

Idealist implications of contemporary science

[3rd revision July 2023]

Abstract

Recent developments in contemporary natural science (including the evolutionary study of perception, cognitive science, and interpretations of quantum physics) incorporate central idealist positions relating to the nature of representation, the role our minds play in structuring our experience of the world, and the properties of the world behind our representations. This paper first describes what these positions are, and how they are introduced in the relevant theories in terms of precisely formulated scientific analogues. I subsequently consider how this way of looking at philosophical idealism through selected parts of contemporary science can help us to pursue new ways of developing key idealist questions in a way that is integrated with a naturalistically supported endeavour to understand central features of reality.

0. Introduction

Idealism has steadily declined in popularity over the last century. Not only do we find hardly any contemporary philosophical defenses of idealism,¹ the fact that a theory has idealistic consequences is sometimes taken to be a sufficient argument for its absurdity.² This is somewhat surprising if we take into account that idealist theories of various forms played a major part in the philosophical history of the world, including intellectual traditions as diverse as Berkeley's idealism, the transcendental idealism of Kant and Schopenhauer, Advaita Vedānta, Yogācāra Buddhism, or Kāśmir Śaivism, to name but a few.³ Still, one might think, despite the venerable ancestry of idealisms of various stripes, theoretical progress means that approaches shown to be conceptually and empirically unsatisfactory will find few defenders, and this is exactly what has been happening to idealism. It is then particularly puzzling to see a current comeback of idealist concepts, carried along not by philosophy, but by developments in contemporary science.⁴ Moreover, this idealistic resurgence does not seem to be confined to a single field or researcher, but arises in areas as distinct as the theory of perception, neuroscience, or the interpretation of quantum physics.

¹ With a few notable exceptions: Foster 1982, 2008, von Kutschera 2006, Sprigge 1983, Taber 2020. See also Farris/Göcke 2022, parts 3-5.

² Anderson 2017: 5-6: "I most certainly do *not* want to recapitulate the realism vs. idealism and epistemic externalism vs. internalism debates [...] I'll simply assert without argument that [...] any epistemology that implies internalism and its attendant skepticism has made a mistake *somewhere*." Simons (2021: 76) includes idealism in his list of "metaphysical follies", views he considers to be "bizarre, extreme and unbelievable".

³ Dunham et al 2011 and Goldschmidt/Pearce 2017 provide good surveys of the diversity of idealist traditions.

⁴ For some discussion of global advance of science in the late 19th and early 20th century and the contemporaneous rise of philosophical idealism in Europe, India, and China see Garfield/Bushan 2017: 178-216, Makeham 2014.

This paper will look at three specific examples: the Interface Theory of Perception, Prediction Error Minimization Theory, and Quantum Bayesianism. While these theories are generally not put forward as explicit scientific or empirical support for idealism, their core ideas are sufficiently intertwined with idealist concepts that one cannot fully endorse these theories without at the same time taking their idealist frameworks seriously. I will, in the course of this essay, not be able to say much to defend these theories, though they strike me all to be of considerable empirical and theoretical interest. One conclusion of this essay will be hypothetical: *if we want to sign up to these theories, then we should, at the same time, embrace the idealist positions that play a central role in their formulation.* My main point, however, is something else. I set out to show how idealist concepts play a fundamental part in the formulation of the three theories discussed and are formulated in terms of scientifically precise analogues. This, I argue, opens up routes to develop the discussion of idealism in a contemporary, scientifically literate manner.

In the context of this discussion I do not want to describe idealism simply as the view that at the fundamental level everything is mental, since the conciseness of this characterisation does not compensate for its vagueness (what exactly is meant by 'mental' here?) or its excessive narrowness (what about non-foundationalist forms of idealism?). A better

approach is to understand idealism as a family of views aligned with at least the following three principles:⁵

1. *Representation*: We do not encounter the world in a direct manner, but through representations. Our access to the world takes place through perception, reasoning, and introspection. Perception, reasoning, and introspection operate on representational entities, rather than on the represented entities themselves.

2. *Formation*: The way the world appears to us is extensively and essentially shaped by structural features of the human mind. As everything appears to have a greenish tint for someone wearing green-coloured lenses, our experience of the world is coloured by the structure of the representations through which we perceive it.

