


ARTICLE

Cognitive Penetration and Cognitive Realism

Majid D. Beni 

Department of Philosophy, Middle East Technical University, Ankara, Turkey
Email: m Beni@metu.edu.tr

(Received 5 April 2021; revised 25 May 2021; accepted 8 July 2021)

Abstract

The paper addresses the issue of theory-ladenness of observation/experimentation. Motivated by a naturalistic reading of Thomas Kuhn's insights into the same topic, I draw on cognitive neuroscience (predictive coding under Free Energy Principle) to scrutinise theory-ladenness. I equate theory-ladenness with the cognitive penetrability of perceptual inferences and argue that strong theory-ladenness prevails only under uncertain circumstances. This understanding of theory-ladenness is in line with Thomas Kuhn's view on the same subject as well as a cognitive version of modest realism rather than downright antirealism.

Keywords: Cognitive penetration; cognitive structural realism; free energy principle; theory-ladenness of observation

1. Introduction

It may be justifiably assumed that scientific realism is somehow related to perceptual realism, in the following sense. The validity of representational credentials of observational chunks of scientific theories is taken for granted by scientific realists. The situation of theoretical terms is more complicated. After all, realists usually need to put forth adroit strategies to establish quasi-reality of theoretical parts of the theory.¹ And it happens to be the case that the observational/theoretical distinction has become a relic of bygone ages. However, it is not clear to what extent theories load/lead the observations, nor are the bearings of theory-ladenness completely clear for the realist vs. antirealist debate. The paper focuses on two interrelated issues in this context. First, there is the question of the scope of theory-ladenness. The paper draws on scientific evidence to deal with this question. It equates theory-ladenness with penetration of perceptual inferences by cognition and builds upon scientific evidence from predictive coding under Free Energy Principle to show that theory-ladenness is not usually strong. Second, the paper argues that a moderate form of theory-ladenness is in line with a modest version of cognitive realism. The paper incorporates insights from recent

¹Instead of outright realism about the referential prowess of theoretical terms (or realism about the referents of those terms), scientific realism has been evolved into sophisticated stances on laws and dispositions (Mackie 1977; Mumford 2006), properties (Chakravartty 2007) and structures (French 2014). These variegated approaches to realism, amongst others, take the representational veracity of observational terms for granted.

computational neuroscience as well as some philosophical material from Thomas Kuhn's discussion of theory-ladenness to substantiate its arguments.

The relation between theory-ladenness and scientific realism can be complicated. The question of how to set a meaningful distinction between the observable and the unobservable has been perennial in the history of the philosophy of science, both for the pioneering philosophers of science and those who moved in their wake (Carnap 1956; Maxwell 1962, 1971; van Fraassen 1980; Psillos 2000b; Friedman 2011). Denying that there is such a distinction, or assuming that observations/experimentation is led/loaded by theories can result in intricate implications that may lend support to diverse and at times conflating stances, from diehard realism to downright antirealism. Tampering with the border between observable and unobservable may also lead to accounts that can be construed as sophisticated forms of realism or non-realism. A good example of a sophisticated stance is Thomas Kuhn's (1962) theory of science which is centred on his insights into theory-ladenness and incommensurability of theories across different historical/social contexts. The present paper draws on some recent breakthroughs in cognitive/computational neuroscience to shed some light on the issue of theory-ladenness, its scope, and its bearing on scientific realism.

The paper would unearth the psychological root of theory-ladenness by discussing the issue of cognitive penetrability of perceptual inferences under the Free Energy Principle (FEP). FEP underlies a flourishing approach to cognition, action, perception, etc., in terms of minimisation of the free energy or prediction error (Friston 2012, 2013, 2019). The paper will argue that the picture of theory-ladenness (or cognitive penetrability) under FEP is too sophisticated to lead to an outright confirmation or rejection of scientific realism. The relevance of the present enterprise to the works of Kuhn is this. Kuhn, too, has conjured scientific evidence from Gestalt psychology in support of his views on theory-ladenness and incommensurability. In the wake of Kuhn, others, such as Churchland (1979) and Fodor (1984) have discussed theory-ladenness with an eye to evidence from experimental psychology. The topic has been of philosophical interest until very recently (Brewer and Lambert 2001; Karaca 2013; Franklin 2015). Against this background, the paper would pursue its twin goals. Firstly, it draws on some advancements in cognitive neuroscience to argue that strong cognitive penetrability/theory-ladenness is not a prevalent phenomenon. Secondly, the paper argues that weak cognitive penetrability/theory-ladenness is compatible with modest forms of cognitive realism, as being pioneered by Giere (1988, 1992) and restored by myself recently (Beni 2019a). The paper also examines its suggestions against a naturalist reading of Kuhn's insights into theory-ladenness. Kuhn's reliance on psychology sets a precedent for the scientifically informed approach of the paper to theory-ladenness, and, the moderate form of cognitive realism that is advocated here is in harmony with Kuhn's naturalism. The historical connection between cognitive realism and Kuhn's ideas (through Bird's naturalist reading of Kuhn and Paul Churchland's and Ronald Giere's cognitive approaches to science) is canvassed in broadest terms, whereas psychological grounds for theory-ladenness are explored closely.

The paper is structured in the following manner. It discusses issues of theory-ladenness and incommensurability. After establishing that it is alright to draw on cognitive psychology (and in this case, cognitive penetrability) to discuss the issue of theory-ladenness, the paper canvasses the Free Energy Principle and unfolds the implications for theory-ladenness. After arguing that scientific evidence supports only a weak version of theory-ladenness, the paper unfolds the consequences for the realist vs. anti-realist debate to indicate that the FEP-based account of theory-ladenness is in line with a moderate cognitive form of realism rather than full-blooming versions of either realism or antirealism.

2. Theory ladenness and cognitive penetration

Generally, scientific realism conveys a commitment to the existence of the mind-independent world (this is the ontic component). It also includes epistemic and semantic theses, indicating that science provides reliable knowledge of the world and that theories are truth-apt and theoretical posits have veridical references. There might have been versions of realism and antirealism that are committed to a clear distinction between observational and theoretical parts of theories.² But now it is generally believed that even observational parts of theories are (at least to some extent) laden by the theoretical content, and as such, there is no clear distinction between the observational and the theoretical. This thesis/conjecture is called ‘theory-ladenness of observation/experimentation’, and we just call it ‘theory-ladenness’ for the sake of brevity. There are interesting ramifications for the philosophy of science. On the one hand, rejecting the observational/theoretical dichotomy may support scientific realism as opposed to versions of empiricism that are committed to the referential capacity of the observational but not theoretical parts of theories. At the same time, because theory-ladenness rejects the possibility of testing the theory by theory-independent observations or experimentations, it may well result in scientific antirealism. All that I said about theory-ladenness also applies to the issue of cognitive penetrability of perceptual inferences. Hence the relation between perceptual realism – which is concerned with “whether perceptual beliefs are true” (Raftopoulos 2008: 64) – and scientific realism. Perceptual realism consists of two theses: firstly, experience has nonconceptual content, and secondly, nonconceptual content of experience represents the structure of the mind-independent world veridically (also see Raftopoulos 2008). The assumption of the cognitive penetration of perceptual inference infringes the first thesis. This paper considers the issue of cognitive penetration of perception in tandem with the bearing of theory-ladenness (and the range of its scope) on scientific realism (assuming that ‘cognitive penetration’ can be equivocated with ‘theory-ladenness of observation’).

