Euro Jnl Phil Sci (2017) 7:393–409
DOI 10.1007/s13194-016-0167-8

ORIGINAL PAPER IN PHILOSOPHY OF SCIENCE

Reconstructing the upward path to structural realism

Majid Davoody Beni[1]

Received: 26 June 2016 / Accepted: 22 November 2016 / Published online: 16 December 2016

Abstract In his Analysis of Matter, Russell advocated an epistemic form of Epistemic Structural Realism (ESR) which held that our knowledge of the external world is structural. This approach has been criticised by M.H.A. Newman (Mind, (146), 137–148, 1928) and Stathis Psillos (Philosophy of Science, 68(S3), S13–S24, 2001). The paper aims to reconstruct Russell's version of ESR, and defend its experimental and philosophical plausibility. The basic assumption is that without getting a viable experimental handle on the linkage between the structure of perceptions and causal structure of the stimuli, the so called upward path to SR couldn't be defended conclusively. In this vein, the paper explores the resources of the recent experimental theories of perception (these include the theory of brain's difference-based coding, and the theory of prediction error minimization). These help establishing a structuralist account of the perception, and provide an account for the causal linkage between the structure of perceptions and the structure of the stimuli. I argue that it is possible to settle for objectivity of this linkage according to the criteria of comprehensiveness, simplicity, and above all rationality, in a way that meets the standards of the scientific realists.

Keywords Structural realism · Nemwan's objection · Theory of perception · Sparse coding · Predictive coding · Inference to the best explanation · Scientific realism

In his *Analysis of Matter,* Russell had advocated an epistemic form of Structural Realism (hereafter, SR) which held that our *knowledge* of the external world is structural. Russell's view was based on the idea of the existence of commonality between the structure of percepts and the structure of stimuli. Russell's formulation of SR was notoriously criticised by (Newman 1928). More recently, Stathis Psillos (2001) launched a lethal attack on Russell's approach to SR. Psillos called it the upward

✉ Majid Davoody Beni
m.davoody@auta.c.ir

[1] Department of management, science and technology, Amirkabir University of Technology, No. 424, Hafez Street, 3313-15875, Tehran, Iran

path to SR, because, being an empiricist, Russell attempted to begin from the empiricist premises and reach a sustainable realist position.

The aim of the present paper is to reconstruct the upward path to SR and defend its experimental and philosophical plausibility. The basic assumption is that without getting a viable experimental handle on the linkage between the structure of perceptions and the causal structure of stimuli, the so-called upward path to SR couldn't be defended conclusively. To achieve its goal, the paper explores the resources of the computational neuroscience in the way of mending Russell's upward path to SR. These resources consist of theories of sparse coding and difference-based coding, as being hypothesised by Georg Northoff among others, and the theory of prediction error minimization (PEM for short).

My proposal for repairing the upward path to SR consists of two steps. First, I argue that taking difference-based coding as the brain's main strategy for encoding sensory inputs helps to establish a structuralist account of the perception. Next step is to explain how the PEM theory helps to account for the causal linkage between the structure of perceptions and the structure of the stimuli. Furthermore, I will argue that we are entitled to settle for the objectivity of the linkage between the structure of perception and the structure of the world according to the criteria of comprehensiveness, simplicity, and above all rationality, in a way that meets the standards of the scientific realists. The moral that I draw from the discussion is that the upward path to the SR is repairable.

1 Russell's upward path to structural realism

SR is a philosophical thesis about the structural foundations of reality. There is a schism in the SRists' camp, between the advocates of the Epistemic and Ontic forms of SR (respectively, ESR and OSR). According to ESR, what we can *know* about the world is its structure. OSRists, on the other hand, strive to argue that structure is all that there is. We don't need to engage in this debate for the time being. Suffice it to say that, in retrospect, Russell's version of the SR can be recognised as a form of ESR. Before going any further, I have to draw attention to another distinction. Let us call this the distinction between the theoretical and the empirical approaches to SR.[1]

The theoretical arguments for SR are quite well-known. One form of it has been offered as a reaction to a famous antirealist argument. Pessimistic induction builds upon the instances of falsification of the successful scientific theories of the past and refers to the historical discontinuity of the theories to conclude that the success of our present scientific theories could not be an indicator of their truth (Laudan 1981). Structural realist parries the attack by remarking that the continuity could be found at the level of form, not content, of the changing theories (Worrall 1989). Another argument, particularly favoured by the advocates of OSR, relies on the ontology of Quantum Mechanics and General Relativity. Modern physics has not been kind to the notion of individual object, and being inspired by the breakthroughs in QM and GR, the advocates of OSR argued that it is recommendable to forsake the stale object-oriented ontology in favour of a structuralist one (French and Ladyman 2003; Lam and Esfeld 2012). Russellian (or

[1] This point has been brought to my attention by one of the reviewers of the paper.

empirical) SR, on the other hand, relies on a different kind of argumentation. The arguments for this kind of SR have to be experimentally informed. They aim at establishing an objective link—in terms of a structural relationship—between knowledge and its source, on the basis of experimental theories of perception. On the basis of such theories, the advocates of empirical SR tend to argue that although our perceptions do not reproduce the nature of stimuli, they do still represent the *structure* of the stimuli in a reliable manner. Let me elaborate.

SRists of various stripes agree that there is a structural similarity between the theories on the one hand and the reality that underlies the theoretical representations on the other hand. Russell's view is generally in accord with this consensus: "if the hypotheses as stated were correct, the objective counterparts would form the world having the same structure as the phenomenal world, and allowing us to infer from phenomena the truth of all propositions that can be stated in abstract terms and are known to be true of phenomena" (Russell 1919, p. 61). Therefore, if phenomenal experiences present a three-dimensional world, the world that gives rise to such experiences should have three dimensions, and so forth. In his *Analysis of Matter*, Russell elaborated on the same point to argue that what could be *inferred* from the structure of the percepts is the logico-mathematical structure of the world. The first-order (non-structural) properties of the external world, on the other hand, have to remain inherently unknowable to us (Russell 1927, p. 226). Hence Russell's delicate empirically informed the approach to ESR.

