In tandem, we relieve the reductionist tendency of the orthodox forms of SR even more and in this fashion support the developing Cognitive version of SR.

The paper consists of two main parts. In the first part, we will briefly review the support that SR has received from the fulfilment of the unification program in the field of physics. To do this, we outline SR and focus on Esfeld and Lam's attempt at reconciling philosophical interpretations of Quantum Mechanics (QM) and GR to one another through advocating a structural realist account of space–time.

In the second part of the paper, we draw attention to the structural features of the brain and its neural activity and build upon the relevant theories in the field to develop this observation into a structural realist account of the brain. Despite the schematic nature of our proposal, we substantiate our argument by providing a rather detailed presentation of the recent breakthroughs in contemporary computational neuroscience. We will specifically discuss the brain's coding strategy, i.e., how it codes and encodes its neural activity, and explain how the brain's spontaneous activity results in the formation of statistically-based spatiotemporal structures. The formation of statistically-based spatiotemporal structures is based on the way the brain encodes environmental stimuli, e.g., events or objects, into its own neural activity. We proceed to argue that the brain's main neuronal activity appears to be weaving the spatiotemporal patterns through a particular neural code that underpins the encoding of statistically-based differences (between different stimuli rather than the stimuli themselves). We claim that this model of the brain's activity paves the way to making a connection between our structuralist model of the brain and the variety of structural realism that has been advocated by Esfeld and others in the field of philosophy of physics. When construed in this fashion, the brain's neural code (as well as the brain's spatiotemporal structure that has been based on its neural code, e.g., difference-based coding) conforms to the structural realist understanding of the structural nature of reality. That is to say, the scientifically informed philosophical image of the brain is compatible with—or in a stronger statement, requires—being spelled out in terms of the non-eliminativist version of SR.

The outcome of our proposal is philosophically significant in two interconnected ways. On the one hand, it provides a reliable philosophical model of the brain that is inspired by (and is consistent with) recent breakthroughs in neuroscience. On the other hand, it provides all the more reason for embracing the ambitious metaphysical claim that underlies SR, according to which metaphysics consists in the unification in sciences.

## 2 Structural Realism and Its Metaphysics

SR, in its different branches, aimed at overcoming the historical pessimism about the verisimilitude of scientific theories. SR-theorists claim that the scientific findings, especially in fields of sub-particle physics, and—in the wake of Esfeld and Lam’s endeavours—GR, support SR more readily than the alternative entity realism. The thesis has been associated with strong metaphysical claims about the structural foundation of reality. This claims, however, are scientifically informed, and SR-theorists have heralded their reluctance to domesticate the scientifically evolved picture of the world to the habituated intuitions of the orthodox metaphysics. Metaphysics, according to OSRists, simply consists in *unifying* the outcome of scientific activity in different branches of scientific enterprise (Ladyman and Ross 2007, 2). This statement of OSR may indicate that the world solely consists of structures. The objection that may be raised is that where there is a structure there is the structured. Being aware of such problems, in developing our theory, we do not eliminate relata completely and lean toward a moderate form of OSR (we will explain this in the next section).

SR metaphysics has been constantly presented as consisting in the unification of the sciences. Here, the general insight is that “we have a unified world-view to the extent that we use a smaller rather than a larger number of argument patterns in science, and to the extent that what get used as schematic sentences in these argument patterns are themselves derived from other non-ad hoc argument patterns” (Ladyman and Ross 2007, 44). We adhere to this unificationist take on metaphysics in our enterprise in this paper, albeit without going so far to claim that metaphysics, *in toto*, consists in the unification of sciences (nor do we presume that unification mandates reduction to physics). Our reason for stating our unificationist tendency in a modest form is twofold. For one thing, we are interested in improving a specific metaphysical approach in the domain of philosophy of science, and we do not aim to provide a comprehensive definition of metaphysics in its entirety. Moreover, we do not assume that the unification of sciences consists in reducing special sciences to physics. In fact, we endeavour to find some leverage to back up psychology as a scientific discipline that can play more than just a subsidiary role to physics in the project of unification of sciences. We respect the autonomy of scientific disciplines that are mature enough to get related to other scientific disciplines, without paying homage to reductionism or domesticating advancements in psychology to the outcome of theories of physics. Be that as may, the foundation stone for the unificationist metaphysics has been solidly laid down by Ladyman and Ross, but the mansion is not completely erected yet. There is still much that can be achieved by developing a version of unificationism that gives special sciences their due. In harmony with previous attempts (Beni 2019) at relying on psychology as an appropriate base for a scientifically-informed metaphysics of sciences, in this paper, we show how the project of unification can be carried out without subverting psychology or making it subsidiary to physics.

Before going further, it should be remarked that there is a schism between the advocates of OSR, between those who strive to exterminate the relata that feature in

the structure completely (Muller 2011), and those who seek to argue that objects and relations have to be treated on the same ontological footing, with the objects being characterised only by the relations in which they stand. The arguments of moderate SR has been voiced by (Ladyman and Ross 2007; Ladyman 2007) among others. Here we focus on (Esfeld and Lam 2008, 2010; Esfeld 2013) which strive to apply the idea with an eye to the space–time notion of contemporary physics. They sought to substantiate their thesis by arguing that non-eliminativist SR (NSR), when applied to space–time, provides a convincing understanding of the nature of space–time points in the standard tensor formulation of general relativity. To flesh out a more basic argument, Esfeld and Lam brought the unification idea to the fore once more. They recommended NSR on the basis of the reconciliatory framework that it provides for spelling out the shared metaphysics of GR and QM.