The relationship between cognitive penetrability and scientific realism has been already discussed in the literature. Fodor (1984) and Kitcher (2001), amongst others, have endeavoured to unpack the ramifications of psychological accounts of background categories and preconceptions for the scientific realism vs. constructivism debate. Even some of the most interesting claims of no less significant figure than Kuhn (1962: 62–3) himself draw on psychological studies³ to establish the point that perceptions or observations are mediated by conceptual schemes of the agents.⁴ Kuhn has developed this insight into his account of ‘incommensurability’, which aims to characterise the conceptual disparity across the preceding and succeeding scientific paradigms that are

²Such naïve realist views may accept the dichotomy, but assume that observational terms and theoretical terms both refer to real entities (of observational and non-observational kinds). Naïve empiricism may indicate that only observational terms but not theoretical terms have veridical references. There have been scientific empiricists who toyed with the idea of reducing theoretical parts to observational parts (for example, by using observational terms in reductive pairs (Carnap 1936) or invoking the Ramsey-sentence approach, by showing how the theoretical sentences of the theories could be existentially quantified without making specific commitments to particular theoretical entities, given that the experimental content of the theory could be exhausted in terms of the referential capacity of the observational components of the theory (Psillos 2000a; Friedman 2011).

³Such as Jerome Bruner’s and Leo Postman’s (1949) experiment about identifying anomalous cards such as the black four of heart and the red six of spades.

⁴Kuhn drew on these cases to argue that scientists’ experience of the evidence is intermediated by their theoretical schemes. For example, Kuhn argues that in the case of combustion for example, Priestley and Lavoisier were considering the same phenomena (now called oxidation) as the evidence for their diverse accounts of the combustion process (Kuhn 1962: 59).

disjoined across scientific revolutions (Kuhn 1962: 148 ff). He refers to cognitive and Gestalt psychology – in works of Bruner and Postman – in support of his statement of theory-ladenness and incommensurably (Kuhn 1962: 62, 111). Bird (2012) argues that these attempts result in an unappreciated form of Kuhnian naturalism that predated Quine's naturalised epistemology. According to Bird (2005), owing to practical reasons – such as the urge to have the philosophical credentials of his work recognised by his contemporaries – Kuhn has ignored the earlier naturalist leaning of his conception of theory-ladenness. Bird's work aims to bring this naturalist aspect into focus.

Before going further, I just point out that naturalism – which is also connoted by Kuhn's work – may be at least to some extent compatible with realism because in some sense it invites trusting in scientific theories as the venues of providing reliable guidelines to philosophy. If that is true, the existence of naturalist elements in Kuhn's work indicates that his work is not totally incompatible with a moderate form of (cognitive) realism. In time, this psychologically informed account of a theories-world relationship evolves into a version of scientific realism that builds upon scientific grounds (from cognitive science and psychology) to substantiate insights into the relationship between theories and the objective structure of the world (Giare 1988; Beni 2019a).

3. Kuhnian naturalism and cognitive realism

Kuhn's views on incommensurability and theory-ladenness do not indicate that science is irrational (Hoyningen-Huene 1993), nor do they lead to downright antirealism. It is true that Kuhn's ideas have evolved into confuting convergent realism – according to which science gets ever closer to the truth (Laudan 1981) – or social constructivism – which makes scientific knowledge mainly a matter of social factors and forces (Latour and Woolgar 1979). However, Kuhn (1992) himself repudiates strong social constructivism, which according to him makes science a matter of power and interest, rather than rationality, evidence, and representing nature. The relationship between scientific realism and Kuhn's ideas is not straightforward, especially as his enterprise could be construed as an attempt at demystifying scientific rationality, rather than dethroning it. This point has also been nailed down in Kuhn's (1977: Ch. 13) own later interpretation of the original statement of his views, where he remarks that the invitation to trust the collective judgment of the scientists does not make theory choice a matter of mob psychology. Although (in light of theory-ladenness of theories) it cannot be assumed that there is decisive proof or theory-independent evidence to determine the choice of theories or paradigms, Kuhn concedes that the choice is constrained by some maxims, norms, or values. The list, which is not exhaustive, includes notions of accuracy, consistency (with other theories), broadness of scope, simplicity, and fruitfulness (Kuhn 1977: 321–2). None of these maxims is sufficient for objectively dictating the choice of theories, and none of them can be interpreted uniquely. Even when applied collectively, the said maxims do not eliminate the problem of choice for individual scientists. But this does not mean that the choice is totally drained from objectivity. In fact, Kuhn questions the plausibility or usefulness of an idealized distinction between subjective factors and objective ones (that is to say, he challenges the distinction between the context of discovery and the context of justification). Accordingly, although Kuhn's theory does not deny the role of personal and idiosyncratic factors in theory choice, it does not lead to downright antirealism because according to Kuhn (1977: 325), “every individual choice between competing theories depends on a mixture of objective and subjective factors, or of shared and individual criteria”. If that is true, a moderate cognitive realist reading of the relation between theories and reality does not directly defy Kuhn's theory of

science. The version of cognitive realism that I have in mind is mainly inspired by works of Giere (1988) (which might be somehow in line with Kuhn's late views (Giere 2013)). But before speaking about cognitive realism, I will elaborate a bit more on the naturalist interpretation of Kuhn's ideas.

We cannot continue by scrutinising the thorny issue of 'scientific paradigm' in the context of Kuhn's work. That issue is simply too complicated to be unfolded here. Suffice it to say that paradigms afford puzzles that scientists are supposed to solve, given that scientific activity mainly consists in puzzle solving (Kuhn 1962: 36 ff). Solutions to the puzzles, too, are supposed to be evaluated against the criteria that the paradigm provides. Roughly stated, there are no objective criteria such as observation or experiment to evaluate the paradigms. Observations/experimentations are laden with theories that are firmly instilled in the cognitive system of scientists through rigorous training and indoctrination processes. Scientists could not conjure theory-independent evidence in support of their theories. Diverse paradigms are *incommensurable*, they are taxonomically and conceptually separated from one another. Paradigms are indeed prone to expiration and change, say when instances of crisis escalate and anomalies amass beyond a certain threshold. But as long as scientists work within a paradigm, it is rather the paradigm that provides the standard for gauging the eligibility of scientists (this means that the paradigm tests the scientists, not the other way around). It is important for our purpose in this article to note that puzzle-solving within a paradigm is not a matter of the application of rules of rationality, but it is a matter of pattern recognition (under Gestalt psychology according to Kuhn) (Bird 2005: 103). For Kuhn pattern recognition under a paradigm takes the place of the application of rational rules. Bird also argues that since Kuhn replaces scientific rationality with puzzle-solving and pattern recognition, it would be "natural to appeal to connectionism (neural network models, parallel distributed-processing (PDP)) as showing how it is realised in the scientist's brain" (Bird 2005: 109). This is the link between the naturalist reading of Kuhn's and a version of cognitive realism that will be introduced in the final section of the paper. I use this warrant, about it being natural to appeal to connectionism, to support a weak notion of theory-ladenness by recent developments in unsupervised learning in neural networks.