This empirical approach manifests through taking the epistemic legitimacy of the percepts (i.e. the building blocks of perceptions) for granted and trying to infer the structure of reality from the structure of the perceptions. This approach is different to the classical theoretical approach to SR in an interesting way. The phenomenal experiences find a clear logico-mathematical expression in our scientific theories, but this doesn't mean that content of the experience could be completely carried into sciences theories. This is mainly because the intrinsic character of stimuli remains unknown in this picture, but we don't need to dismay; nothing in physics depends on the intrinsic properties of the stimuli, after all (Russell 1927). The relations that are regimented in theories of physics are not identical with those that we perceive in the visual field but merely correspond to them in such a way that preserves their logico-mathematical properties. The phenomenal properties, in their turn, share the structure of the reality. The philosophical assumption that underlies this insight is that same causes imply same effects (if two stimuli are identical, then so are the resulting percepts). Equally, one can say that difference in effects (here, the percepts) is originated by the difference in causes (i.e. stimuli). Psillos calls this "Helmholtz-Weyl" principle (Psillos 2001, p. S14) because the principle (I call it HWP for short) has been endorsed by Helmholtz and Weyl on different occasions. HWP indicates that the relation between percepts represents the relation between the stimuli. The point is very important because it was upon such basis that Russell claimed we have a great deal of knowledge "as to the structure of stimuli" (Russell 1927, p. 227).

Before reviewing Psillos' objection I have to add a few lines about the viability of HWP. HWP may appear to be a version of what is often called 'Curie's Principle' (CP for short). In its original formulation by of Pierre Curie, CP indicated that when certain causes produce certain effects, the elements of symmetry of the causes must be found in the effects produced by those causes (Curie 1894). Since the converse of CP is not true, the effects should be more symmetric than causes. CP also provides a causal explanation for the cases of symmetry breaking in sub-particle physics. When applied to deterministic laws, CP holds for all transformation under which these laws are invariant. There are different physical processes that constitute a refutation of CP; the counterexample could be found in the fields of Lagrangian and Hamiltonian mechanics, QM, QFT, probabilistically stochastic theories, or even deterministic systems. CP has been also discredited as a vacuous, analytic principle, or a truism, which could be only occasionally applied to cases (see Earman 2004; Ismael 1997; Norton 2014, unpublished data; Roberts 2013). One way to face the counterexamples is to chalk them up to a deficiency in the postulation of the cause and its effect. But that attitude would be a hindrance to the discovery of new theories (Chalmers 1970, p. 148). So, the bearing of the counterexamples on the status of CP has to be taken seriously. However, I do not think that the existence of a similarity between CP and HWP is strong enough to soil the status of HWP or block the upward path to SR for good. I shall explain why.

Even without going into the detail, one may observe that counterexamples to CP are all extracted out of fields of classical mechanics or sub-particle physics. Considering the fact that CP itself had been presented as a principle of physics, the choice of the domain of counterexamples is just appropriate. However, despite mildly leaning towards physicalism, SR-theorists generally do not defend micro-physicalism or the generalised causal exclusion argument, which holds that only fundamental physics nestles genuine causation (Ladyman 2008; Ross et al. 2007 chapters 5, 7). Neither did Russell's philosophy harbour any straightforward commitment to the causal completeness of physics. Quite to the contrary, Russell (1917) was one of the first philosophers who contended that the attempt to domesticate modern physics to the orthodox causal metaphysics roots in "anthropomorphic superstitions" formed by habituated memories that rely on the asymmetry of past and future (Russell 1917). To make a long story short, although there are good reasons for being sceptical of the validity of CP in particular, or causal metaphysics in general, there is no reason to presume that the same counterexamples could rule out the viability of HWP or its substitute. This means, being a principle of the autonomous field of psychology, HWP could retain its status regardless of the demise of CP. Of course, this is not enough reason to establish the validity of HWP. However, once we complied that CP and HWP do not stand or fall together, we can proceed to explore further experimental grounds for the existence of a representational relation between the structure of perceptions and the range of stimuli, Psillos' well-posed scepticism notwithstanding.

Psillos' criticism began with the following remark: saying that we could infer the structure of the stimuli from the structure of the percepts, indicates that we have to be able to set structural isomorphism between the set of percepts and the set of stimuli. Psillos contended that HWP is not strong enough to establish the necessary isomorphism (because at least, the converse of HWP should be true too). It should be added that seeing isomorphism as an unnecessarily strong relation, Russell had complied with a 'roughly one-one relation' in his definition of a representational relation. A roughly one-one relation between stimuli and percepts was good enough according to Russell, for the purpose of inferring the mathematical properties of the stimuli from the knowledge of percepts (Russell 1927, pp. 226–227). Psillos, however, did not find the proposal convincing. It doesn't even make good sense to talk of roughly one-one relation according to Psillos because the relation is either one-one or it is not so (Psillos 2001, p. S15). Moreover, from a realist viewpoint, it should at least in principle be possible to assume that the reality possesses some "extra structure" that can't be manifested in the structure of the phenomena.[2] This indicates that the relation between phenomena and the reality could be modelled in terms of embeddability, instead of isomorphism. The problem is that, as Psillos remarked, if Russell had taken that step, then all the original attraction of his strategy for reconciling empiricism to realism had been lost (Psillos 2001, p. S15). One way or another, Russell's pioneering work on SR collapses into a form of structural empiricism:

> So, the Russellian upward path to SR faces an important dilemma. Without the converse of the Helmholtz-Weyl principle, it cannot establish the required isomorphism between the structure of the phenomena and the structure of the (unobservable) world. Hence it cannot establish the possibility of inferential knowledge of the latter. With the converse of the Helmholtz-Weyl principle, it guarantees knowledge of the structure of the world, but at the price of conceding-a prior-too much to idealism. (Psillos 2001, p. S16)

Votsis (2005) ventured to reply Psillos' challenge. Although Votsis' replies are fairly valid, they are not quite enough to establish the plausibility of Russell's SR. Votsis convincingly argued that although the unobservable world could have some extra structure that is not manifested in the structure of the phenomena, yet this point does not endanger the validity of ESR which simply holds that structures of phenomena mirror the structures of the unobservable world (p.1367). This has been beautifully associated with technical explanations about the role of morphism and injections. Votsis concluded that the existence of a variance between the structure of the external world and the structure of perceptions is enough for establishing the viability of ESR's claim. So, what is established by Votsis' argument is that *if we accept HWP* then some inferential knowledge about the structure of the external world can be safeguarded. However, as Votsis himself granted, there may be reasons for being sceptical of the plausibility of experimental resources of Russell's realism. As we saw, the plausibility of HWP is not beyond a shadow of reasonable doubt—considering its resemblance to CP. Therefore, some further (experimental) ground should be provided if we intend to establish the experimental plausibility of the foundations of Russellian SR, beyond a shadow of reasonable doubt. That is to say, a fully convincing argument for the plausibility of Russell's SR needs to go beyond refutation of Psillos' formal worries. Considering the empirical foundations of Russell's path to SR, we need to demonstrate that the experimental resources of the upward path to SR are really renewable.