We offer to provide further grounds for optimism about the metaphysical doctrine of SR. we do this by arguing that metaphysics of brain could be best developed in terms of the meta-theoretical framework of NSR. Our argument for this claim is quite straightforward. Based on empirical evidence about the brain's coding strategy, we argue that the brain's neural activity is based on statistically-based differences and results in formation of a statistically-based spatiotemporal structure. Such empirical characterisation of the brain's neural activity in terms of structure and relation is in accordance with the philosophical model of the brain that could be presented in terms of OSR. Our construal of difference-based coding and the emerging spatiotemporal structure allow for retaining a thin notion of relata, as the spatio-temporal points that are captured by the brain's neural processing. Before unfolding our account of the structural realist model of the brain, we provide a detailed review of Esfeld and Lam's formulation of NSR.

### 3 Moderate Structural Realism

Modern physics has been unkind to the notion of individual objects, and both Bose–Einstein and Fermi–Dirac formulations of statistics give way to the construal of quantum particles as ‘non-individuals’. Quantum particles are subject to the permutation invariance and thus indiscernible. This indicates there is no universal way for reconciling this non-individualistic picture with the individualist or monadic ontology that has prevailed upon the orthodox metaphysics for centuries (French 2014, pp. 34–35). One can seek to modify this pictures and put it in the framework of traditional metaphysics, at the cost of ignoring what scientific theories try to tell us, but that is not an option that a scientific-minded philosopher may want to grant. OSR solution is that orthodox object-oriented metaphysics has to be replaced with the metaphysics of SR (French and Ladyman 2003). Now, as I remarked, OSR could be developed into an eliminativist form that doesn't leave any room for the existence of the relata. This radical form of SR, however, is liable to some criticism. Aside from the conceptual issue concerning the notion of relata-less relations, the radical OSR seems to rely too absolutely on the new developments of sub-particle physics to take the philosophical implications of other areas of science into account. This is at odds with the spirit of unification. In reaction to this

situation, Esfeld and Lam (2008) sought to develop OSR into a more moderate venue that holds:

1. Relations require relata, that is, objects that stand in the relations.
2. *It is not the case* that these objects necessarily have intrinsic properties over and above the relations that they bear to one another. (Esfeld and Lam 2008)

The view retains a thin notion of object. This notion is deprived of any intrinsic properties. Objects are characterized contextually, and in accordance with their role in the structure. So, the relata and relations are to be treated on an equal ontological footing. The relation between objects and structure is one of mutual dependence, and the identity conditions of the objects are provided by the relations themselves. Aside from being true with regard to quantum systems, Esfeld and Lam argued that delegating the issue of the identity conditions to the relations is also true with regard to the space–time points, which can stand in exactly the same spatiotemporal relations and yet be numerically distinct (Esfeld and Lam 2008). This provides a nice alternative to understanding the objects as either bundles of intrinsic properties or bundle of relations. This approach does not mandate acceptance of haecceity or thisness as regards space–time points, and therefore, it is particularly consistent with the underlying assumptions of SR.

One significant virtue of Esfeld and Lam’s moderate SR is that it offers to extend the SR into some untrodden grounds. Although the advocates of OSR laid much emphasis on the compatibility of their view with modern physics, they were mostly focusing on QM, and have been more or less oblivious of the relation between their metaphysics and the developments in GM. Esfeld and Lam’s attempt at connecting OSR to GR in its standard tensor formulation and the fibre bundle formulation provides new grounds for appreciating the unifying virtues of the structuralist enterprise. Advocates of the radical SR seek to go around the problem by eliminating the space–time points on which the tensor fields could be defined. But this would lead to a failure in explaining how the space–time structure can be conceived without any reference to constituents. NSR is successful at fulfilling the explanatory job by treating the space–time points and the relations that prevail among them on an equal ontological footing, albeit without going so far to accept the space–time points are anything but mathematically individuated gloss that could be characterized contextually.

SR-theorists have been so enthused about the unifying virtues of the SR that in the very beginning of their defence of their metaphysical creed, they openly acknowledged that their naturalised metaphysics simply means “a metaphysics that is motivated exclusively by attempts to unify hypotheses and theories that are taken seriously by contemporary science” (Ladyman and Ross 2007, 1). It is in this spirit that NSR promises to provide a framework for integrating our philosophical interpretation of our theories of QM and GR. Although, the qualitative, physical nature of relations of the QM is essentially different from metrical relations that underwrite GR, both kinds of structures could be accommodated by the flexible metaphysical framework of moderate NSR (Lam and Esfeld 2012b; Esfeld 2013a). We offer to take the unifying program beyond the fields of QM and GR.