3. *Non-correspondence*: This principle rejects the assumption of a world of mind-independent represented objects behind our representations that correspond, at least in broad structural outlines, to the entities featuring in contemporary physical and mathematical theories. As such there are no spatial, temporal, causal, or mathematical structures 'out there' which are isomorphic

⁵ My claim is not that every position in the history of philosophy ever described as idealism endorses all of these principles, but rather that they subsume a large enough subset of the family to provide sufficient content for the discussion of idealism presented below.

with the structure of our experience, or with the structure of theoretically systematising this experience.⁶

Note the essential role of the principle of non-correspondence in this list. Even if we assume that our access to the world is largely indirect, and strongly shaped by our mind as the tool making this indirect access possible, we might still think we are able to isolate the basic aspects of our experience that correspond to basic physical or mathematical aspects of the represented world beyond our experience. This view, however, would be a mitigated form of realism, not a form of idealism.

The purely negative formulation of this principle is also responsible for the variety of different views that can be subsumed under this conception of idealism, diverging in how they spell out references to 'the world beyond our representations'. Two obvious ways of understanding this phrase might be called *noumenalism* and *limitationism*. For the noumenalist, such a world beyond our representations exists, though we cannot say anything else about it. Since all our conceptual resources for describing entities in terms of space and time, causation, probability, and so forth are restricted to the realm of representations, we cannot use these resources to speak about the world beyond the representations, for example by saying that it

⁶ I believe the three principles can be brought into sharper focus by considering their direct opposites, many of which are widely accepted philosophical positions. *Representation* contradicts various forms of direct or naïve realism which assumes that our perception puts us into a direct relation with entities 'out there' in the world, without the need for a representational intermediary (for the popular disjunctivist variety of this view see Soteriou 2016). An antithesis of *formation* is the common-sense realism defended, for example, by Michael Devitt ("Tokens of most current observable common-sense and scientific physical types objectively exist independently of the mental.", 1997: 24), while *non-correspondence* is directly opposed to epistemic realism, "the reigning orthodoxy among philosophers for almost a generation", the view that "especially in the 'mature' and well-developed parts of the physical sciences, scientists have come very close to discerning the way the world *really* is." (Laudan 1997: 138).

is the cause of what appears to us. Furthermore, we do not have any other conceptual tools at our disposal for describing this world. The limitationist, on the other hand, holds the even more restrictive view that since the only thing we can say anything about are representations any attempt to speak even about the existence of the world behind the representations is impossible and results in nonsense. All our talk needs to be restricted to the level of representation.

1. Interface theory of perception

The interface theory is a theory of perception developed by Donald Hoffman and his collaborators over the last decades.⁷ It is based on the largely uncontroversial idea that our perceptual abilities are the product of evolutionary development. They evolved because they provided us with increased fitness, increasing our chances to pass on our genes by enhancing our abilities to detect food, competitors, and potential mates. The evolutionary origins of perceptual abilities are usually not regarded as in conflict with their potential to accurately represent the world. On the contrary, it seems that our ability to detect food can only equip us with added fitness if it manages to represent the presence of food in the world accurately.⁸ Nevertheless, Hoffman argues, perceptual strategies aiming

⁷ See Hoffman et al 2015a, b. Hoffman 2019 presents a popularized account of the theory. Hoffman also sets out to build a comprehensive theory of consciousness (“conscious realism”, the view that “the objective world [...] consists entirely of conscious agents”, 2008: 103) on the basis of the interface theory. These further theoretical developments are not part of our present discussion.

⁸ See Simons 2017: 32.

for evolutionary fitness, and perceptual strategies aiming for faithful representation of the world pull in different directions. The central support of this claim is a result in evolutionary game theory called the fitness-beats-truth (FBT) theorem.⁹ The FBT theorem says that if perceptual strategies increasing evolutionary fitness are pitched against strategies aiming to accurately represent the environment ('truth'), the former regularly outperform the latter and will drive them to extinction, at least in cases where fitness payoff and resource quantity do not vary monotonically.¹⁰

The consequence the interface theory draws from the FBT theorem is that because our perceptual abilities have been picked by an evolutionary process, they will have been selected for increasing evolutionary fitness, not accurate representation. The probability that the perceptual abilities we ended up with perform both roles by accident (having been selected for one feature, but by coincidence also displaying another) is very low.¹¹ This applies independently of whether we use our perceptual abilities to navigate a busy road or to construct a theory of the fundamental constituents of matter.¹²

⁹ Prakash 2021.

¹⁰ Hoffman et al. 2015a: 1486, Prakash 2021: 326-327. This is frequently the case. Not enough water is as bad for an organism as too much; that some intake of water increases fitness does not imply that any intake will.