Let us recap before proceeding. The relation between neural network models and Kuhn's notion of scientific activity is this: Kuhn specifies scientific activity both at the individual and large-scale levels as pattern recognition. Because neural distributed processing underlies reliable accounts of sequential thinking and pattern recognition, it can provide a nice formal-scientific framework for regimenting Thomas Kuhn's view on scientific practice. For example, paradigm shifts can be regimented in terms of the shift in the configuration of activation spaces in a connectionist network. In the same vein, incommensurability can be explicated in terms of "a difference in pattern-recognition capacities and cognitive habits – between subjects" (Bird 2005: 113). Tacit assumptions and inferential habits that underlie pattern recognition are mapped onto the configuration of activation spaces in connectionist networks and owing to having different patterns of activation spaces, subjects could be in disagreement about what is a reasonable solution to a problem (hence a PDP formulation of Kuhn's insights is viable).

Let us connect Bird's interpretation of Kuhn's work to the enterprise of the present paper. To the extent that embodied cognitive organic cognitive systems are at issue, neural networks are implemented in organisms on the basis of unsupervised learning in hierarchically organised parallel distributed organization with the capacity of forming generative models and optimising them. A state-of-the-art account of PDP in organic systems has been presented in terms of minimising variational free energy (see Kiefer

2020). As Bird has already said, it is natural to draw on PDP and connectionism to explicate Kuhn's account of pattern recognition. If that is the case, it will as well be appropriate to draw on more recent accounts of pattern recognition in unsupervised networks to account for theory-ladenness. Hence the relevance to Kuhn's ideas to the present project (more on this later in the paper). In the remainder of the paper, I first explore state-of-the-art cognitive sciences to shed some light on the murky issue of theory-ladenness, and then, after showing that at least some flourishing theories of psychology do not support strong theory-ladenness, I present a moderate realist account of the theories-world relationship. This enterprise is inspired by Kuhn's scientifically informed approach to theory-ladenness.

4. Free energy principle and predictive processing

Thus far, I have argued that it is legitimate to draw on theories from cognitive psychology to enlighten the discussion of theory-ladenness. Thomas Kuhn had drawn on theories of Gestalt psychology, and Bird furthered Kuhn's discussion by considering PDP in neural networks in terms of connectionism. In order to make an improvement on this discussion, I will focus on predictive coding under Free Energy Principle (FEP). The relevance of using FEP in the present context is that not only does the theoretical framework of FEP include an updated version of unsupervised learning in neural networks (which is used by Bird to reinforce a naturalist reading of Kuhn's work), it also furthers some seminal insights of Gestalt psychology (which are used by Kuhn in support of his account of theory-ladenness). Orlandi (2016) provides an interesting account of the historical relationship between the Bayesian accounts of perception (such as FEP), connectionism, and Gestalt psychology. The general insight behind Orlandi's discussion is that all of these accounts are loyal to the basic insight of (Kohler's version of) Gestalt psychology, according to which "the brain is a dynamic, physical system that converges toward an equilibrium of minimum energy" (Orlandi 2016: 348). This means that the mode of unsupervised learning in connectionist networks is in harmony with this Gestalt idea, given that networks strive to reduce errors by adjusting connection strengths based on feedback from the environment (or training set). FEP provides an advanced account of processes that have been already contrived in Gestalt psychology and connectionism. We will be able to elaborate on the relationship between FEP and Gestalt theory later in this section, after saying a bit more on the free energy principle. For the time being, suffice to say that the present paper draws on computational neuroscience (and FEP in particular) to inform the philosophical discussion of theory-ladenness (by the same token, FEP can be used to illustrate the discussion of the theories-reality relationship). Below, I shall furnish some further details about predictive coding under FEP.

Friston and colleagues have discovered and presented the Free Energy Principle (FEP) (Friston 1994, 2010; Friston *et al.* 2006) as an equation that unifies accounts of cognition and action. They argue that FEP provides a grand unifying framework that comprises theories of cognition, perception, learning, emotion, memory, consciousness, and action and integrates them (Hohwy 2013; Clark 2016; Ramstead *et al.* 2017; Friston *et al.* 2020; Seth and Hohwy 2020) (there are of course also those who question the unificatory scope of FEP, for discussion see Beni (2018, 2019b), Colombo and Hartmann (2015) and Colombo and Wright (2018)). The general idea behind FEP is that organisms tend to resist the dispersing force of the environment by achieving homeostasis or reaching non-equilibrium steady states with their environment. If organisms do not reach homeostasis, their existence will be terminated. In order to resist dissolving, the organisms must make their internal physical and chemical

milieu stable enough. To do so, they must avoid surprise, which is the negative log probability of an outcome. Organisms systematically avoid surprises by putting an upper bound on their internal entropy, which is the average of the surprise (Friston 2010: 127). This, namely minimising free energy, is what keeps the organism in non-equilibrium steady states. Accordingly, making the *internal* milieu stable enough mandates setting some sort of barrier between internal and external states. The encroachment upon the boundary would lead to endangering the steadiness of the states of the system. The requisite boundary has been characterised in terms of Markov blankets (Kirchhoff *et al.* 2018; Parr *et al.* 2020).

Markov blankets are used to model the conditional dependencies between internal and external states. The Markov blanket of a node consists of the node, its parents, its children, and co-parents of its children. Parents are the cause of children and blankets model patterns of conditional dependencies around the node. Blankets consist of sensory states and active states, where the former states affect the internal states but are not affected by them, and the active states affect the external states but are not affected by them. Given that internal and external states do not affect one another directly, Markov blankets serve to model conditional independence/dependence between the internal and the external (Hipólito *et al.* 2021). The organism's interaction with the environment – how the organism garners evidence for its existence by actively searching the environment and how it decreases the free energy by acting on the environment – can be modelled by using Markov blankets. The question of whether Markov blankets are embodied (or real/physical) entities within organisms or they are just abstract modelling tools has been the subject of some debate (Bruineberg *et al.* 2020; van Es and Hipólito 2020; Beni 2021).