## 4 Going Beyond Physics

Esfeld and Lam's attempt at unifying different regions of physics their view has been still suffering from certain limitation. It is not a limitation of their theory if it did not outgrow its promised framework to actually extend the unifying scheme of NSR to the field of special sciences (such as biology, psychology, economics, etc.). What they did in the way of unifying the foundational physics was worthwhile enough. However, when they tried to flag ontological reductionism—indicating that any concrete relations in the natural world are identical with physical relations—they mildly underplayed the prospect of extending the approach to the field of psychology. That is to say, when remarking on the relation between cognitive system and physical objects, they went so far to claim that “these relations do not have any special ontological status” (Esfeld and Lam 2008, 29). Without intending to read too much into this short interjection, or proposing to evaluate Esfeld and Lam's work on the basis of this ephemeral hint, we argue that the naturalised metaphysics of SR could be somewhat open to ascribing ontological status of the domain of special sciences. It is true that OSR has been based on an unpretentious form of physicalism that holds:

Special science hypotheses that conflict with fundamental physics, or such consensus as there is in fundamental physics, should be rejected for that reason alone. Fundamental physical hypotheses are not symmetrically hostage to the conclusions of the special sciences. (Ladyman and Ross 2007, 44)

Ross and Ladyman's version of physicalism is open-minded in the sense that despite pursuing the goal of unification, they are not committed to a strictly reductionist agenda of, say, Oppenheim and Putnam (1958), who assumed that all phenomena (even phenomena studied by special sciences) could be explained by reference to theories of fundamental physics. Ladyman and Ross are not committed to a bottom-up form of micro-reductionism according to which all scientific theories, from psychology, sociology, biology, and chemistry reduce to fundamental physics. Their conception of unification (to some extent) acknowledges the autonomy of special sciences and does not give way to the hierarchical model of the world that puts microphysical at the bottom level (Ladyman and Ross 2007). In doing this, SR-theorists joined forces with the renowned advocates of anti-micro-physicalism (Hüttemann 2004; Schaffer 2003). We must immediately add that despite shunning the strict form of micro-reductionism, in view of its commitment to the principle of the primacy of physics, Ladyman and Ross' account of unification remains dedicated to physicalism. Let me elaborate.

In order to establish the principle of the primacy of physics, Ladyman and Ross (2007), 41 ff.) mainly rely on historical facts, which indicate that progress in the life sciences and psychology was possible only to the extent that these sciences were invoking same quantities and laws that were employed in physical theorizing. They argue that “in the history of science a succession of specific hypotheses to the effect that irreducibly non-physical entities and processes fix the chances of physical outcomes have failed” (Ladyman and Ross 2007, 43). As the statement shows, they

are advocating physicalism in the broadest sense, which merely underlines the consistency between physics and theories of special sciences. However, though they understand physicalism in the broadest sense, their view still implies that in case of any conflict between hypotheses in physics and hypotheses in psychology, it is not possible to question the hypotheses of physics. But it is not entirely clear why physics should always have it right, given the autonomy and maturity of special sciences. Nor is it clear why hypotheses regarding emergent properties in psychology would have to be consistent with fundamental physical principles.<sup>1</sup>

A vehement advocate of the autonomy of special sciences may argue that we cannot reduce psychology to fundamental physics because we cannot find the right sort of identity statements for mapping theories of the former discipline onto the theories of the latter (see Fodor 1974). It follows that because there is no precise mapping or translation between these two domains, we cannot ensure that there is consistency between physics and psychology. We appreciate this point and do not make our account of unification hostage to reductionism and physicalism, meaning that we do not take any a priori stance on the question of physicalism. This means that, although we suspect that there are inter-theoretical relations between physics and psychology, we do not go so far as to assume that hypotheses regarding emergent properties in psychology would have to be consistent with fundamental physical principles.

Most importantly, unlike Ladyman and Ross, we do not think that the relationship between physics and special sciences needs to be asymmetric, meaning that, say, in case of a conflict between physics and psychology, we do not presuppose that physics would always be right and psychology would always be wrong. We appreciate the broadness of Ladyman and Ross' conception of physicalism, but in speaking of brain structure, we make ontological commitments to statistical patterns of neuronal activity which may be characterized as informational structures, without specifying information as either a physical or non-physical commodity. Specifically, we propose that the world's statistical structure provides the basis for any subsequent structures—ontological relation is statistical structure. If so, these very same statistical structure should be manifest on different levels of the world. For instance, statistical structure may be manifest in the kind of structures described in physics, i.e., physical structure, and, in extension, in the brain too. Conceived within the framework of ontological SR, the brain may be characterized by statistical structures which, ultimately, can be traced to the world's statistical structure—statistical structure, formalized ontologically as relation in SR, do then provide the “common currency” (Northoff et al. 2019) of world and brain, and ultimately of science and world. Note that statistical structure is here meant in a most basic way, on a purely non-representational level of processing inputs and output stimuli; that must be distinguished from the cognitive more representational level where statistical structure can refer to contents as for instance in the statistically-based organization of contents in perception (as in gestalt theory). In the remainder of the paper, we will flesh out our views slowly.

---

<sup>1</sup> We owe this remark to one of the reviewers of our paper.



It is our claim that sciences that deal with the brain and mind are mature enough now to provide us with reliable knowledge about the relationship between physical objects and our cognitive systems. Also, we may rely on the information bestowed by our best scientific theories to philosophise about the structural constitution of the cognitive systems themselves. Because theories of neuroscience and the cognitive sciences that could be invoked to ground this form of structuralism are not in open contradiction to physics, their ontological implications could be embraced by SRists as the missing pieces of the metaphysical puzzle of unification. It is in this light that we argue that philosophical implications of best scientific theories as regards the spatiotemporal structure of the brain (and its relation to the physical world) could be surmised as in harmony with the philosophical picture of the world that is submitted by QM and GR. Importantly, the notion of spatiotemporal structure of the brain is not meant in a cognitive representational sense. Instead, it is understood in a dynamic sense, referring to the constitution of time and space by the brain in its relation to, i.e., alignment to the world (Northoff et al. 2019); such dynamic sense of spatiotemporal structure intrinsically connects the brain to the world and its spatiotemporal dynamic, i.e., world-brain relation (Northoff 2018). That very same spatiotemporal nature of world-brain relation, in turn, makes it possible for us yielding metaphysical assumptions and scientific investigation of the world.