¹¹ Hoffman 2008: 112 and Prakash 2020 for the relevant technical details behind this claim. We can show that the more experiences a perceiver is able to distinguish the greater the probability that perceptual strategies aiming at fitness the perceiver might adopt will outperform strategies aiming at accurate representation.

Prakash et al 2020 examines a set of structures of our representation of the world (total orders, permutation groups, cyclic groups, and measurable spaces), arguing that if these structures were also instantiated in the world, it would be exceedingly unlikely that evolutionarily successful fitness functions ever mirror them.

¹² Hoffman 2015b notes that "the fundamental dynamical properties of physics—including position, momentum, and spin—do not describe reality as it is, but are instead products of—that is, creations of—the measurement process " (1554) and that "physics is almost surely not causally complete". (1571).

As a consequence, we should conceive of our perceptions as an evolutionary developed interface that enables us to successfully navigate the world in a way that maximises the likelihood of passing on our genes. Successful navigation of the world, however, does not imply accurate representation, in fact the FBT theorem shows that it is exceedingly unlikely that our perceptual interface shares a significant amount of structure with the world it represents.¹³ As such, the interface theory argues, we can conceive of this interface along the lines of a computer's graphical user interface: useful for operating the computer, but highly misleading if we understand the structure of the interface as indicative of the structure of the software and hardware generating the interface.

1a. Representation

The representationalism characterizing the interface theory is evident even after this short description. According to this theory, we never interact with the world itself, but only with an interface of evolutionary developed perceptual representations. The interface theory of perception moves the majority of things we usually consider to be part of the non-mental external world (fundamental particles, neurons, medium-sized dry goods, time, space) into the perceptual interface and treats them as constituents of this interface.¹⁴ The interface theory's representationalism postulates a boundary with representations on one side, and whatever it is these representations represent on the other.

¹³ Such a sharing of structure is assumed, for example, by O'Brien/Opie 2004: 15, who argue that "it is a relation of structural resemblance between mental representing vehicles and their objects that disposes cognitive subjects to behave appropriately towards the latter."

¹⁴ Hoffman et al 2015: 1484, 1501, 1503.

1b. Formation

Interface theory also relies crucially on the principle of formation. The representations that make up the interface have the form they have because they have been shaped by evolutionary selection. The way the world appears to us as a collection of discrete entities spread out in space and time does not reflect an underlying ontology of spatio-temporally related individuals but derives from structural features of the human mind, features which are, in turn, the result of selection pressure favouring representations that succeed best in enhancing evolutionary fitness.¹⁵ The reason why the world appears in three-dimensional space to us, for example, the proponents of the interface theory argue, is because structuring representations in this way constitutes an error-correcting code.¹⁶ Representing information about fitness ‘spread out’ in space allows for redundancy, and hence for the detection and correction of transmission errors. According to this view, our spatial experience is not a reflection of a three-dimensional spatial world of material objects external to the observer, but the result of an evolved procedure for presenting information about fitness to the organism in a reliable way.

1c. Non-correspondence

¹⁵ Hoffman et al 2015: 1496.

¹⁶ Fields et al 2017: 272-273. ‘Errors’ are here understood not as misrepresentations of the environment, but as actions that decrease fitness (288). See also Prakash 2020: 121 which discusses a formal development of the idea that “a conscious agent can consistently see geometric and probabilistic structures of space that are not necessarily in the world *per se* but are properties of the conscious agent itself”.

The form of the non-correspondence principle adopted by the interface theory appears to be best described as a form of noumenalism, arguing that we cannot say anything about the world behind the evolutionary generated enterprise apart from asserting it exists. This, I believe, also explains the rather curious claim by its proponents that the interface theory does not provide an argument for idealism.¹⁷ For a theory suggesting that we “cannot assume that physical objects have genuine causal powers, nor that space-time is fundamental, since objects and space-time are simply species-specific perceptual adaptations”¹⁸ this might appear somewhat disingenuous. Their claim is best construed, I believe, as aiming for metaphysical neutrality. The interface theory’s key claim, based on the FBT theorem, is to deny the plausibility of a structural correspondence between our representations and the world represented. As such it presents a purely negative thesis that leaves it wholly open what ‘the world behind the interface’ might amount to. It is essentially unknowable, and the interface theory restricts itself to transcendental claims about what structure would have to be present in the world *given that we can do science at all*. Of course it might be the case that not even this obtains, but if we are not even willing to entertain the conjecture that empirical investigations can yield knowledge we are probably not interested in the discussion surrounding the interface theory in the first place. It turns out that the structural assumptions the interface theory makes about the world behind the representation are exceedingly minimal. Its proponents sometimes suggest that both the set of our

¹⁷ Hoffman et al 2015: 1490.