It is also worth mentioning that FEP not only underpins a theoretical framework for explicating life, it has also been applied to the brain. The brain is in charge of controlling the action of the organism and regulating its interaction with the environment. Prediction error minimization is roughly the same thing as free energy minimisation. The brain contributes to maximising the survival of the organism by optimising the precision of the organism's representations of the causal structure of the world. To do so, the brain does not just passively receive raw data from the environment. It forms predictions about the environment first and then, through top-down approximate Bayesian mechanisms that are realised in a hierarchical structure, it minimises the discrepancy between predictions and their target systems in the environment. The discrepancy will be minimized by either updating generative models or changing the world (Limanowski and Friston 2020). This model specifies the brain “as a hierarchical filter that takes the non-linear time series of sensory input and seeks to filter out regularities at different time scales” (Hohwy 2017: 77). Across this hierarchical organization, the brain applies approximate Bayesianism to calculate the discrepancy between its generative models – internal models of likelihood of data, given their causes – and the causal structures in the environment. Then the organism minimises the discrepancy by either updating models of priors in the brain or changing the environment.

Let consider the resemblance that the FEP framework bears to Gestalt psychology once more. According to FEP, the organism aims to stay in non-equilibrium steady states by minimising its free energy. In this context, the brain is the approximate Bayesian inference machine that minimises the prediction error of generative models. This bears resemblance to Köhler's (2007 [1920]) idea of the brain as a holistic physical system (or a physical Gestalt) that evolves towards stable states whose energy is minimum. In fact, Wagemans *et al.* (2012: 86) suggest that to amend for the empirical shortcoming of the original Gestalt account of brain functions, we may implement physical Gestalt “as recurrent networks with closed feedback loops, which can be proven

to converge to an equilibrium state of minimum energy” (as stated by FEP) (also see Van de Cruys and Wagemans 2011). More generally, according to a version of Gestalt psychology that is represented by Max Wertheimer, Kurt Koffka, and Wolfgang Kohler, “perception was not the product of sensations but it arose through dynamic physical processes in the brain” (Wagemans *et al.* 2012: 4). And this notion speaks directly to the FEP-based account of perception according to which the brain is capable of optimising a (free-energy) bound on surprise or its complement (Friston 2010). Of course, none of these points is directly about Kuhn’s views on theory-ladenness, but the seminal point about the resemblance between the physical Gestalt and the Bayesian brain under FEP could be developed into accounts of perception that support Kuhn’s insights into theory-ladenness. For example, FEP’s account of cognition does not begin from the stimuli but from the generative models which are in the business of forming predictions about the world at the top levels of cognitive architecture. Past experiences are accumulated in these generative models, and as such, experience is not directly dependent on the unadulterated intrinsic properties of the stimuli, but (also) on patterns of correlations in past experiences as well as the overall neural structure and the global activity of the cognitive system. This means that the interpretation of stimuli is penetrated by the configuration of the activation patterns in the cognitive architecture, which provides a basis for saying that observation is theory-laden. We will delve into this account of theory-ladenness in the next section.

Let us recap. FEP presents a holistic account of cognition/perception according to which the perceptual experience is not decomposable to the received sensory inputs. This occasions a resemblance to ideas of Gestalt theorists, who argued that “the world that appears to us is not a literal copy of the external world but that there are events beyond the visual input, in the brain, which are responsible for the structural appearance that the world has in our consciousness” (Van de Cruys and Wagemans 2011: 239). Consequently, FEP can be conjured to support the reconstruction of Kuhn’s accounts of theory-ladenness and incommensurability, in the same capacity that Gestalt theories have been applied in Kuhn’s original statement of his views.

5. The FEP-based account of theory-ladenness

Brewer and Lambert (2001) (whose work inspired the present enterprise) have investigated the bearing of cognitive psychology on the analysis of theory-ladenness. Their reference to cognitive psychology is based on the theoretical models of interaction of top-down theoretical information and bottom-up sensory information, but they also draw on specific experiments – such as the Müller-Lyer illusion (which is also discussed by Fodor 1984) and the old woman/young woman picture. Brewer and Lambert discuss these experiments to conclude that perception and attention are indeed theory-laden, but they give a spin to their argumentation to assert that instances of theory-ladenness are strong only when the perceptual evidence is ambiguous or degraded. For example, the observations of states of Pluto could be laden by theories only when there are fuzzy regions on photos of Pluto. We will proceed to consolidate results that are extracted by Brewer and Lambert. To do so, we discuss the theoretical implications of FEP as a grand theoretical framework rather than focusing on specific experiments. The result may be worthwhile because the FEP framework provides a solid foundation for construing diverse experiments that are relevant to the issue of theory-ladenness. Clark (2016: 54 ff.), for example, draws attention to a few experiments. However, we can make sense of the diverse experiments best when we use a theoretical framework that accounts for why human cognitive faculties reply to the experiment’s setting in the way that they do. We rely on FEP (rather than individual experiments) to account

for theory-ladenness in the same fashion that Kuhn drew on Gestalt psychology. I assume that relying on a theoretical framework rather than individual experiments provides a more profound insight into theory-ladenness. Let me elaborate.

One of the several experiments that are mentioned in Clark's account of FEP's top-down architecture concerns the sine-wave speech (Remez and Rubin 1990). In the first stage of the experiment, a vaguely audible replica of the normal speech that is stripped of almost all recognisable sounds and retains only a skeletal acoustic outline is played for the subjects. Not surprisingly, subjects mainly hear incomprehensible noises. In the second round, the experimenter lets the subjects first hear the original soundtrack – in which the sentences are clearly audible – and then plays the sine-wave replica. This time, the subjects can hear the sounds of clearly intelligible speech even when listening to the sine-wave replica. The conclusion is that theories have a top-down influence on the subject's sensory perceptions. Clark (2016: 54 ff.) also mentions some other experiments that confirm this conclusion (the initially indistinguishable picture of the shadowed Dalmatian dog is another famous example). The take-home point is that “a whole pantheon of knowledge-based and contextual effects upon perceptual inference, since it makes what we know (both consciously and, more often, non-consciously) about the world a prime player in the construction of perceptual experience itself” (Clark 2016: 56). Each of the mentioned experiments could be used to support a strong form of theory-ladenness. And by accounting for the results of these experiments in a unifying manner, FEP provides yet more fundamental support for theory-ladenness. The brain constructs generative models and applies them to the actual sensory inputs. Actual sensory inputs are perceived on the basis of the generative model at the top levels of the hierarchy, and evidence/sensory inputs must conform to the hypothesis/theory that is exhibited by the generative models. Therefore, observation is theory-laden. However, closer scrutiny may indicate that FEP does not unconditionally support theory-ladenness.