2 One or two subtle methodological points

Even a scientific realist (and certainly an advocate of ESR) could comply with the existence of *some amount* of discrepancy between the scientific models and facts that

[2] In this respect, Psillos' objection is old wine in new bottle. Newman had criticised Russell's SR on the same grounds. According to Newman "the doctrine that *only* structure is known involves the doctrine that *nothing* can be known that is not logically deducible from the mere fact of existence, except ("theoretically") the number of constituting objects" (Newman 1928, p. 144 original emphasis).

they represent. Models don't correspond to their target systems in all respects and with complete faithfulness, because some aspects would be ignored and others would be underlined in the processes of idealisation and representation (see Teller 2011; Weisberg 2013). So, although SRists should be able to show there is no momentous divergence between the structure of phenomena and the structure of the world (as Psillos demanded), she does not need to be able to prove that the respective structures conform to one another completely. To strike a fine balance between these two extremes, the advocates of SR have to propose a viable way for constraining the divergence, without going so far to say there is a complete isomorphism between the structures. This reply is consistent with SRists' idea of partial isomorphism (Bueno et al. 2002; da Costa and French 2003), and Votsis' above-mentioned reply to Psillos. But, as I have already remarked, to meet the challenge in the spirit of Russellian empirical approach to the SR, we have to show that the claim about the convergence (if not complete correspondence) between the structures could be supported by the experimental theories of the field.

Some of Russell's contemporaries or predecessors (such as Poincare, Eddington, Cassirer), assumed that there is a structural relation between the theories of the physics on the one hand, and the psychological and cognitive processes on the other hand (see French 2014 chapter 4). It was in this venue that Russell (1919, 1927) alluded to the commonality between the structure of perceptions on the one hand, and the logico-mathematical structure of scientific theories on the other hand. That is to say, although it is true that the aim of Russell's (1927) enterprise was to investigate the philosophical outcome of modern physics, yet his philosophical system was resting on the foundation of the theory of perception, whose philosophical significance lies in the fact that it explains how we happen to know about the unperceived parts of nature. Note that Newman's (1928) objection—which gave rise to a constant challenge to all varieties of SR—had been originally targeting Russell's *causal theory of perception*. Psillos (2001) toyed with the idea of mending the psychological foundation of Russell's SR by marking the significance of Helmholtz's psychological theory, but not being a fan of ESR, he dismissed the possibility hastily. Be that as it may, the foundations of Russellian SR rests on the reliability of scientific theories (e.g. HWP or its substitutes) that account for the relation between the structure of perceptions and the structure of the stimuli in a way that constrains the divergence between two structures. I advocate this methodological assumption and argue that getting a scientifically informed handle on the structural relation between stimuli and perception provides all that is needed for repairing the upward path to SR. The first step is to propose a viable structuralist theory of perception. Then we should inquire as to how this structure latches onto the structure of the stimuli.

3 The structuralist account of perception

Before showing how the structure of perceptions latches onto the structure of reality, I have to explain why the system of perceptions could be best explained when we model it along structuralist lines. I argue that the experimental researches recommend sparse coding as the basic strategy that underlies the cognitive mechanisms of perception vis-à-vis the representation of the stimuli. Moreover, I suggest that as sparse coding is arguably based on the difference-based coding, then it is possible to account for it in terms of a structuralist approach. Afterwards, I will draw attention to the fact that this structuralist account of perception does not mandate a one-one relation between the perceptions and stimuli.

Sparse coding is allegedly the natural neural coding that the brain applies in order to encode the information, that is, to capture the features of the stimulus. Considering the brain as a neural network with on/off neurones, it is possible to demonstrate that only a small or 'sparse' number of neurones in the network are activated at any given time. I marginally distance from the standard definition and follow in Georg Northoff's footsteps, who claimed that at any given time, the brain's sensory input processing takes place "by a number of active neurons lower than the number of stimuli" (Northoff 2014a, p. 4). The history of the subject is at least as old as Barlow's interesting work (Barlow 1972), but the literature seems to be gaining momentum recently (Brenner et al. 2000; Field 1987; Hassoun 1993; Olshausen and Field 1997, 2004). Northoff proposed that sparse coding is the original metric or measure that brain applies to the processing of its own neural activity (Northoff 2014a, 2014b). It applies to any kind of neural activity in the brain, including both resting-state activity and stimulus-induced activity. It provides a "common currency" for linking the different levels of brain's activity. Here, the basic insight is that the activity of brain across cellular, population and regional levels. The experimental evidence that supports the theory is plentiful (Olshausen and Field 2004; Poo and Isaacson 2009; Simoncelli and Olshausen 2001; Zylberberg et al. 2011). I proceed to skim over some of the reasoning behind the theory.

The experimental data suggest that "dense coding" strategy, i.e., a one-to-many relation between the stimulus and neurones is highly inefficient. This is mainly because the amount of the neural activity that is invoked to code certain amount of information is not economical. "Local coding" or one-to-one coding strategy is not plausible either because it requires that the neurones be tuned to give distinct responses to extremely specific sensory inputs (Northoff 2014a, pp. 5–6; Vinje and Gallant 2000). This leaves us with the sparse coding which allows for encoding of sensory inputs' structure via the activity of an economical number of neurones. This strategy makes room for the functional establishment of multi-layered and time-conditioned neuronal networks (Molotchnikoff and Rouat 2012). The emphasis on the "sparseness" indicates that the sensory inputs are processed and coded in a sparse way, and as I remarked, the number of active neurones that are engaged in the processing is lower than the number of stimuli. This strategy allows for maximal information transfer at the cost of minimum involvement of active neurones. Seeing the brain's neuronal activity in terms of sparse-coding indicates that relation between the domains of stimuli and perceptions is not one-one, after all. Psillos' objection notwithstanding, turns out that "a roughly" one-one relation is not totally insensible. This provides a valuable purchase for defending the viability of Russell's approach to SR. But we still need to show how sparse coding theory could be incorporated into structuralist account of perception. To flesh out this point, it is enough to lay emphasis on the fact that sparse coding is underwritten by difference-based coding.