¹⁸ Hoffman et al 2104: 20.

experiences and the set of world-states have a specific set-theoretic structure, that is, that they form a measurable space.¹⁹ What this amounts to is essentially to say that for every set there is its complement, and for each finite collection of sets there is a set containing all their members. This is required if we want to speak about the probabilities of world-states and experiences,²⁰ for talk about probabilities necessitates talk about events not happening (complementation) and events happening together (union). If these resources are not available, talk about the probability of some event happening cannot be properly formulated. Elsewhere, however, they suggest an even more parsimonious picture. According to this we only have to assume that the *set of experiences* forms a measurable space, the required structure of the world-states is then induced via the perceptual mapping. In this way we then “need postulate no *a priori* structure of any kind on [the set of world-states] *W*.”²¹

For the interface theory the nature of the world behind the representational interface is thus radically and inescapably underdetermined. Like the Kantian *noumenon*, we still have to assume that it exists (for the interface theory requires our experiential states to be mapped to *something* they represent), but there is nothing we otherwise can or ever could say about it. Evidently it is not implied that the world behind the interface is mental in nature — hence the claim that the interface theory cannot be understood as an argument for idealism.²² This

¹⁹ Hoffman et al 2015: 1483.

²⁰ Hoffman 2015b: 1565.

²¹ Hoffman 2015b: 1563.

²² Despite the neutrality of the interface theory’s conception of the world behind the representations its proponents use it as a basis on which to build an idealist metaphysics (‘conscious agent theory’), see Hoffman et al 2015: 1502). For more on conscious agent theory see Hoffman 2008, Hoffman/Prakash 2014.

position is not in tension with the view developed in the present paper (that the interface theory, like the other scientific approaches discussed, is fundamentally intertwined with idealist concepts), since the conception of idealism deployed here is not one entailing that ‘everything is mental in nature’.

2. Prediction-error minimization theory

Prediction-error minimization theory sets out to provide a comprehensive theory of perception and cognition.²³ To discuss which, if any, idealist conclusions prediction error-minimization theory might support we will primarily consider two aspects: the idea of prediction-error minimization itself, and the notion of a Markov blanket separating the internal states of the perceiver from its environment.

2a. Prediction-error minimization

The central idea behind prediction-error minimization theory is that the brain’s activity is to be understood first and foremost as concerned with minimizing prediction error. In order to achieve this the brain creates a model of the external and internal world, using information coming in through the various sensory channels (including introspective information) in order to predict future incoming information. The discrepancy between what the model predicts and the actual input received is the prediction error. In order to minimize the prediction error in the future the brain can

²³ Helpful surveys of the prediction-error minimization theory are provided by Hohwy 2013, 2020, as well as in the collection of papers available at <https://predictive-mind.net>.

then either change the model (as one might, for example, adapt one's beliefs so as to represent the environment as aptly as possible), or it might change the state of the organism. This latter possibility links prediction-error minimization with the explanation of action: raising my arm is my brain's way of minimizing the discrepancy between modelling my arm as being presently up, and present incoming information reporting my arm as being down.

2b. Markov blankets

A Markov blanket is a specific network of nodes in which we can move from some nodes to other nodes with a specific transition probability.²⁴ If we consider the nodes to be particular states, a transition probability of 0.6 between state *a* and state *b* means that there is a 60% conditional probability of state *b* obtaining, given that state *a* obtains. An obvious example of this can be found in weather forecasting. A 0.6 transition probability between sunny weather on day *d* and rain on day *d*+1 means that there is a 60% probability of it raining tomorrow if it is sunny today. Transition probabilities are also used in the analysis of medical symptoms. If a patient has a cough, there is a certain probability that he also has a raised temperature, which is the transition probability from the 'cough'-state to the 'raised temperature'-state.