The plausibility of the bigger picture—in which the theories of physics and special sciences are neatly interwoven to one another—is the incentive that persuades us to attempt to promote SR-metaphysics as providing a parsimonious model of the brain-world relationship.

#### **4.1 Being a Structural Realist About the Brain I: Difference-Based Coding as the Brain's Neural Code**

We begin to extend SR to the field of brain studies. But we must make a disclaimer first. Our intended sense of structuralism is not precisely the same as what sometimes is called structuralism about mental representations in the philosophy of mind, which aims to explain mental representations without appealing to non-physical or representational properties (see O'Brien and Opie 2004). Our conception of structuralism is borrowed from the philosophy of science, and before this occasion, it has been applied to the philosophy of mind on a few occasions (Beni 2016; Laughlin 1979; Northoff 2018). Here, we aim to apply SR to the brain and show how ontological commitments are to be made to brain structures. In this sense, our version of structuralism does not need to be in contrast with functionalism, or in fact with any of the other fashionable theories of the philosophy of mind. This is because coming from a background of the philosophy of science, our intended sense of structuralism does not map precisely with the extant theories of the philosophy of mind. That said, we acknowledge that to the extent that correspondence with the extant theories of philosophy of mind is at issue, our theory is compatible with functionalism (in the sense of Block 1980; Piccinini 2004).<sup>2</sup> Roughly stated, functionalism is a metaphysical program which holds that mental states can be

<sup>2</sup> We owe this remark to one of the reviewers of this journal.

individuated according to their place in the organization of inputs, outputs, and other mental states, rather than according to their content. This is in complete agreement without our understanding of scientific structuralism which identifies the objects in accordance with relations in which they stand and the relations that they bear to one another. However, while there are philosophers who developed functionalism with an eye to folk theories of mind (Lewis 1972) or within the context of metaphysical discussion of consciousness (Shoemaker 1975), we develop our version of scientific structuralism in a scientific context and with the explicit aim of showing how brain structures could be the subject of ontological commitments. After this clarification, we proceed to draw on some the recent view of characterizing the brain in terms of spatiotemporal dynamic (Northoff and Huang 2017; Northoff et al. 2019) to flesh out our insight. We explain how the brain's fundamental activity could be spelled out in structuralist terms. Afterwards, we show how this activity forms the spatiotemporal structure of the brain, and show how our proposal adds up to NSR. Let us begin by showing how we can model the dynamic and spatiotemporal (rather than perceptual or cognitive) the features of the brain's neural activity. If we can characterise the brain's neural activity in terms of spatiotemporal structure and relation rather than in terms of relata or intrinsic properties, then we could claim that there is room for extending SR to the field of neuroscience. We begin with the basic unit of the neuronal activity: neural code. There are numerous approaches to the question of the neural code (See Northoff 2014a for an overview). Here we offer to highlight only one particular approach that focuses on encoding.

The brain shows spontaneous activity and stimulus-induced task-evoked activity—the latter is triggered by stimuli or tasks while the former arises within the brain itself (see Northoff 2014a, b). The interaction between spontaneous activity and stimuli can occur because the two share something like a common code or “common currency” that underlies their differences (Northoff et al. 2019). Specifically, we understand by “common currency” shared features between spontaneous activity and stimuli independent of and prior to the latter's subsequent perception and cognition on the basis of the former. “Common currency” refers here to the processing level which, as one aspect thematized in signal detection theory, concerns the relationship between signal and noise: the processing of inputs is more specific targeting single inputs if the signal-to-noise ratio is high which, in contrast, decreases the sensitivity to a wide variety of different inputs. Generally, one may want to say that a balanced signal-to-noise ratio may be best suitable for establishing high degrees of sharing, i.e., high “common currency” of neural activity and input stream. One way to construct the needed bridge would be to code stimuli and spontaneous activity in direct relation to each other on the basis of their different statistical frequency distribution across time and space, i.e. in terms of spatiotemporal structure (we will speak more of this structure in the end of this section). Spontaneous activity shows continuous change which results in a certain statistical frequency distribution that could be described as “neuronal statistics” (Northoff 2014a, b). The stimuli themselves follow and occur in a certain statistical frequency distribution, i.e., their “natural statistics” (Barlow 2001) (See also the group around J. L. Gallant for investigating the encoding of natural statistics).