In what follows, I proceed to show how relations or "differences" between multiple stimuli are encoded by a comparatively sparse number of neurones, into the patterns of neuronal activity of the brain. I argue that since the processing of the assembly of neurones (whose number is sparser than the number of stimuli) encodes the relations and differences between the stimuli, then it could be argued that sparse coding "presupposes" difference-based coding. The same data that recommend sparse coding also indicate that brain doesn't encode different stimuli separately. What the brain captures is the structure of stimuli, that is to say, the brain encodes frequency distribution of the stimulus across its different discrete points in physical space-time (Northoff 2014a, p. 5). The statistical frequency distribution of a stimulus across different discrete positions in time and space is called the brain's "natural statistics" (ibid), and it indicates that neurones' responses would be adjusted to the stimuli according to the statistical variance in the occurrence of the stimuli. So, the brain's neural activity is based on encoding the differences between the stimuli rather than being based on encoding stimuli themselves. Because it is the difference between spatiotemporal points that is encoded by a relatively smaller number of neurones, it could be argued that sparse coding presupposes difference-based coding. In the light of this hypothesis, Northoff claimed that "sparse coding is supposed to be based on the encoding of spatial and temporal difference values as extracted from the stimuli's statistical frequency distribution across different discrete points in physical time and space" (Northoff 2014a, p. 13). The brain is predisposed, according to this approach, to encode the spatial and temporal differences, rather than the stimulus itself. This adds up to the conclusion that "spatial and temporal differences between different stimuli rather than the stimuli themselves are the common measure or metric in the brain's encoding of neural activity" (Northoff 2014a, pp. xviii–xxi).

Please note that difference-based coding is the brain's main encoding strategy across cellular, population, and regional levels. The indications of this point for structuralist construal is rather important. Were the difference-based coding only defined at the level of the processing of a single neurone—e.g., were we saying that difference-based coding occurs when relations or "differences" between multiple stimuli can be encoded by a single activated neurone—it would be almost impossible to extract a structuralist picture out of the difference-based coding theory. I explain this point via a thought experiment.[3] Suppose that by turning on a lamp, I express the distance between Toronto and Montreal in terms of structural information. Suppose that I have one such lamp for each pair of Canadian cities. Then it does not follow that the structure of my lamp collection matches the structure of Canadian cities. Indeed, it appears that the lamps could be moved around to an arbitrary degree and still capture the same structure. If this were the case, the way to extracting a structuralist account out of the difference-based coding would be almost blocked for good. There are good reasons, however, for optimism. The explanation follows immediately.

The encoding of stimuli is not a result of processing of a single neurone. There are assemblies of neurones that are activated during the processing. Each neurone engaged in the processing calibrates its activity with respect to its respective neighbouring neurone, hence it is possible to understand the neurone's activity in relationship to the activity of other neurones. The point has been established in a number of different experimental studies. One significant case that suggests itself is Grammont and Riehle's study of the dynamics of precise spike synchronisation and rate modulation in a population of neurones recorded in monkey motor cortex during performance of a delayed multidirectional pointing task (Grammont and Riehle 2003). Northoff relied on

[3] The example has been brought to my attention by one of the reviewers of this paper.

this, and similar studies, to argue that a neurone's activity (i.e. its difference-based coding) relies on the information from the respective others via encoding its own activity in relation to them. The encoding of the stimuli's relational properties is not taking place through the activity of a single neurone, but rather through the processing of a neuronal assembly. The difference-based processing of the population of neurones is tightly interdependent, because the coding activity of a single neurone in the assembly depends on its relationship to the activity of its neighbouring neurones, and vice versa. The activity of the single neurone could be understood in relation to the activity of the whole population. This indicates that neuronal processing of the brain is relational, and by the same token, could be regimented in a structuralist framework. This structuralist picture is based on the relational properties of the assembly of the neurones rather than the intrinsic properties of a single neurone. On the other hand, the differences between the space-time points, rather than the points themselves, are encoded into the neuronal patterns of the brain's coding. Although the relation between the pattern of neuronal processing and the structure of space-time is not one-to-one, it is not true that any given pattern of the neuronal processing (with any arbitrary arrangement) could encode certain spatiotemporal structure. This indicates that there is a structural similarity between the structure of stimuli and patterns of the brain's neuronal processing.[4]

Let me recap. On the one hand, the patterns of neuronal activity of the brain could be described in relational, difference-based, and structural terms. The spatiotemporal structures that the brain encodes could be surmised in equally relational, structural terms. This paves the way to a structuralist interpretation of perception, and it meshes nicely with the implications of Russell's view on the structure of perceptions.

Before ending this section, I re-emphasise the point that the brain is not merely reacting to the external stimuli. The brain is active when it scans the spatial and temporal differences and encodes them into the cellular and regional levels (Northoff 2014a parts III-IV). This indicates that the brain possesses some kind of intrinsic spatial-temporal structure against which it measures and compares the differences between the external stimuli. This is why the brain is predisposed to use difference-

[4] To cement the argument, we may take the argument beyond the activity of neuronal patterns in the brain of one subject. There are a number of interesting studies that confirm the existence of structural resemblance between the structure of stimuli and different patterns of neuronal processing within the brains of different subjects (Lindenberger et al. 2009; Sänger et al. 2012). These studies confirm the existence of structural similarity between different groups of neuronal codes (in different persons) that support the representation of a joint action. To demonstrate this claim, Sänger et al.'s (2012) focused on the analysis of the simultaneously recorded EEG from the brain of 12 guitar duets who repeatedly played a modified Rondo in two voices. The study indicates that phase locking as well as within-brain and between-brain phase-coherence connection strengths were enhanced in frontal and central sites of the brains of the subjects during the periods of high demand for musical coordination. This experimental setting easily demonstrates that the similarity of patterns of neuronal activity within the brain of the subjects is proportional to the structure of the stimuli to which they are exposed. Note that these are not the contents of stimuli that cause the structural similarity of the neural patterns, so much as the differences between the stimuli i.e., the differences between notes and tunes. The changes in the functional interbrain networks that emerges within and between the brains of duet of partners in delta (1–4 Hz) and theta (4–8 Hz) frequency ranges are proportional to the changes in the musical structure (changes in metre and tempo) of the segments of the adapted version of the Rando in D-Major. Interestingly, this relation, as well as the relation between the brain of the leader and the follower guitarists in the duet, could be modelled by small world graph, which indicates that the neuronal networks that represent the structure of stimuli finds a well-defined form.

based coding. Studies concerning the brain's "resting state" activity and its "default-mode-network" could demonstrate this point with enough experimental plausibility (Northoff 2012, 2014a chapters 4–5). The brain's intrinsic activity forges an intrinsic spatial and temporal structure, which lies at the foundation of its encoding activity and constitutes a baseline layer (Northoff 2014a, pp. 234–244). The underlying structure is the most important element of a thoroughly structuralist account of perception. What remains to be done, in the way of repairing the upward path to a realist form of structuralism, is to connect the structure of perceptions to the reality.