If some set of nodes forms a Markov blanket around a given node *x*, all the information necessary to determine the probability of *x* obtaining is located in the probabilities of *x*'s 'direct neighbours' (more precisely: in its

²⁴ For more discussion of Markov blankets see Westerhoff 2020: 37-40, Menary/Gillett 2021, Bruineberg 2022.

parents, children, and children's other parents) that form part of the Markov blanket. There is no necessity to go further afield to states separated from x via long chains of intermediate states. As such the Markov blanket constitutes a mathematical model of a boundary such that if you want to know what is going on beyond the boundary you simply have to look at what is happening in the boundary itself. It thereby provides a useful abstract way of conceiving of the epistemologist's veil of perception separating the 'in here' from the 'out there'. Some researchers equate the boundary between the brain and the rest of the organism with a Markov blanket, arguing that the brain receives sensory input through "exteroceptive, proprioceptive, and interoceptive receptors" and delivers, as an output "proprioceptive predictions [...] mainly in the spinal cord."²⁵

2c. Representation

As such, the notion of the Markov blanket locates the prediction-error minimization framework firmly in a representationalist setting in which all we ever directly²⁶ interact with are parts of the brain-generated representational interface.²⁷ The Markov blanket constitutes an epistemic boundary between the representations 'in here' and whatever is represented 'out there'.

Note, however, that this *environmental seclusion* postulate,²⁸ claiming that we do not have direct access to the environment, but infer it as the hidden causes of our perceptions is by no means accepted by all

²⁵ Hohwy 2016: 276.

²⁶ For further discussion of the notion of 'directness' relevant in this context see Snowdown 1992, Westerhoff 2020: 22-32.

²⁷ Hohwy 2016: 283.

²⁸ Wiese/Metzinger 2007: 1.

proponents of predictive processing.²⁹ Some oppose it by arguing that the Markov blanket is not a fixed boundary, but can be conceived as expanded or contracted depending on context.³⁰ In a case where an agent is, for example, crucially reliant on a notebook in his cognitive operations, the defender of enactive cognition would consider the notebook to be part of his mind, and hence locate notebook is inside the Markov blanket. In other contexts, the same notebook might be part of the environment beyond the Markov blanket and thus a member of the set of ‘hidden causes’. The natural response for the defender of the environmental seclusion postulate is to say that in both cases the notebook, being an extracranial entity is merely an inferred entity,³¹ and not anything we are directly aware of, a move that critics challenge by claiming that interaction with notebooks-in-the-model is not sufficient. What is required to minimize prediction error, they claim, is “concrete, physically realized embodied interaction of an organism with its local environment”.³² A point that has often not been sufficiently stressed when discussing this dispute between the ‘internalist’ interpretations of predictive processing and the ‘externalist’ ones concentrating on embodied, enactive, and extended cognition is that they are in an important respect on a par. The “exteroceptive, proprioceptive, and interoceptive receptors”,³³ the skull, the brain, and spinal cord of the former and the “arms, legs, notebooks, and sticky-notes”³⁴ of the latter cannot be conceived as differing in

²⁹ For some accounts that challenge this postulate see Anderson 2017, Clark 2017, Fabry 2017a,b.

³⁰ Clark 2017: 17.” See also Parr et al 2022: 108-109.

³¹ Hohwy 2016: 275.

³² Fabry 2017b: 406.

³³ Hohwy 2016: 276.

³⁴ Clark 2017: 16.

ontological status.³⁵ Independent of its interpretation, the prediction error minimization framework is formulated in terms of physical entities that, on the one hand, account for the implementation of the prediction error being minimized and are, on the other hand, themselves part of the inferred world of hypothesized entities generated through the framework. It would be peculiar for either interpretation to claim that amongst the inferred 'hidden causes' there are some (skulls, brains, notebooks etc) somewhat less hidden, where the inferred entity matches the object it represents as it really is, while there are others where that object is completely inaccessible to us. If the prediction error minimisation's notion of hidden causes makes any sense at all, they must all be equally hidden.³⁶

What prediction error minimization takes cognition to be, independent of specific implementations consisting of brains, assemblies of brains, notebooks and so forth is drawing inferences from probabilistic patterns in sensory input.³⁷ These patterns are patterns in representations, and as such it is no exaggeration to say that everything we are cognitively engaged with according to prediction error minimization theory are representations.³⁸

2d. Formation

The resulting picture begins to look noticeably Kantian, with the *phenomena* of models based on statistical regularities opposed to the

³⁵ On this point see Clark 2017: 16.

³⁶ A point accepted by Hohwy: "the brain is itself a hidden cause" (2016: 268). See also Shand 2014: 245, note 5.

³⁷ Hohwy 2013: 179, 221.