Rather than coding each stimulus by itself, Barlow suggested that the brain codes and represents “chunks of stimuli” and their details together. He calls the results of this process “gathered details” (Barlow 2001, p. 248). Take the example of a complex scene with a breakfast table covered with various items of food and plates etc. In that case our glance first falls on the big tea pot in the middle, then we wander to the bread basket, from there to the cheese plate, the jams, and the various plates. All items are located at different spatial positions on the table and are not perceived simultaneously by us—we rather perceive them sequentially by letting our glance wander around the table and its various items. If one were encoding each single stimulus by itself, one would not connect all items together and consider them to belong to one and the same table, the breakfast table. Nor could one link all stimuli together under a particular category, e.g., breakfast in our example; instead, different stimuli may be grouped together under different schema (as it is described in schema theory). The outline of a structuralist account of the brain’s activity looms large. The metaphor of the breakfast table shows that the important element in formation of the breakfast table is the relation between the constituents. The constituents themselves cannot be eliminated from the table, but they are not to be characterised with regard to their intrinsic properties,<sup>3</sup> so much as being identified with regard to their position in the general pattern of the table (coffee, whose intrinsic virtues are held in high esteem by us, and indeed all overworked philosophers, could be an exception). Metaphors are not precise in all respects, but all in all the breakfast metaphor sheds light on how neural system encodes the stimuli. Despite their spatial and temporal differences, the different stimuli (and hence the different items) must be encoded in conjunction. Once they are put together during encoding, they come to constitute what Barlow describes as “chunks of stimuli” and “gathered details”. These “chunks of stimuli” and their “gathered details”, in turn, may shape and thereby provide the basis of the contents of subsequent perception and cognition. As they are “put together” on the basis of the underlying spatiotemporal dynamic of the brain’s spontaneous activity, the contents of perception and cognition have a strong dynamic or spatiotemporal layer which, on a deeper level, complement their cognitive surface, i.e., the contents as we perceive and cognize them.

Yet another example is the phenomenology of experiencing music, i.e., of a melody. The phenomenologist Edmund Husserl already described early on in his book on “Inner time consciousness” (Husserl 1991) that we do not hear any single tone in isolation but perceive the present tone in relation to the previous one and often make predictions about the next forthcoming tone. This is only possible if we encode the present tone in relation to the previous one thus putting both together as “chunks of tones” with “gathered details” by means of which we constitute a temporal stream of tones which we then hear as melody. This is only possible, according to Barlow, if our brain encodes the occurrence of the tones (and stimuli in

---

<sup>3</sup> This simply means that we do not have access to real essence, hidden causes, or quiddities. All physical properties, such as the effects of implied momentum of a moving target could be specified structurally, that is in terms of say, the relationship between mass and velocity. We do not need to know the intrinsic essence of the moving target, but identify its momentum structurally on the basis of the physical structure in which the mass and velocity are connected in a certain way.

general) in terms of their statistical occurrence in time and space. The closer temporally the tone follows the preceding one, the more likely both tones are encoded and processed together as “chunks of tones”. The same holds obviously for the spatial dimension: in the case of the breakfast table, the various items are spatially near to one another and are therefore highly likely to be encoded together as “chunks of stimuli”.

We proceed with presenting further details, to indicate how our model of the brain’s activity conforms to SR. Let us start with what is *not* encoded into neural activity, since that will make it easier for us to better understand the brain’s actual encoding strategy. When perceiving a melody, for example, Barlow proposes that the sensory cortex does not encode each tone by itself. Instead of encoding single stimuli by themselves, the brain seems to encode the distribution of the stimulus as, being more specific, melodic contour and interval size rather than in terms of absolute (auditory) frequencies. In the case of a bird’s song, for example, the brain will encode the distribution of a particular tone across discrete points in physical time. And the brain may also encode the spatial position of the bird’s tone relative to, for instance, a nearby rustling of leaves. What is encoded into neural activity is thus the statistical frequency distribution of stimuli across different discrete points in physical time and space. This is what Barlow describes as the encoding of the stimuli’s “natural statistics,” the statistical frequency distribution of a stimulus across discrete positions in time and space.

Having described natural statistics, it is now imperative to clarify the nature of “neuronal statistics” which can be seen as manifestation of the world’s statistical structure within the brain (see above for details including the relationship to signal detection theory). Externally-generated events in the environment are encoded in terms of their statistical frequency distributions, or natural statistics, into the brain’s neural activity, the result of which is stimulus-induced activity. The same holds, analogously, for the brain’s spontaneous activity itself. Internally-generated events within the brain are encoded in terms of their statistical frequency distributions, or neuronal statistics, the result of which is spontaneous activity. This has major implications. The encoding of the external stimuli’s natural statistics into neural activity is only possible through interaction with the neuronal statistics that characterises spontaneous activity. The interaction between external stimulus and spontaneous activity can consequently be sketched as an interaction between two different statistics, natural and neuronal.

Now we set forth the kernel of our structuralist model of the brain. The brain and its spontaneous activity’s neuronal statistics encode stimuli as statistical frequency distributions across different points in time and space. The resulting neural activity, the stimulus-induced activity, then reflects the statistically-based differences between the spontaneous activity’s neuronal statistics and the stimuli’s natural statistics. This amounts to “difference-based coding” (see Northoff 2014a for experimental details). One instance of difference-based coding is predictive coding as developed by Friston (2010). Predictive coding refers to the fact that stimulus-related activity results from the comparison of an anticipated or predicted input, i.e., an empirical prior, in the pre-stimulus period with the actual stimulus input: the degree of their divergence or convergence is described as prediction error which

determine the amplitude or magnitude of stimulus-related activity. Since the prediction error is based on the difference between predicted and actual stimulus input, it can be viewed as one instance of difference-based coding. Difference-based coding extends beyond the relation of pre-post-stimulus intervals as described by predictive coding: difference-based coding is supposed to describe any neural activity changes that are assumed to be based on the relative differences between different time points (Northoff 2014a).

Thus, statistically-based differences provide the “common currency” between spontaneous activity’s neuronal statistics and stimuli’s natural statistics. This common currency, we contend, constitutes the relation that brains bear to the wider world in which they exist.