4 Interacting with reality

As I remarked, the brain is not the passive site of reacting to sensory inputs. To be more precise, as Northoff remarked "the data and their interpretation suggest that the brain's spontaneous activity is both passively affected, or shaped, by environmental events and, at the same time, actively modifies, e.g., amplifies or attenuates that very same affection" (Northoff 2016a, p. 18). The facts about active modification of the experience by amplification or attenuation constitute the main evidence for the existence of a bilateral dynamical relation between the world and the brain. The existence of the bilateral dynamical relation provides enough reason to defend Russell's upward path against the charge of collapsing into structural empiricism. Below I shall spell out the experimental details.

The relevant experimental evidence indicates that cerebral processes involved in the encoding of early life events in the spatiotemporal structure of brain's spontaneous activity could be arrayed into three layers (Duncan et al. 2015; Sadaghiani and Kleinschmidt 2013). First, there is the "surface" or "stimulus-induced" layer, which embeds the task-evoked activity in response to aversive stimuli. Then there is the "hidden layer" of brain's neural activity on which one can observe inter-individual differences in glutamate levels and entropy in the spontaneous activity. Finally, there is the "basis or bottom layer" wherein the spontaneous activity of brain's neural network latches onto the early childhood events. From the assessment of this layered picture, Northoff concluded that the difference between the "life events as real" and "life events as perceived" could be traced back to the difference between the environment as real on the one hand, and the environment as perceived on the other. The phenomenon of "active modification by amplification or attenuation" manifests when brain's spontaneous activity encodes environmental events. It is only on the basis of the assumption of the brain's direct interaction with the environmental context that this phenomenon, i.e. the modification of early life events through brain's intrinsic encoding activity, could be explained satisfactorily. Note that because "[t]he brain only shows a limited spatiotemporal scale or range when compared to the one of the world" (Northoff 2016b, p. 17), we can conclude that the brain and the spatiotemporal structure of its spontaneous activity do not mirror the world and its spatiotemporal structure in the one-one fashion. Let's consider Psillos' objection in this light.

Psillos' objection, we remember, held that without establishing the requisite isomorphism between the structure of the phenomena and the structure of the (unobservable) world, the upward path to SR cannot establish the inferential route to the knowledge of the latter. However, the above-mentioned experimental data clearly suggest that we can infer the knowledge of reality from the knowledge of the structure of the perceptions even in the absence of the one-one relation between. This means it is possible to avoid the first horn of Psillos' dilemma. However, as our access to the structure of the stimuli is mediated by and limited within the borders of our own perceptual field, the scepticism concerning resemblance of the structure of the unperceived stimuli to the structure of perception is not completely unrooted. In the next sections, I offer further experimental and philosophical reasons that help to assuage this lingering scepticism.

5 Predictive coding theory

In this section, I am using the resources of the prediction error minimization (PEM) theory in order to allay scepticism and show how the structure of perceptions latches onto the structure of realty.

Let us begin with emphasising the fact that, as Andy Clark (2013) remarked, PEM theory is a legacy of Helmholtz's contribution to physiological optics. Helmholtz's approach had presumed that sensory systems are in the business of inferring sensory causes from their bodily effects. It is possible, accordingly, to model perceptions in terms of a process of probabilistic, knowledge-driven inference. Russellian structuralism has been supposed to angle the upward path when boosted by a principle of Helmholtz's psychology. In recent years, the Helmholtzian psychology was reinforced by the resources of computational and neuroscientific advancement and gained considerable experimental and theoretical strength. PEM theory has been born in this context.

PEM model generally presumes that the brain has access to both its own internal estimation and the "true states of affairs" in the external world. Accordingly, PEM-theorists endeavour to explain how the brain could mark a difference between these two pictures and compare them in order to predict its error and minimise it in the way of approximating the reality. The prediction error is a feedback signal that the brain receives from the world in response to its effort for testing the hypothesis against the world so that the cognitive system could refine its hypotheses by taking the prediction error into account. The conclusion is that the PEM system is supervised and impacted by the truth (Hohwy 2014, p. 266).

Although there are several ways to model the PEM (Clark 2008; Hohwy 2014; Huang and Rao 2011; Kilner et al. 2007; Rao and Ballard 1999), I stay with Northoff in unfolding the technical aspects of the theory, mainly because his account of the PEM is deliberately consistent with the difference-based encoding strategy (as well as with the implied structuralism). Afterwards, I will develop my argument to conclude that it is possible to make connection between the structure of the perception and the structure of the world in a way that is demanded by the critics of the upward path to SR.

Northoff spelt out his version of the PEM theory (called "predictive coding") in terms of the relationship between sparse coding strategy and the encoding of predictions of extrinsic stimuli into the brain's neural activity (Northoff 2014a part III). This approach presumes that the discrepancy between the predicted stimulus and the actual stimulus determines the degree of stimulus-induced activity. (Northoff 2014a chapter 7). PEM theory is supported by various experimental data (e.g. Blakemore et al. 1999, 2000; Rao and Ballard 1999). Each of these experimental studies is quite telling in its own right. Moreover, as I remarked, Northoff's account of PEM is quite structuralist-friendly. The degree of the prediction error is in part dependent on the level of resting-state activity that immediately precedes the arrival of the actual input. The brain's capacity of predictive coding could be traced back to "rest-stimulus interaction" (Northoff 2014a, p. 152). This meshes quite nicely with the difference-based coding and the structuralist account that has been described in the previous section. The linkage between the perceptions and reality has been described in terms of the difference between the resting-state activity level and the stimulus-induced activity. According to this approach, the active brain (whose relational, neuronal structure has been assessed in the previous sections), is in charge of generating predictions and receiving continuous exteroceptive inputs from the environment. That being so, it could be claimed that different perceptions are informative about the external world, to the extent that they are representing the causal structure of the stimuli. This is mainly because the brain has been endowed with the wonderful capacity to minimise the discrepancy between the predicted inputs and the actual inputs. So, the brain's hypotheses about the structure of the world are capable of approximating the external reality.