³⁸ For the idea of equating the agent with a model see Hohwy 2017: 3.

noumena of 'hidden causes'. The way in which predictive processing takes up themes from Kant has not gone unremarked in the literature.³⁹ In addition to the dichotomy between phenomena and noumena, and the fact that for predictive processing, as for Kant, perception is not a process of 'reading off' information from incoming sensory data, but an active process of constructing hypotheses for which the world provides feedback, an interesting example of how specific ideas in predictive processing can be understood as contemporary developments of Kantian themes is the case of hyperpriors.⁴⁰ The concept of hyperpriors also provides a good example of how the principle of formation fits into the prediction-error minimization framework. Prediction-error minimization crucially relies on the generating of predictions or hypotheses on the basis of incoming data. As the number of predictions that can be based on a single set of data is vast, some form of constraint on the set of predictions needs to be assumed. In the process of Bayesian inference, predictions are developed by adapting the values of priors in the light of new incoming information. In order to know which priors to have in the first place, priors about priors — hyperpriors — are required. Hyperpriors provide an interesting analogue to Kant's forms of appearance, space and time. In the same way as these constrain and give form to our experience, hyperpriors constrain and give form to the set of predictions which manifest as the perceived world. For sure, hyperpriors assumed in the predictive processing literature do not

³⁹ See, for example, Swanson 2016, Beni 2018, Piekarski 2017, Zahavi 2018, Gładziejewski (2016: 574), Anderson/Chemero (2013: 204). Links between predictive processing and Berkeley's idealism are explored by Shand 2014 and Norwich 1993, chs 15, 17.

⁴⁰ Swanson 2016: 5-6.

just concern space and time,⁴¹ but assumption about these constitute a good example of these ‘constraints on constraints’ necessary to make the required Bayesian computations tractable.⁴²

2e. Non-correspondence

One crucial difference between prediction error minimization and Kant’s account we should not lose sight of concerns the significant amount of structure many theorists ascribe to the realm of ‘hidden causes’. Not only are these “causal states of affairs in the real world” outside of the Markov blanket causally responsible for our sensory input⁴³ they also causally interact with one another.⁴⁴ In addition to this causal structure some also regard it as the locus of the probabilistic structure the brain’s computations track.⁴⁵ I would argue, however, that neither of these constitutes as much of a conflict with the Kantian picture as one might have thought, despite the severe limitations the latter places on properties we can ascribe to the noumenal realm. When it comes to probabilistic structure it is worthwhile to note that regarding probabilities as degrees of beliefs, and hence as representations, rather than as properties of a mind-independent world appears to be the most natural interpretation of prediction error minimization theory’s Bayesianism.⁴⁶

There is hence no need to locate probabilistic structures in the ‘hidden

⁴¹ They also include some very specific assumptions, such as that light comes from above (Hohwy 2013:116).

⁴² Possible examples of hyperpriors mentioned by Clark (2013:196) include the basic assumption about space that every location contains only a single object, and the basic assumption about time that we can only carry out a single action at any given moment.

⁴³ Hohwy 2013: 220.

⁴⁴ Hohwy 2013: 60, 81.

⁴⁵ Gładziejewski 2016: 571.

⁴⁶ Feldman 2013.

causes' if the Bayesian understanding of probabilities as subjective credences locates them more naturally in our model of the world.

When it comes to causal structure, there are some interpreters of Kant who assume existence of some analogue of causation at the level of the noumenal.⁴⁷ Those who, like me, believe that it makes no sense for a Kantian to speak of causal relations between noumena, or, for that matter, between noumena and phenomena,⁴⁸ rather than exclusively between phenomena, can revert to the suggestion that the reference to causation "in the real world" should be treated along the lines of references to the skull, the brain, and the spinal cord we find in the expositions of prediction error minimization theory. These are essential ingredients for stating the theory but are revealed, once the theory is properly understood, to be nothing over and above parts of the model of the world generated through the process of prediction error minimization. It can therefore be plausibly argued that prediction error minimization theory accords with key ideas delineating the outlines of Kantian transcendental idealism, including the split between the phenomenal and the noumenal, the role which specific features of the mind play in constructing very general properties of the world we live it, such as time and space, and the inaccessibility of the noumenal by any of the conceptual resources we employ for navigating the world.

⁴⁷ Westphal 1997: 231.

⁴⁸ As for the Kantian the causal relation is to be confined exclusively to the phenomenal realm. See Rescher 1974: 178.