## **5 Being a Structural Realist About the Brain II: Space Time Relations Versus Spacetime Points**

Traditionally, stimulus-induced activity is supposed to be superseded and thus added onto the ongoing level of spontaneous activity (He 2013; Fox et al. 2006, 2007). Ontologically speaking, such additive rest-task interaction would mean that there is no real relation let alone interaction between the spontaneous activity and the external stimulus. However, recent findings shed doubt on that as they show non-additive interaction: stimulus-induced activity is not simply added onto the ongoing level of spontaneous activity but interacts with the latter such that the former is higher or lower than their mere addition (He 2013; Huang et al. 2017). Ontologically speaking, non-additive interaction means that there is relation of spontaneous and stimulus-induced activity with both being mutually dependent upon each other (Northoff et al. 2010).

We can now explain how difference-based coding makes the non-additive interaction between spontaneous and stimulus-induced activity possible. Non-additive interaction is possible only if the spontaneous activity can directly interact with the stimulus and impact the degree to which it elicits stimulus-induced activity in the brain. Different degrees of non-additive interaction are mediated by different degrees of statistical-difference-based matching between the spontaneous activity’s neuronal statistics and the stimuli’s natural statistics. This means that the better their respective statistics match in their statistically-based differences, the more strongly the spontaneous activity’s neuronal statistics can impact the stimulus and its natural statistics. This would lead to higher degrees of non-additive interaction. Now, brain spontaneous activity could be surmised in terms of difference-based coding. And statistical-difference-based coding paves the way for formation of a structuralist understanding of the brain’s activity. It is the statistically-based differences between spatiotemporal points, rather than the points themselves, that are encoded by the brain. The brain has the interesting tendency to form internal models of the environment by capturing the relations, e.g., statistically-based differences between different stimuli (or events/objects) as in, for instance mental imagery or propositional forms of representation (Kosslyn 2007). The relata themselves, i.e., the space–time points could not be eliminated from this picture entirely. Even so,

they have to be identified and determined with regard to the statistically-based relation that they bear to each other, rather than on the basis of their intrinsic properties. This paves the way for construing the brain's spontaneous activity in terms of NSR. Notice that stimulus-based coding does not refer solely to the external stimulus alone but also internal stimuli; the term stimulus refers here to one specific time point in neural activity which can be related to both internally-oriented stimuli, as during spontaneous activity, or externally-oriented stimuli as during what is referred to as stimulus-induced activity.

To establish our point, we offer a thought experiment. Imagine there were stimulus- rather than difference-based coding. In that case, the stimulus would only be encoded in an isolated way, in its discrete point in time and space, remaining untethered by any statistically-based relation to other stimuli or the brain's spontaneous activity. This would make any direct interaction (e.g., reciprocal modulation) between spontaneous activity and stimuli impossible: the stimuli would then only be processed in terms of their own discrete points in time and space (and their specific frequency) without any interaction with other discrete points in time and space (and other frequencies) as provided by their interaction with the brain's spontaneous activity. In other terms, the temporal and spatial features of the stimuli would no longer be modulated as being spatiotemporally extended in by the brain's spontaneous activity. In such cases, stimulus-induced activity would supervene on the ongoing spontaneous activity in a merely additive way: as it is limited to one specific discrete time point, stimulus-based coding cannot link different points in time like pre-stimulus activity and stimulus onset let alone allow for their direct interaction. That, in turn, renders impossible direct interaction between the different time points of prestimulus activity levels and the external stimulus; the effect of the latter can consequently be only added on top of the former entailing additive interaction.

In short, stimulus-based coding precludes non-additive interaction between spontaneous activity and stimuli. And this is experimentally implausible.<sup>4</sup> Therefore, the coding process that underlies non-additive interaction is difference-based coding. The difference-based coding results in more viable ways of encoding the information, when compared to the stimulus-based coding, whose bearing on the person's perception and cognition of external events would be questionable. This is because temporally separate stimuli could no longer be integrated and linked: One starts looking at the eyes in the face of a person and then continues to the nose and the mouth without being able to integrate the different ingredients of the face; this is well compatible with some of the laws described in

<sup>4</sup> One may want to sketch a slightly distinct scenario though. In that case there would be no interaction between the different stimuli but interaction between single stimulus and spontaneous activity. Would that still be an instance of difference-based coding? Yes, with regard to the stimulus interaction with spontaneous activity, no to the stimulus' interaction with other stimuli. Even if deemed logically conceivable, such a scenario remains problematic and ultimately even logically incoherent. It is the spontaneous activity itself that links and connects the different stimuli with each other by means of which it constitutes the statistically-based differences between the different stimuli. That remains impossible without the stimuli' interaction with the spontaneous activity though. In other terms, the two interactions, stimuli-stimuli and stimuli-spontaneous activity, are linked together. However, that remains impossible without the interaction of the stimuli with the spontaneous activity.

Gestalt theory which, due to the reasons of space, cannot be exploited here. Difference-based coding, on the other hand, allows for encoding the statistically-based temporal differences between eyes, nose, and mouth as incidences of natural statistics, which makes possible their integration and relation as is present when we perceive them as part of one face, albeit without going so far as to deny the significance of the incidences of natural statistics altogether. That being so, the experimental grounds indicate that it is best to understand the brain's neural activity in terms of NSR. Note that this relation is here determined only on the spatiotemporal level of neural activity leaving open how the latter shapes the psychological level and its specific principles and laws (like the gestalt laws).