To provide a reliable understanding of FEP's bearing on the issue of theory-ladenness, it is important to note that top-down influence of prior knowledge is constantly modified through the interface with the bottom-up current of sensory inputs, where the brain forms estimations of the precision of prediction errors and uses these to balance the interaction between predicted and actual inputs. As Hohwy (2013: Ch. 6) argues, although top-down prediction of sensory input bears significantly on perceptual inference, it is impossible to say how extensive the influence of theories actually is (see Hohwy 2017). The implication is important for our discussion of theory-ladenness because it shows that although theory-ladenness is ubiquitous, it is not strong in normal circumstances (when vagueness is not maximal). This is in line with Brewer and Lambert's aforementioned remark.

Hohwy submits that there are indeed genuine instances of dependence of perceptual inference on top-down predictions which form the whole hierarchical structure of prediction error minimisation. However, the observation does not immediately confirm a strong form of theory-ladenness. The important point is that the relation between generative models and sensory inputs is not unilateral, because bottom-up prediction errors, too, affect the brain's predictions via a process that leads to updating generative models. When translated in terms of theory-ladenness, this means that the theory could not completely determine the observation. This adds up to the conclusion that the strong form of theory-ladenness (or cognitive penetration) prevails only under highly uncertain circumstances (Hohwy 2013: 123). Given that the circumstances are never completely certain, perceptual inferences are always to some extent cognitively penetrable. However, uncertainty is not always (or even normally) maximal. Therefore, observation is not always determined by observation. Below, I shall furnish some more details.

Consider the process of the formation of any perceptual inference. The amount of uncertainty is low at earlier stages of processing when the prediction error is hugely suppressed. Hohwy (2017) explicates this situation in terms of the learning rate, where a high learning rate indicates that perceptual inference is determined by new evidence rather than prior beliefs. Under such circumstances active inferences efficiently contribute to updating one's generative models and constrain the penetrability of perceptual inferences – this is the case, for example, in the Müller-Lyer experiment which underlies Fodor's (1984) counterargument (see Hohwy 2013: 127). On the other hand, in cases of overwhelming vagueness and uncertainty, the agents would prefer to put their trust in their theories/hypotheses rather than wavering sensory inputs. This is the case in duck/rabbit and old woman/young woman experiments when the picture itself is too vague to determine the perceptual inferences (Hohwy (2017) accounts for this in terms of deviation from optimal learning rate). This adds up to the conclusion that “beliefs, in the shape of prior expectations, are capable of determining perceptual content in quite profound ways, and that this is especially evident in cases of high uncertainty” (Hohwy 2013: 137). This conclusion, about the relationship between strong cognitive penetrability and highly uncertain circumstances, consolidates Brewer and Lambert's (2001) insight into the correlation between strong and weak theory-ladenness on the one hand and the amount of ambiguity on the other. (It is also in line with the general trajectory of the works by Raftopoulos (2008) and Franklin (2015).) To the extent that scientific evidence from FEP indicates, strong theory-ladenness prevails only under highly uncertain circumstances. Otherwise, theory-ladenness is only weak. In the next section, I will unpack the consequences of the discussion for the realism vs. antirealism debate in the philosophy of science.

6. Moderate cognitive realism

We rifled through parts of the literature on cognitive penetrability under FEP (also see Drayson 2017; Hipólito and Kirchhoff 2019). ‘Cognitive penetrability’ speaks directly to theory-ladenness in the philosophy of science. This provides a key to naturalising notions of theory-ladenness based on recent theories of cognitive neuroscience. The consequences could be rather significant for the realist/antirealist debate in the philosophy of science. If the strong form of theory-ladenness was correct, it would lead to something like strong constructivism or antirealism. If a strong form of theory-ladenness had prevailed, then observation or experimentation wouldn't have been able to confirm, corroborate, or falsify theories, because there were no such things as theory-independent observations. Scientific theories would have been dislodged from nature, and scientific realism would have been in dire straits. However, scientific evidence from recent cognitive psychology only supports the assumption of a modest form of theory-ladenness, given that theory-ladenness prevails only under mounting uncertainty. Because only under ambiguous and uncertain circumstances strong forms of cognitive penetrability arise, science is not completely dislodged from the objective world and there could be relatively independent experiments/observations to establish the objectivity of theories. For, cases of high ambiguity are not ubiquitous. Thus, although strong theory-ladenness can raise its head occasionally, it can be granted that most of the time theory-ladenness is constrained. This understanding of the relationship between theories and reality does not lead to antirealism. Nor does it endorse a full-blooming version of scientific realism (given that theory-ladenness could never be completely uprooted). This leaves room only for a moderate form of cognitive realism. The form of realism that is at issue here is ‘cognitive’ in the sense that it draws on what science reveals about the cognitive relation of theories (which are constructed by

organisms) and the world to establish its epistemic and semantic components. Below, I shall continue this discussion by showing how our scientifically informed enterprise fits in with a naturalistic reading of parts of Kuhn's views on scientific activity.

As I have remarked previously in this paper, Bird's attempt at drawing a connection between PDP and Kuhn's ideas has been based on his (1962: 62–3) reliance on psychology in support of his views on theory-ladenness and incommensurability. This reliance on experimental research provides enough grounds for construing Kuhn's ideas along the lines of naturalism. As I have remarked before, naturalism may lead to a moderate form of realism. While there does not exist an incontrovertible account of the relationship between scientific realism and naturalism, there have been naturalists such as Quine (1960 [2015], 1995) who advocated sophisticated and at the same time robust forms of scientific realism. This does not demonstrate that Kuhn was a scientific realist too, but they might indicate that Kuhn's ideas are not totally incompatible with a gentle form of scientific realism, given that scientific research is accredited enough by him to be conjured in support of his account of theory-ladenness and incommensurability.

I have already mentioned Bird's construal of Kuhn's naturalism, but I have to also draw attention to another creditable restatement of Kuhn's works in terms of PDP and connectionism. This other interpretation is more overtly in line with cognitive realism. Although not being discussed by Bird, Churchland (1989) had paraphrased some of Kuhn's ideas in terms of connectionist models of cognition. Churchland's goal was to replace the stale propositional conception of theories – which specifies theories as language-like entities – with a picture that is informed by breakthroughs in AI and connectionism. Theories are characterised in terms of configurations of activation spaces in artificial neural networks (this move was a part of Churchland's contribution to the cognitive models of science approach, which was also advocated by others (Giere 1992; Nersessian 1999; Thagard and Findlay 2012)). Churchland regards Kuhn's enterprise congenial to his own project, remarking that Kuhn too endeavours to dispense with the propositional conception of theories in favour of a paradigm-based picture, where paradigms are sets of mathematical, conceptual, or instrumental resources, rather than language-like reports. In this context, Churchland argues that Kuhn's notion of indoctrination into a paradigm can be modelled by neural networks, when the networks are trained to represent “a weight configuration that produces some well-structured similarity space whose central hyper-volume locates the prototypical application(s)” (Churchland 1989: 95). In this vein, Churchland provides a viable cognitive representation of some of the central Kuhnian themes such as paradigms, the resistance to a paradigm change, and incommensurability in terms of changes in various values of connection weights at the hidden level of neural networks. And thus, in a work that anticipates Bird's interpretation of Kuhn's naturalism, Churchland explains how Kuhnian notions such as incommensurability could be explicated in terms of perceptual pattern recognition in neural networks.