It is still possible to claim that scepticism concerning the structural resemblance of perceptions to the stimuli cannot be eradicated conclusively by appealing to merely experimental grounds. Although I am inclined to believe that plausibility of the Russellian approach to SR could be based on the experimental evidence from the field of cognitive neuroscience, I will proceed to disclose further philosophical reasons to reinvigorate the realist component of Russellian SR.

6 Scepticism, IBE, and rationality

Although PEM theory has been associated with a realist construal as soon as it was introduced to philosophy (Clark 2012), it is still possible to argue that PEM theory cannot go beyond the internalist model of the brain completely (Hohwy 2014). In this section, I argue that even this internalist-inferentialist model of PEM could be used to account for the linkage between the structure of the perceptions and the structure of stimuli, on the basis of agreeable philosophical reasons SR.

The PEM theory could be understood, according to Hohwy, in terms of an inference to the best explanation (IBE) that could be expressed along the lines of a Bayesian model (Hohwy 2013, 2014). The winning hypothesis about the world is the one with the highest posterior probability, and this hypothesis provides the best explanation of the sensory input, in a context-dependent fashion (Hohwy 2014, p. 263). Upon such basis, we may argue that the brain provides a reliable representation of the world. To substantiate this claim, Hohwy draws on Hempel's account of self-evidencing explanation (Hempel 1965), according to which the information or assumption of occurrence of an event forms an indispensable part of the only available evidential support that the hypothesis may receive. If we see PEM theory as being based on IBE, we have to accept that it relies on self-evidencing explanation too. This is because we cannot crawl out of our skulls to obtain independent evidence for the veracity of the predictive error coding. This causes the problem of the explanatory-evidentiary circle (EE-circle), according to which, if someone raises novel doubts about the occurrence of an evidence, then the higher posterior probability of the relevant hypothesis, acquired through its explanatory prowess, cannot be used to dispel the doubt. The circularity is vicious under the circumstances because the probability of the occurrence of the evidence cannot be established independently of the hypothesis (Hohwy 2014, p. 264). The EE-circle establishes an evidentiary boundary between the evidence and hypothesis: evidentiary because it is defined by the occurrence of the evidence, boundary because causes beyond it can be only inferred (Hohwy 2014, p. 264). The PEM model is bound to remain internalist and inferentialist in this reading: "mental states do not extend into the environment, and the involvement of the body and of action in cognition can be described in wholly neuronal, internal, inferential terms" (Hohwy 2014, p. 288). Does this show that the upward path to realism can't be traversed? I proceed to explain why this is not the case.

We have to note that in receiving the impact of stimuli the brain operates actively, in the sense of being able to optimise the statistical inferences. The brain encodes a conditional probability density function that reflects the relative probability of this state of affairs as well as its alternatives, given the available information. In this way, PEM theory offers a reliable strategy for decreasing the subjective elements of the inference through approximating reality in the long run. It is the self-evidencing brain's responsibility to approximate the optimal Bayesian inference by inferring its own priors as it goes along by using its best current model at the higher level as the source of the priors on the level below. So, by engaging in a process of "iterative estimation" that allows priors and models to co-evolve across multiple linked layers of processing, the brain succeeds at minimising the prediction error (Clark 2013, p. 3). This confirms the claim that the brain's PEM relies on the tractable biophysical-computational resources so as to form Bayesian inferences and approximate the reality. All in all, PEM theory seems to provide our best account of verisimilitude of perception, not only in virtue of the reliability of the experimental researches that support its main thesis but also on account of its explanatory and unifying power. In this way, PEM theory promises to bring cognition, perception, action, and attention together within a common statistical framework which is induced by hierarchical generative models as our basic means of representing the world (Clark 2013, p. 10). Let me remind that metaphysics of SR (even in its ontic, strong expression) is based on measuring the ontological value of the theories according to their unifying explanatory power (Beni 2016; Ross et al. 2007). In their statement of the metaphysical core of the SR, Ross et al. (2007) remarked that it is the *raison d'^etre* of a useful metaphysics to provide "the conceptual framework within which we then consider relationships among contemporary theories in different sciences so as to construct a unified world picture" (Ross et al. 2007, p. viii). Metaphysics is understood in terms of unification of science according to the authors, and it "consists in maximising the ratio of kinds of phenomena we can explain to the number of kinds of causal processes" (Ross et al. 2007, p. 31). PEM theorists' emphasis on the unifying virtue of PEM meshes quite nicely with SRists' stress on the role of unification. Moreover, it juxtaposes with PEM-theorists attempt at introducing the PEM as the best explanation of several cognitive and perceptual adaptive mechanisms.

To make a long story short, the claim concerning the efficiency of the PEM theory in accounting for the brain-reality relationship could be defended on a philosophical basis. Philosophical reasons persuade us to think of PEM theory as a reliable vehicle that comes in handy in repairing the upward path to SR. The theory is both simple and comprehensive (considering its unifying, explanatory power). Since scientific realists themselves didn't hesitate to use IBE and Bayesian arguments to defend the plausibility of their realist stance, they couldn't protest to the application of IBE and Bayesian model to the case of PEM theory. I proceed to explain this point before ending this section.

A number of credible scientific realists used IBE and Bayesian inference to establish the viability of the realist viewpoint. Psillos' (1999) own hint at Maxwell's (1970) Bayesian defence of realism (Maxwell 1970) could be quite illuminating here. Maxwell took instrumentalism and realism as the competing theories that have to entail the data about the success of scientific theories. According to Maxwell, when two inconsistent hypotheses entail the same piece of evidence, the only way in which the evidence can be pushed to support one hypothesis in the face of alternative is via weighing up the initial probabilities of the competing hypotheses. Bayes theorem has to be called into play. The prior probabilities matter, because evidently, the posterior probabilities of the success of empirically equivalent hypotheses (i.e. realism and instrumentalism) are equal. Maxwell's claim is that if we compare the initial probabilities of realism and instrumentalism in this system, it turns out that prior probability of realism should be greater than the prior of instrumentalism. Maxwell invoked the criteria of simplicity and comprehensiveness and claimed that these support the realist hypothesis in the face of the instrumentalist one (see Psillos 1999, p. 72). Interestingly enough, to sanctify this Bayesian approach and to declare the attributed prior probabilities are not mere *subjective* degrees of belief, Psillos argued that assignment of higher probability to realism is *rational*, and hence *objective* (Psillos 1999, p. 73). I don't intend to evaluate the validity of this kind of argumentation. I only want to remind that the scientific realists did not hesitate to use Bayesian model to justify the realist hypothesis.