This would imply that prediction-error minimization endorses a noumenalist form of the non-correspondence principle, according to which a world behind the appearances is said to exist, though we cannot say anything else about it. After all, prediction error minimization theory postulates the existence of an external world of 'hidden causes' out there which bring about the specific experiences we have, and thereby play an essential role in generating the prediction errors our cognitive system tries to minimize. We should note, however, that once interpreted through the lens of prediction error minimization theory this no longer means quite what we might think it means. First, all talk of 'out there' and 'in here', of 'bringing about' and 'generating' only has a place within our model of the world. Second, if we take the idea of the epistemic inaccessibility of the 'hidden causes' seriously there is no point in describing them as mental, material, representations, mind-independent, parts of the model, abstract, and so forth. The only entities we cognitively interact with are parts of our model of the world, and hence representations, and everything we cannot cognitively interact with we cannot cognitively interact with. This includes describing it in specific ways. If this is right, though, prediction-error minimization appears to align more closely with limitationism. It restricts our theorizing to representations and rejects the coherence of talk about the world behind the representations.

3. Quantum Bayesianism

The history of explicitly or implicitly idealist interpretations of quantum mechanics, as well as that of forms of idealism inspired by quantum mechanics is a large topic well beyond the scope of a single paper.⁴⁹ We will focus here on one specific, relatively recent interpretation that is of particular interest when considering the link between contemporary scientific developments and idealism: Quantum Bayesianism.⁵⁰ Quantum Bayesianism or QBism sets out to interpret quantum mechanics in terms of Bayesian probability theory.⁵¹ Its key idea is that the quantum state is not a micro-state of some physical object ‘out there’ but a subjective probability or degree of belief about the result of a specific measurement.

3a. Advantages of the QBist interpretation

One advantage of QBism is the way it accounts for the collapse of the wavefunction. In a particular experiment, the calculated wavefunction provides the prior probability. Once the experiment is performed and the observation made, the experimenter acquires new information as a consequence. The experimenter then updates their beliefs or subjective probabilities. According to the QBist understanding the collapse of the wavefunction is simply the updating of specific probability judgements.⁵²

This account of the nature of the wavefunction allows for a simple

⁴⁹ Apart from popular classics in this genre like Capra 1975 and Zukav 1979 see, for example, Goswami 1989, Wendt 2015, Kastrup 2021.

⁵⁰ I also lack the space to discuss the work of Markus Müller, whose approach has stronger idealist implications than QBism. See, for example Ball 2017, Müller 2018, 2020.

⁵¹ The philosophical implications of QBism are subject to extensive discussion in the literature, and not all philosophers and physicists writing on this agree about all of them. Nevertheless, I hope that the account given below faithfully represents the view of key contributors to this debate.

⁵² Fuchs et al 2014: 749. For criticism of this understanding of the wavefunction see Brown 2019.

resolution of the paradox of Wigner's friend, as well as making it possible for QBists to diffuse problems involving non-locality that arise from entangled quantum states.⁵³

3b. Representation

QBism has strong representationalist inclinations, since it is exclusively concerned with experiences, not with whatever objects these experiences might represent. According to some interpreters, the "entire purpose" of the QBist enterprise is to "enable any single agent to organize her own degrees of belief about the contents of her own personal experience."⁵⁴ As such, "everything it [i.e. QBism] purports to be about, and everything we could ever hope to possibly assert on its basis, occurs inside the consciousness of some lonely agent."⁵⁵ There is no pre-theoretical stratification of experiences into those reflecting reality's deep structure and those of a superficial or illusory nature:

Everything experienced, everything experiencable, has no less an ontological status than anything else. [...] A child awakens in the middle of the night frightened that there is a monster under her bed, one soon to reach up and steal her arm – that *we-would-call-imaginary* experience has no less a hold on onticity than a Higgs-boson detection event would if it were to occur at the fully operational LHC.⁵⁶

⁵³ For some discussion see von Baeyer 2016 135-137, Fuchs et al 2014: 750.

⁵⁴ Fuchs et al 2014: 750.

⁵⁵ Norsen 2016: 234.

⁵⁶ Fuchs 2010: 21-22.

There appears to be no place in QBism for a direct encounter with the world, for acquaintance with entities outside of specific, subject-relative tokens of experiences. Our entire cognitive access to the world takes place through the middle-man of experiential representations. As such QBism assumes the existence of a boundary between the representation (in the form of experiences) and the represented. Quantum physics and, by extension, all empirical investigation of the world is taken to be exclusively concerned with the former.