Both Churchland's and Bird's interpretations seek to update Thomas Kuhn's account by drawing on the scientific literature on neural networks and PDP. Previously in this paper, I have argued that FEP amounts to an advanced form of top-down processing in hierarchically organized PDP. I also referred to Orlandi's work to argue that predictive processing under FEP is in line with insights from Gestalt psychology as well as connectionism. To be more precise, although FEP retains some of the basic insights into distributed structural representations of connectionist types (Friston and Stephan 2007: 420), it makes a great improvement on traditional connectionism in the following sense. Instead of presenting “a story about how an internal reconstruction of the external world is constructed in the brain” (as a traditional connectionist account would do), FEP underlies an ontology of “the set of semantic structures (a large part of

which are learned through experience from immersion in specific contexts) that a living creature uses, implicitly or explicitly, to parse and order the flow of its sensory states” (Ramstead *et al.* 2020: 18). This makes FEP an ideal framework for garnering some evidence that could help us to adjudicate the possibility and scope of theory-ladenness. Our enterprise to use FEP to address the issue of theory-ladenness is not wide of the mark, when compared against Kuhn’s own scientifically informed account of theory-ladenness. Perhaps it is also worth mentioning that FEP has been adopted into a cognitive account of scientific representation recently. I shall unpack this last remark immediately.

Parts of Churchland’s insights have been recently adopted into a form of the cognitive realist theory of scientific representation (Beni 2019a), which aims to account for scientific representation in terms of the (individual or collective) cognitive engagement of scientists with their environment, where the cognitive engagement is characterised in terms of minimisation of prediction error under FEP. It is worth mentioning that aside from the reliance on Churchland’s work, a project that I have developed a while ago, namely cognitive structural realism (CSR),⁵ is also heavily influenced by Giere’s (1988, 1992, 1994) pioneering work on cognitive models of science (for discussion see Beni 2019a: 27–8; 41–2; 101–2 among other places). Giere takes a cognitive – rather than an exclusively formal or methodological – approach to the nature of scientific theories, assuming that “the cognitive sciences have reached a sufficient state of maturity that they can now provide a valuable resource for philosophy of science” (Giere 1992: XV). This cognitive take on scientific rationality is in line with Kuhn’s rich and sophisticated account of science. For, cognitive models of science – which are grounded in cognitive capacities of actual agents – can register and represent historical, social, and political aspects of science, and as such, they do not need to be understood in terms of purely logical-methodological tools (also see Giere 2013). This means that the cognitive model approach could as well capture a part of the subtle historical-social tendency of Kuhn’s approach. Indeed, Giere develops his proposed meta-theoretical cognitive framework with the explicit aim of bridging the gap between the logical empiricist understanding of theories and Kuhn’s historical approach to science (Giere 1992: XV; Beni 2019a: 28).⁶

Be that as it may, cognitive realism, as being developed in the wake of Giere’s and Churchland’s pioneering works, indicates that scientific theories are embodied in the brain of actual scientists (who form scientific communities), and scientific representations get realized through active inferences and dynamic free energy minimisation relation between cognitive agents and their environment. By minimising their prediction errors scientists endeavour to enhance the precision of their representations of the causal structure of the world. Because scientists are organisms, their collaborative enterprise to construct scientific representations could be basically explicated in terms of their organic-natural capacity to minimise the discrepancy between their (shared) generative models and the causal structure of the world.

Let us recap. By minimising their free energy or prediction errors, organisms succeed in enhancing the precision of their representations of the structure of the world. Some organisms are even capable of forming reliable estimation of the precision of their representations of the world based on the temporal depth of the generative models

⁵Aside from approaches that are discussed here, as its name indicates, the version of cognitive realism that I have developed before is also related to the structural realist programme in the contemporary philosophy of science, but we do not need to discuss that point in the present paper.

⁶Another interesting connection between Giere and Kuhn’s work is that Giere construes works of at least the later stage of Kuhn’s (2000) work along the lines of perspectival realism (see Giere 2013).

that are in use in their free energy minimization mechanisms (Beni 2019a: Ch. 7). My aforementioned account of scientific representations (in CSR) is as modest as its naturalist tendency is radical. It is dubbed *cognitive realism* mainly because it draws on theoretical achievements in cognitive sciences to establish the viability of its defence of the epistemic and semantic components of the realist thesis (instead of orthodox arguments such as the no miracles argument). The present paper does not aim to substantiate the plausibility of this recent form of cognitive realism (see Beni 2017, 2019c, 2020 for elaboration and defence). It simply aims to show that the moderate interpretation of theory-ladenness (along the lines of cognitive penetrability under FEP) is not only compatible with a naturalist reading of Kuhn's ideas but also with a moderate form of cognitive realism (which might as well be compatible with a naturalized version of Kuhn's ideas in its own right). This remark puts our engagement with the issues of theory-ladenness and cognitive penetrability in the context of the realist-antirealist debate in the philosophy of science.

7. Conclusion

The paper engaged with the issue of theory-ladenness within the philosophy of science. Drawing on the works of Bird, Churchland, Fodor, Giere, Kitcher, and Thomas Kuhn, amongst others, the paper argued that it is commendable to build upon theoretical and experimental psychology to get a well-informed perspective on theory-ladenness. Although it has been taken for granted that observation/experimentation is theory-laden, the scope of theory-ladenness has not been clearly drawn before. The paper approaches the issue of theory-ladenness by equating it with the cognitive penetrability of perceptual inferences before drawing on the theoretical resources of the minimisation of free energy (FEP) to provide a viable account of the scope of theory-ladenness. Grounds from cognitive neuroscience indicate that theory-ladenness (operationalised in terms of cognitive penetrability) is not normally strong, and strong theory-ladenness prevails only under exceedingly uncertain circumstances. Then the paper proceeded to unfold the ramifications of the point for the realist vs. antirealist debate. I argued that the weak form of cognitive penetrability/theory-ladenness (which receives support from scientific evidence) is compatible with modest forms of cognitive realism (that have been pioneered by Giere and restored in the cognitive structural realism project) rather than with strong versions of either realism or antirealism. I also argued that the scientifically informed view of theory-ladenness, as well as the account of its affinity with cognitive realism, is in line with a naturalist reading of Kuhn's perspective on science.⁷

References

- Beni M.D. (2017). 'Reconstructing the Upward Path to Structural Realism.' *European Journal for Philosophy of Science* 7, 393–409.
- Beni M.D. (2018). 'The Reward of Unification: A Realist Reading of the Predictive Processing Theory.' *New Ideas in Psychology* 48, 21–6.
- Beni M.D. (2019a). *Cognitive Structural Realism: A Radical Solution to the Problem of Scientific Representation*. Cham: Springer.
- Beni M.D. (2019b). 'Conjuring Cognitive Structures: Towards a Unified Model of Cognition.' In A. Nepomuceno-Fernández, L. Magnani, F. Salguero-Lamillar, C. Barés-Gómez and M. Fontaine (eds), *Model-Based Reasoning in Science and Technology. MBR 2018*, pp. 153–72. Cham: Springer.
- Beni M.D. (2019c). 'The Benacerraf Problem as a Challenge for Ontic Structural Realism.' *Philosophia Mathematica*. <https://doi.org/10.1093/phimat/nkz022>.