Scientific realists have also appealed to approximate truth and IBE in the way of establishing the realist stance. As Psillos remarked, the "best explanation of the instrumental reliability of scientific methodology is that background theories are relevantly *approximately true*" (Psillos 1999, p. 77 my emphasis). So, the realist is willing to make an explanatory connection between the empirical success of the theories and the correctness of the theories' reports about the unobservable parts of the world, without going so far to assert that everything that theory is claiming about the world is precisely true. Realist's resolve remains unshakable in the face of worries voiced by the antirealists (Fine 1986; van Fraassen 2008).

Now, returning to the upward path to SR, it could be observed that precisely the same kind of arguments could be invoked to defend the integrity of this empirical approach to SR. To establish the link between the structure of perceptions and the causal structure of stimuli, PEM theory mainly relies on the IBE and Bayesian model. That being so, it should pass the filter of scientific realists who appeal to precisely the same criteria in order to endorse the realist stance in the face antirealism. There are philosophical reasons to prove PEM theory come in handy for the purpose of reconstructing the upward path to SR. The upward path to SR is repairable. Q.E.D.

7 Concluding remarks: in defence of roughly one-one relation

I drew on the sparse coding and difference-based coding theories to argue that mechanisms of perception could be best interpreted along the lines of structuralism.

After that, to establish the realist component of the upward path to SR, I made use of the interactive model of the brain and PEM theory. We saw that the brain is endowed with the interesting capacity to minimise the discrepancy between its perceived inputs and the actual, stimuli-induced inputs. The plausibility of PEM theory could be defended on experimental grounds. But the experimental arguments could also be reinforced by philosophical reasoning, e.g. IBE. As scientific realism has been defended on the basis of similar arguments, the validity of the arguments can hardly be denied by the scientific realists. The same criteria of rationality and objectivity that supports scientific realist claim could be conjured in the way of establishing the plausibility of the methods that we used to repair the upward path to SR.

Acknowledgements Georg Northoff read an earlier version of this paper and made some useful comments. Also, the two anonymous referees of this journal made constructive comments that helped to promote the style and quality of the arguments. All of these debts are gratefully acknowledged.

References

Barlow, H. B. (1972). Single units and sensation: a neuron doctrine for perceptual psychology? *Perception, 1*(4), 371–394.

Beni, M. D. (2016). Structural realist account of the self. *Synthese, 193*(12). Springer Netherlands: 3727–40. doi:10.1007/s11229-016-1098-9.

Blakemore, S.-J., Wolpert, D., & Frith, C. (1999). The cerebellum contributes to somatosensory cortical activity during self-produced tactile stimulation. *NeuroImage, 10*(4), 448–459.

Blakemore, S.-J., Wolpert, D., & Frith, C. (2000). Why can't you tickle yourself? *Neuroreport, 11*(11), R11–R16.

Brenner, N., Bialek, W., & de Ruyter van Steveninck, R. (2000). Adaptive rescaling maximizes information transmission. *Neuron, 26*(3), 695–702.

Bueno, O., French, S., & Ladyman, J. (2002). On representing the relationship between the mathematical and the empirical. *Philosophy of Science, 69*, 497–518.

Chalmers, A. F. (1970). Curie's Principle. *The British Journal for the Philosophy of Science, 21*(2), 133–148. doi:10.1093/bjps/21.2.133.

Clark, A. (2008). *Supersizing the mind: embodiment, action, and cognitive extension*. Oxford: Oxford University Press.

Clark, A. (2012). Dreaming the whole cat: generative models, predictive processing, and the Enactivist conception of perceptual experience. *Mind, 121*(483), 753–771. doi:10.1093/mind/fzs106.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences, 36*(3), 181–204. doi:10.1017/S0140525X12000477.

Curie, P. (1894). Sur la symétrie dans lesphéno enes physiques, symétrie d'un cham electrique et d'un champ magnétique. *J. Phys. Theor. Appl, 31*(31), 393–415. doi:10.1051/jphystap:018940030039300>.

da Costa, N. C. A., & French, S. (2003). Science and partial truth. Oxford University Press, Oxford. doi:10.1093/019515651X.001.0001

Duncan, N. W., Hayes, D. J., Wiebking, C., Tiret, B., Pietruska, K., Chen, D. Q., et al. (2015). Negative childhood experiences alter a prefrontal-insular-motor cortical network in healthy adults: a preliminary multimodal rsfMRI-fMRI-MRS-dMRI study. *Human Brain Mapping, 36*(11), 4622–4637. doi:10.1002/hbm.22941.

Earman, J. (2004). Curie's principle and spontaneous symmetry breaking. *International Studies in the Philosophy of Science, 18*(2–3), 173–198. doi:10.1080/0269859042000311299.

Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America A, 4*(12), 2379. doi:10.1364/JOSAA.4.002379.

Fine, A. (1986). Unnatural attitudes: realist and instrumentalist attachments to science. *Mind , 95*(378), 149–179. doi:10.1093/mind/XCV.378.149.*XCV*

French, S. (2014). The structure of the world metaphysics and representation. *Journal of Chemical Information and Modeling, 53*. doi:10.1017/CBO9781107415324.004.

French, S., & Ladyman, J. (2003). Remodelling structural realism: quantum physics and the metaphysics of structure. *Synthese, 136*(1), 31–56. doi:10.1023/A:1024156116636.

Grammont, F., & Riehle, A. (2003). Spike synchronization and firing rate in a population of motor cortical neurons in relation to movement direction and reaction time. *Biological Cybernetics, 88*(5), 360–373. doi:10.1007/s00422-002-0385-3.

Hassoun, M. H. (1993). *Associative neural memories: theory and implementation. Associative neural memories*. Oxford: Oxford University Press.

Hempel, C. (1965). *Aspects of scientific explanation and other essays in the philosophy of science*. New York: Free Press.