## 6 Concluding Remark

In this paper, we drew on the attempts of Esfeld and Lam to use the unifying trajectory of structural realism in developing a metaphysical framework for integrating the philosophical interpretations of quantum mechanics and general relativity; this served the purpose of extending non-eliminative structural realism to the field of mind and brain studies. Relying on some recent breakthroughs in neuroscience, we argued that the basic neural mechanisms of the brain which are at work in formation of a neuro-dynamic conception of space–time could be construed in structuralist terms. It was in this spirit that we endeavoured to push the borders of structural realism further, and go beyond the domains of quantum mechanics and general relativity, so as to extend SR to the field of brain studies.

The enterprise for extending structural realism into the field of brain and mind studies is beneficial to the structural realist project too. Projecting structural realism into the field of mind and brain studies paves the way for weaving the argument patterns and spreading explanatory paradigm of structural realism more confidently. It is, “the *raison d'être* of a useful metaphysics”, after all, “to show how the separately developed and justified pieces of science (at a given time) can be fitted together to compose a unified world-view” (Ladyman and Ross 2007, 58). This proves that the promise of unifying different patterns of scientific activity across different fields of study is indeed feasible, and thus there are further grounds for being optimistic about the success of metaphysics of structural realism.

## References

- Barlow H (2001) The exploitation of regularities in the environment by the brain. *Behavioral and Brain Sciences* 24(4):602–607
- Beni MD (2016) Structural realist account of the self. *Synthese* 193(12):3727–3740. <https://doi.org/10.1007/s11229-016-1098-9>
- Beni MD (2018) A structuralist defence of the integrated information theory of consciousness. *J Conscious Stud* 25(9–10):75–98
- Beni MD (2019) *Cognitive structural realism: a radical solution to the problem of scientific representation*. Springer Nature, Cham
- Block N (ed) (1980) *Troubles with functionalism*. In: *Readings in the philosophy of psychology*. Harvard University Press, Cambridge, vols 1 & 2, pp 268–305

- Esfeld M (2013) Ontic structural realism and the interpretation of quantum mechanics. *Eur J Philos Sci* 3(1):19–32. <https://doi.org/10.1007/s13194-012-0054-x>
- Esfeld M, Lam V (2008) Moderate structural realism about space-time. *Synthese* 160(1):27–46. <https://doi.org/10.1007/s11229-006-9076-2>
- Esfeld M, Lam V (2010) Ontic structural realism as a metaphysics of objects. In: Bokulich A, Bokulich P (eds) *Scientific structuralism*. Springer, Dordrecht, pp. 143–59. [https://doi.org/10.1007/978-90-481-9597-8\\_8](https://doi.org/10.1007/978-90-481-9597-8_8)
- Fodor JA (1974) Special sciences (or: the disunity of science as a working hypothesis). *Synthese* 28(2):97–115. <https://doi.org/10.1007/BF00485230>
- Fox MD, Snyder AZ, Zacks JM, Raichle ME (2006) Coherent spontaneous activity accounts for trial-to-trial variability in human evoked brain responses. *Nat Neurosci*. <https://doi.org/10.1038/nn1616>
- Fox MD, Snyder AZ, Vincent JL, Raichle ME (2007) Intrinsic fluctuations within cortical systems account for intertrial variability in human behavior. *Neuron*. <https://doi.org/10.1016/j.neuron.2007.08.023>
- French S (2014) *The structure of the world*. Oxford University Press, Oxford. <https://doi.org/10.1093/acprof:oso/9780199684847.001.0001>
- French S, Ladyman J (2003) Remodelling structural realism: quantum physics and the metaphysics of structure. *Synthese* 136(1):31–56. <https://doi.org/10.1023/A:1024156116636>
- Friston KJ (2010) The Free-Energy Principle: a Unified Brain Theory? *Nat Rev Neurosci* 11(2):127–138. <https://doi.org/10.1038/nrn2787>
- Hasselmann F, Seevinck MP, Cox Ralf F A (2010) Caught in the undertow: there is structure beneath the ontic stream. *SSRN Electron J*. <https://doi.org/10.2139/ssrn.2553223>
- He BJ (2013) Spontaneous and task-evoked brain activity negatively interact. *J Neurosci*. <https://doi.org/10.1523/JNEUROSCI.2922-12.2013>
- Huang Z, Zhang J, Longtin A, Dumont G, Duncan NW, Pokorny J, Qin P et al (2017) Is there a nonadditive interaction between spontaneous and evoked activity? Phase-dependence and its relation to the temporal structure of scale-free brain activity. *Cerebral Cortex (New York, N.Y.: 1991)*. <https://doi.org/10.1093/cercor/bhv288>
- Husserl E (1991) On the phenomenology of the consciousness of internal time (1893–1917), trans. J. B. Brough, Kluwer, Dordrecht. <https://doi.org/10.1007/978-94-011-3718-8>
- Hüttemann A (2004) *What's wrong with microphysicalism?*. Routledge, London
- Kosslyn S (2007) *Clear and to the Point: 8 Psychological Principle for Compelling PowerPoint Presentations*. Oxford University Press, Oxford. <https://doi.org/10.1007/s13398-014-0173-7.2>
- Ladyman J (1998) What is structural realism? *Stud History Philos Sci A* 29(3):409–424. [https://doi.org/10.1016/S0039-3681\(98\)80129-5](https://doi.org/10.1016/S0039-3681(98)80129-5)
- Ladyman J (2007) On the identity and diversity of objects in a structure. In: *Proceedings of the aristotelian society, supplementary volumes*. Oxford University Press, The Aristotelian Society. <https://doi.org/10.2307/20619100>
- Ladyman J (2008) Structural realism and the relationship between the special sciences and physics. *Philos Sci* 75(5):744–755. <https://doi.org/10.1086/594520>
- Ladyman J, Ross D (2007) *Every thing must go*. Oxford University Press, Oxford. <https://doi.org/10.1093/acprof:oso/9780199276196.001.0001>
- Lam V, Esfeld M (2012) The structural metaphysics of quantum theory and general relativity. *J Gen Philos Sci* 43(2):243–258. <https://doi.org/10.1007/s10838-012-9197-x>
- Laughlin DM (1979) *Structuring the*. pp 53–56
- Lewis D (1972) Psychophysical and theoretical identifications. *Aust J Philos*. <https://doi.org/10.1080/00048407212341301>
- Lloyd EA (1994) *The structure and confirmation of evolutionary theory*. Princeton University Press, Princeton
- Lloyd EA (2015) Evolutionary theory, structure of. In: James W (ed) *International encyclopedia of the social & behavioral sciences*, vol 8. Elsevier, Amsterdam. <https://doi.org/10.1016/B978-0-08-097086-8.81007-7>
- Muller FA (2011) Withering away, weakly. *Synthese* 180(2):223–233. <https://doi.org/10.1007/s11229-009-9609-6>
- Northoff G (2014a) *Unlocking the brain: volume 1: coding*. Oxford University Press, New York
- Northoff G (2014b) *Unlocking the brain: volume 2: consciousness*. Oxford University Press, New York. <https://doi.org/10.1093/acprof:oso/9780199826995.001.0001>