⁷I sincerely thank the anonymous reviewers of this journal for their insightful comments.

- Beni M.D. (2020). 'Reconstructing Probabilistic Realism: Re-enacting Syntactical Structures.' *Journal for General Philosophy of Science* 51(2). <https://doi.org/10.1007/s10838-018-9426-z>.
- Beni M.D. (2021). 'A Critical Analysis of Markovian Monism.' *Synthese*. <https://doi.org/10.1007/s11229-021-03075-x>.
- Bird A. (2005). 'Naturalizing Kuhn.' *Proceedings of the Aristotelian Society* 105, 99–117.
- Bird A. (2012). 'Kuhn, Naturalism, and the Social Study of Science.' In V. Kindi and T. Arabatzis (eds), *Kuhn's The Structure of Scientific Revolutions Revisited*. New York, NY: Routledge.
- Brewer W.F. and Lambert B.L. (2001). 'The Theory-Ladenness of Observation and the Theory-Ladenness of the Rest of the Scientific Process.' *Philosophy of Science* 68(S3), S176–86.
- Bruineberg J., Dolega K., Dewhurst J. and Baltieri M. (2020). *The Emperor's New Markov Blankets*. PhilSci Archive. <http://philsci-archive.pitt.edu/18467/>.
- Bruner J. and Postman L. (1949). 'On the Perception of Incongruity: A Paradigm.' *Journal of Personality* 18, 206–23.
- Carnap R. (1936). 'Testability and Meaning.' *Philosophy of Science* 3, 419–71.
- Carnap R. (1956). 'The Methodological Character of Theoretical Concepts.' *Minnesota Studies in the Philosophy of Science* 1, 38–76.
- Chakravartty A. (2007). *A Metaphysics for Scientific Realism: Knowing the Unobservable*. Cambridge: Cambridge University Press.
- Churchland P.M. (1979). *Scientific Realism and the Plasticity of Mind*. Cambridge: Cambridge University Press.
- Churchland P.M. (1989). 'On the Nature of Theories: A Neurocomputational Perspective.' In C.W. Savage (ed.), *Minnesota Studies in the Philosophy of Science, Volume 14. Scientific Theories*, pp. 59–101. Minneapolis, MN: University of Minnesota Press.
- Clark A. (2016). *Surfing Uncertainty*. Oxford: Oxford University Press.
- Colombo M. and Hartmann S. (2015). 'Bayesian Cognitive Science, Unification, and Explanation.' *British Journal for the Philosophy of Science* 68(2), axv036.
- Colombo M. and Wright C. (2018). 'First Principles in the Life Sciences: The Free-Energy Principle, Organicism, and Mechanism.' *Synthese*. <https://doi.org/10.1007/s11229-018-01932-w>.
- Drayson Z. (2017). 'Modularity and the Predictive Mind.' In W. Wiese and T. Metzinger (eds), *Philosophy and Predictive Processing*, Vol. 12. Frankfurt am Main: MIND Group. Johannes Gutenberg Universität Mainz.
- Fodor J. (1984). 'Observation Reconsidered.' *Philosophy of Science* 51(1), 23–43.
- Franklin A. (2015). 'The Theory-Ladenness of Experiment.' *Journal for General Philosophy of Science* 46(1), 155–66.
- French S. (2014). *The Structure of the World: Metaphysics and Representation*. Oxford: Oxford University Press.
- Friedman M. (2011). 'Carnap on Theoretical Terms: Structuralism Without Metaphysics.' *Synthese* 180, 249–263.
- Friston K.J. (1994). 'Functional and Effective Connectivity in Neuroimaging: A Synthesis.' *Human Brain Mapping* 2(1–2), 56–78.
- Friston K.J. (2010). 'The Free-energy Principle: A Unified Brain Theory?' *Nature Reviews. Neuroscience* 11(2), 127–38.
- Friston K.J. (2012). 'A Free Energy Principle for Biological Systems.' *Entropy (Basel, Switzerland)* 14(11), 2100–21.
- Friston K.J. (2013). 'Active Inference and Free Energy.' *Behavioral and Brain Sciences* 36, 212–13.
- Friston K.J. (2019). 'A Free Energy Principle for a Particular Physics.' *ArXiv*. <http://arxiv.org/abs/1906.10184>.
- Friston K.J. and Stephan K.E. (2007). 'Free-energy and the Brain.' *Synthese* 159(3), 417–458.
- Friston K.J., Kilner J. and Harrison L. (2006). 'A Free Energy Principle for the Brain.' *Journal of Physiology Paris* 100(1–3), 70–87.
- Friston K.J., Wiese W. and Hobson J.A. (2020). 'Sentience and the Origins of Consciousness: From Cartesian Duality to Markovian Monism.' *Entropy* 22. <https://doi.org/10.3390/E22050516>.
- Giere R.N. (1988). *Explaining Science: A Cognitive Approach*. Chicago, IL: University of Chicago Press.
- Giere R.N. (1992). 'Cognitive Models of Science.' *Minnesota Studies in the Philosophy of Science* XV (4). <https://doi.org/10.1177/030631293023004005>.
- Giere R.N. (1994). 'The Cognitive Structure of Scientific Theories.' *Philosophy of Science* 61(2), 276–96.
- Giere R.N. (2013). 'Kuhn as Perspectival Realist.' *Topoi* 32(1), 53–7.