Hohwy, J. (2013). The predictive mind. Oxford University Press, Oxford doi:10.1093/acprof:oso/9780199682737.001.0001

Hohwy, J. (2014). The self-evidencing brain. *Noûs, 50*(2), 259–285. doi:10.1111/nous.12062.

Huang, Y., & Rao, R. P. N. (2011). Predictive coding. *Wiley Interdisciplinary Reviews: Cognitive Science, 2*(5), 580–593. doi:10.1002/wcs.142.

Ismael, J. (1997). CURIE'S principle. *Synthese, 110*(2), 167–190. doi:10.1023/A:1004929109216.

Kilner, J. M., Friston, K. J., & Frith, C. D. (2007). Predictive coding: an account of the mirror neuron system. *Cognitive Processing, 8*(3), 159–166. doi:10.1007/s10339-007-0170-2.

Ladyman, J. (2008). Structural realism and the relationship between the special sciences and physics. *Philosophy of Science, 75*(5), 744–755. doi:10.1086/594520.

Lam, V., & Esfeld, M. (2012). The structural metaphysics of quantum theory and general relativity. *Journal for General Philosophy of Science, 43*(2), 243–258. doi:10.1007/s10838-012-9197-x.

Laudan, L. (1981). A confutation of convergent realism. *Philosophy of Science, 48*(1), 19–49. doi:10.1086/288975.

Lindenberger, U., Li, S.-C., Gruber, W., Müller, V., Decety, J., Chaminade, T., et al. (2009). Brains swinging in concert: cortical phase synchronization while playing guitar. *BMC Neuroscience, 10*(1), 22. doi:10.1186/1471-2202-10-22.

Maxwell, G. (1970). Theories, perception and structural realism. In R. Colodny (Ed.), *The nature and function of scientific theories* (pp. 3–34). Pittsburgh: University of Pittsburgh.

Molotchnikoff, S., & Rouat, J. (2012). Brain at work: time, sparseness and superposition principles. *Frontiers in Bioscience, 17*(1), 583. doi:10.2741/3946.

Newman, M. H. A. (1928). I.—Mr Russell's & quot; causal theory of perception&quot. *Mind* , (146), 137–148. doi:10.1093/mind/XXXVII.146.137.Xxxvii

Northoff, G. (2012). Immanuel Kant's mind and the brain's resting state. *Trends in Cognitive Sciences, 16*(7), 356–359. doi:10.1016/j.tics.2012.06.001.

Northoff, G. (2014a). *Unlocking the brain: volume 1: coding*. New York: Oxford University Press.

Northoff, G. (2014b). *Unlocking the brain: volume 2: consciousness*. New York: Oxford University Press. doi:10.1093/acprof:oso/9780199826995.001.0001.

Northoff, G. (2016a). Neuroscience and whitehead I: Neuro-ecological model of brain. *Axiomathes*. doi:10.1007/s10516-016-9286-2.

Northoff, G. (2016b). Neuroscience and whitehead II: process-based ontology of brain. *Axiomathes*, 1–25. doi:10.1007/s10516-016-9287-1.

Olshausen, B. A., & Field, D. J. (1997). Sparse coding with an overcomplete basis set: a strategy employed by V1? *Vision Research, 37*(23), 3311–3325.

Olshausen, B. A., & Field, D. J. (2004). Sparse coding of sensory inputs. *Current Opinion in Neurobiology, 14*, 481–487. doi:10.1016/j.conb.2004.07.007.

Poo, C., & Isaacson, J. S. (2009). Odor representations in olfactory cortex: "sparse" coding, global inhibition, and oscillations. *Neuron, 62*(6), 850–861. doi:10.1016/j.neuron.2009.05.022.

Psillos, S. (1999). *Scientific realism: how science tracks truth*. New York: Routledge.

Psillos, S. (2001). Is structural realism possible? *Philosophy of Science, 68*(S3), S13–S24. doi:10.1086/392894.

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

Roberts, B. W. (2013). The simple failure of Curie's principle. *Philosophy of Science, 80*(4), 579–592. doi:10.1086/673212.

Ross, D., Ladyman, J., Collier, J., & Spurrett, D. (2007). *Every thing must go*. Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780199276196.001.0001.

Russell, B. (1917). Notion of Cause. In Mysticism and Logic and Other Essays (pp. 173–199). London: Orwin.
Russell, B. (1919). *Introduction to mathematical philosophy*. London: George Allen & Unwin.
Russell, B. (1927). *The analysis of matter*. London: Kegan Paul.
Sadaghiani, S., & Kleinschmidt, A. (2013). Functional interactions between intrinsic brain activity and behavior. *NeuroImage, 80*, 379–386. doi:10.1016/j.neuroimage.2013.04.100.
Sänger, J., Müller, V., & Lindenberger, U. (2012). Intra- and interbrain synchronization and network properties when playing guitar in duets. *Frontiers in Human Neuroscience, 6*, 312. doi:10.3389/fnhum.2012.00312.
Simoncelli, E. P., & Olshausen, B. A. (2001). Natural image statistics and neural representation. *Annual Review of Neuroscience, 24*(1), 1193–1216. doi:10.1146/annurev.neuro.24.1.1193.
Teller, P. (2011). Two models of truth. *Analysis, 71*(3), 465–472. doi:10.1093/analys/anr049.
van Fraassen, B. C. (2008). Scientific representation. Oxford University Press, Oxford doi:10.1093/acprof:oso/9780199278220.001.0001
Vinje, W. E., & Gallant, J. L. (2000). Sparse coding and decorrelation in primary visual cortex during natural vision. *Science (New York, N.Y.), 287*(5456), 1273–1276.
Votsis, I. (2005). "The Upward Path to Structural Realism." Philosophy of Science 72(5): 1361–72. doi:10.1086/508974.
Weisberg, M. (2013). *Simulation and similarity using models to understand the world*. Oxford: Oxford University Press.
Worrall, J. (1989). Structural realism: the best of both worlds? *Dialectica, 43*(1–2), 99–124. doi:10.1111/j.1746-8361.1989.tb00933.x.
Zylberberg, J., Murphy, J. T., & DeWeese, M. R. (2011). A sparse coding model with Synaptically local plasticity and spiking neurons can account for the diverse shapes of V1 simple cell receptive fields. *PLoS Computational Biology, 7*(10), e1002250. doi:10.1371/journal.pcbi.1002250.