- Northoff G (2018) *The spontaneous brain: from the mind-body to the world-brain problem*. MIT Press, Cambridge
- Northoff G, Huang Z (2017) How do the brain's time and space mediate consciousness and its different dimensions? Temporo-spatial theory of consciousness (TTC). *Neurosci Biobehav Rev* 80(September):630–645. <https://doi.org/10.1016/J.NEUBIOREV.2017.07.013>
- Northoff G, Qin P, Nakao T (2010) Rest-stimulus interaction in the brain: a review. *Trends Neurosci*. <https://doi.org/10.1016/j.tins.2010.02.006>
- Northoff G, Wainio-Theberge S, Evers K (2019) Is temporo-spatial dynamics the 'common currency' of brain and mind? In quest of 'spatiotemporal neuroscience'. *Phys Life Rev*. <https://doi.org/10.1016/j.plrev.2019.05.002>
- O'Brien G, Opie J (2004) Notes toward a structuralist theory of mental representation. In: Clapin H, Staines P, Slezak P (eds) *Representation in mind: new approaches to mental representation*. Pergamon, Oxford
- Oppenheim P, Putnam H (1958) Unity of science as a working hypothesis. In: Feigl H, Scriven M, Maxwell G (eds) *Minnesota studies in the philosophy of science, vol II*. University of Minnesota Press, Minneapolis, pp 3–36
- Piccinini G (2004) Functionalism, computationalism, and mental states. *Stud History Philos Sci A* 35(4):811–833. <https://doi.org/10.1016/j.shpsa.2004.02.003>
- Ross D (2008) Ontic Structural Realism and Economics. *Philosophy of Science* 75(5):732–743. [10.1086/594518](https://doi.org/10.1086/594518)
- Schaffer J (2003) Is there a fundamental level? *Nous* 37(3):498–517. <https://doi.org/10.1111/1468-0068.00448>
- Shoemaker S (1975) Functionalism and qualia. *Philos Stud* 27(5):291–315. <https://doi.org/10.1007/BF01225748>
- Worrall J (1989) Structural realism: the best of both worlds? *Dialectica* 43(1–2):99–124. <https://doi.org/10.1111/j.1746-8361.1989.tb00933.x>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Affiliations

Majid Davoody Beni<sup>1,2</sup> · Georg Northoff<sup>3,4,5,6,7,8,9</sup>

<sup>1</sup> Department of Management, Science, and Technology, Amirkabir University of Technology, Tehran, Iran

<sup>2</sup> Nazarbayev University, Nur-Sultan City, Kazakhstan

<sup>3</sup> Institute of Mental Health Research, University of Ottawa, Ottawa, Canada

<sup>4</sup> Mind, Brain Imaging and Neuroethics Research Unit, University of Ottawa Institute of Mental Health Research, Royal Ottawa Mental Health Centre, 1145 Carling Avenue, Rm. 6435, Ottawa, ON K1Z 7K4, Canada

<sup>5</sup> Taipei Medical University, Graduate Institute of Humanities in Medicine, Taipei, Taiwan

<sup>6</sup> Taipei Medical University-Shuang Ho Hospital, Brain and Consciousness Research Center, New Taipei City, Taiwan

<sup>7</sup> National Chengchi University, Research Center for Mind, Brain and Learning, Taipei, Taiwan

<sup>8</sup> National Chengchi University, Department of Psychology, Taipei, Taiwan

<sup>9</sup> Centre for Cognition and Brain Disorders (CBBDD), Normal University Hangzhou, Hangzhou, China