- Hipólito I. and Kirchhoff M.D.** (2019). 'The Predictive Brain: A Modular View of Brain and Cognitive Function?' *Preprints* **2019**, 2019110111. doi: 10.20944/preprints201911.0111.v1.
- Hipólito I., Ramstead M.J.D., Convertino L., Bhat A., Friston K. and Parr T.** (2021). 'Markov Blankets in the Brain.' *Neuroscience & Biobehavioral Reviews* **125**, 88–97.
- Hohwy J.** (2013). *The Predictive Mind*. Oxford: Oxford University Press.
- Hohwy J.** (2017). 'Priors in Perception: Top-down Modulation, Bayesian Perceptual Learning Rate, and Prediction Error Minimization.' *Consciousness and Cognition* **47**, 75–85.
- Hoyningen-Huene P.** (1993). *Reconstructing Scientific Revolutions: Thomas S. Kuhn's Philosophy of Science*. Translated by A.T. Levine. Chicago, IL: University of Chicago Press.
- Karaca K.** (2013). 'The Strong and Weak Senses of Theory-Ladenness of Experimentation: Theory-Driven versus Exploratory Experiments in the History of High-Energy Particle Physics.' *Science in Context* **26** (1), 93–136.
- Kiefer A.B.** (2020). 'Psychophysical Identity and Free Energy.' *Journal of the Royal Society Interface*. <https://doi.org/10.1098/rsif.2020.0370>.
- Kirchhoff M.D., Parr T., Palacios E., Friston K.J. and Kiverstein J.** (2018). 'The Markov Blankets of Life: Autonomy, Active Inference and the Free Energy Principle.' *Journal of the Royal Society* **15**(138), 20170792.
- Kitcher P.** (2001). 'Real Realism: The Galilean Strategy.' *The Philosophical Review* **110**(2), 151.
- Köhler W.** (2007 [1920]). 'Physical Gestalten.' In *A Source Book of Gestalt Psychology*, pp. 17–54. London: Kegan Paul.
- Kuhn T.S.** (1962). *The Structure of Scientific Revolutions*. Chicago, IL: Chicago University Press.
- Kuhn T.S.** (1977). *The Essential Tension: Selected Studies in Scientific Tradition and Change*. Chicago, IL: University of Chicago Press.
- Kuhn T.S.** (1992). 'The Trouble with the Historical Philosophy of Science.' In *Robert and Maurine Rothschild Distinguished Lecture, 19 November 1991; Occasional Publication of the Department of the History of Science*. Cambridge, MA: Harvard University Press.
- Kuhn T.S.** (2000). *The Road Since Structure*. Chicago, IL: University of Chicago Press.
- Latour B. and Woolgar S.** (1979). *Laboratory Life: The Construction of Scientific Fact*, 2nd edition. Princeton, NJ: Princeton University Press.
- Laudan L.** (1981). 'A Confutation of Convergent Realism.' *Philosophy of Science* **48**(1), 19–49.
- Limanowski J. and Friston K.** (2020). 'Active Inference under Visuo-proprioceptive Conflict: Simulation and Empirical Results.' *Scientific Reports*. <https://doi.org/10.1038/s41598-020-61097-w>.
- Mackie J.L.** (1977). 'Dispositions, Grounds, and Causes.' *Synthese* **34**, 99–107.
- Maxwell G.** (1962). 'The Ontological Status of Theoretical Entities.' In H. Feigl and G. Maxwell (eds), *Scientific Explanation, Space, and Time: Minnesota Studies in the Philosophy of Science*, pp. 181–92. Minneapolis, MN: University of Minnesota Press.
- Maxwell G.** (1971). 'Structural Realism and the Meaning of Theoretical Terms.' *Minnesota Studies in the Philosophy of Science* **4**, 181–92.
- Mumford S.** (2006). *Laws in Nature*. London: Routledge.
- Nersessian N.J.** (1999). 'Model-based Reasoning in Conceptual Change.' In L. Magnani, N.J. Nersessian and P. Thagard (eds), *Model-Based Reasoning in Scientific Discovery*, pp. 5–22. New York, NY: Springer US.
- Orlandi N.** (2016). 'Bayesian Perception is Ecological Perception.' *Philosophical Topics* **44**(2), 327–51.
- Parr T., Da Costa L. and Friston K.J.** (2020). 'Markov Blankets, Information Geometry and Stochastic Thermodynamics.' *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **378**(2164), 20190159.
- Psillos S.** (2000a). 'Carnap, the Ramsey-Sentence and Realistic Empiricism.' *Erkenntnis* **52**(2), 253–79.
- Psillos S.** (2000b). 'Rudolf Carnap's "Theoretical Concepts in Science".' *Studies in History and Philosophy of Science Part A* **31**(1), 151–72.
- Quine W. van O.** (1960 [2015]). *Word and Object*. Cambridge, MA: MIT Press.
- Quine W. van O.** (1995). *From Stimulus to Science*. Cambridge, MA: Harvard University Press.
- Raftopoulos A.** (2008). 'Perceptual Systems and Realism.' *Synthese* **164**(1), 61–91.
- Ramstead M.J.D., Badcock P.B. and Friston K.J.** (2017). 'Answering Schrödinger's Question: A Free-energy Formulation.' *Physics of Life Reviews*. <https://doi.org/10.1016/j.PLREV.2017.09.001>
- Ramstead M.J.D., Friston K.J. and Hipólito I.** (2020). 'Is the Free-Energy Principle a Formal Theory of Semantics? From Variational Density Dynamics to Neural and Phenotypic Representations.' *Entropy* **22** (8). <https://doi.org/10.3390/e22080889>.

- Remez R.E. and Rubin P.E. (1990). 'On the Perception of Speech from Time-varying Acoustic Information: Contributions of Amplitude Variation.' *Perception & Psychophysics* 48(4), 313–25.
- Seth A.K. and Hohwy J. (2020). 'Predictive Processing as an Empirical Theory for Consciousness Science.' *Cognitive Neuroscience*. <https://doi.org/10.1080/17588928.2020.1838467>.
- Thagard P. and Findlay S. (2012). *The Cognitive Science of Science: Explanation, Discovery, and Conceptual Change*. Cambridge, MA: MIT Press.
- Van de Cruys S. and Wagemans J. (2011). 'Gestalts as Predictions – Some Reflections and an Application to Art. – PsycNET.' *Gestalt Theory* 33(3–4). <https://psycnet.apa.org/record/2012-02844-005>.
- van Es T. and Hipólito I. (2020). 'Free-Energy Principle, Computationalism and Realism: A Tragedy.' *PhilSci Archive*. <http://philsci-archive.pitt.edu/id/eprint/18497>.
- van Fraassen B.C. (1980). *The Scientific Image*. Oxford: Oxford University Press.
- Wagemans J., Elder J.H., Kubovy M., Palmer S.E., Peterson M.A., Singh M. and von der Heydt R. (2012). 'A Century of Gestalt Psychology in Visual Perception: I. Perceptual Grouping and Figure-ground Organization.' *Psychological Bulletin* 138(6), 1172–217.

Majid D. Beni is presently affiliated with the Department of Philosophy of Middle East Technical University, Ankara. He works at the intersection of the philosophy of science and the philosophy of neuroscience. Amongst other things, he has developed a structural realist theory of selfhood and a cognitive version of structural realism. Presently, he is expanding his cognitive structural realism project to explicate the social aspects of scientific practice, so as to enrich the orthodox version of structural realism with resources from social ontology and social studies of science as well as cognitive psychology and neuroscience.