This point is not undermined by the fact that degrees of belief are updated in the face of evidence. Given that for the QBist no experience has a higher ontological status than another no sense can be made of updating our beliefs 'against the world'. We can only update them against other experiences or other beliefs, and none of these can claim to be in more intimate contact with the extra-representational world than others. Any world beyond the representation remains inaccessible using QBist resources.

3c. Formation

Even though QBists do not have much to say on the way mental structure shapes what we perceive as the structure of the world, they consider what is usually reified as the spatio-temporal structure of the world as a result of the organizing activity of the mind:

To represent the rich spatio-temporal structure of human experience as mathematical points in a space-time continuum is a smart strategic simplification, but we ought not to confuse our actual experience with a cartoon. [...] Space-time is an abstraction that I construct to organize such experiences.⁵⁷

Moreover, to the extent there are objectively correct ways of setting our probability judgements, these do not flow from a world behind the representations,⁵⁸ but from internal constraints on our judgements, in particular from the Born rule and the Dutch book theorem which function as coherence constraints.⁵⁹ The source of the structure of the world as we identify it, and the objective, or, to be more precise, inter-subjective nature of this identification is to be found in features of our minds, not in a mind-independent world our minds mirror.

3d. Non-correspondence

While the representationalist, experience-focused nature of QBism is widely agreed upon, there is no unified take on non-correspondence. QBism rejects the claim that our representations mirror a world of mind-independent spatio-temporal objects,⁶⁰ though different authors disagree on how precisely this absence of correspondence is to be understood.

3d.1 Noumenalism

⁵⁷ Mermin 2014b: 422-423.

⁵⁸ Fuchs 2017: 119, note 5.

⁵⁹ Glick 2021: 17-18, Fuchs 2017: 119.

⁶⁰ Glick 2021: 8.

Several interpreters take QBism to spell out non-correspondence in a noumenalist manner, sometimes explicitly noting the similarities to the Kantian *noumenon*.⁶¹ The existence of a world behind the representations is assumed; Fuchs points out that “[w]e believe in a world external to ourselves precisely because we find ourselves getting unpredictable kicks (from the world) all the time. [...] [I]n such a quantum measurement we touch the reality of the world in the most essential of ways.”⁶² This point seems to be directed in particular against a misunderstanding of idealism which takes it to imply that we can somehow simply decide the contents of our experiences and the outcomes of our measurements. However, amongst all the proponents of idealism throughout the history of philosophy we will find hardly any who regarded our epistemic input as so malleable that we could quite literally transform our experience of an empty glass of water into that of a full glass by simply wanting it to be so. The frequent resistance of the world to our desires and the existence of surprises are experiential data an idealist will not ordinarily deny.⁶³ While the existence of a representation-transcendent world is affirmed in this way, its description is not what QBism aims at. Though scientific theories are often perceived as seeking to provide a specific account of the fundamental nature of the world, QBism is understood as just setting out to predict specific groups of our experiences (namely those connected

⁶¹ Mohrhoff 2020: 29.

⁶² 2017: 121. How, according to QBism, the world manages to ‘kick’ in a way that shows up in our experience remains entirely mysterious. As Brown (2019: 81) notes, the “part of QBism which relates to “a theory of stimulation and response” between the agent and the world is not grounded in known physics.” Nor, one might add, is prediction error minimization’s theory of the interaction of entities on both sides of the Markov blanket.

⁶³ For further discussion of attempts trying to establish a mind-independent world by kicking see Simons 2017, Massin 2019.

with quantum mechanical experiments). Timpson raises this point in defending QBism against the charge of solipsism,⁶⁴ noting that “the quantum Bayesian simply rejects the idea that the quantum mechanical statements one would typically make describe how things are. [...] [Q]uantum mechanical statements do not provide us with a story about how things are with microscopic systems, a set of facts characterising them [...].”⁶⁵

If quantum mechanical statements read qbistically do not describe how things are ‘out there’ they also cannot make the claim that everything ‘out there’ is fundamentally material or fundamentally mental (or, as the solipsist would want it, that everything ‘out there’ is really ‘in here’). If “science is not about ultimate reality but about what we can reasonably expect”⁶⁶ the idea of treating it as a guide to the nature of what is ontologically fundamental appears to be seriously misguided. While the existence of a world behind representations is assumed, we cannot say anything about what the properties of this world are⁶⁷ — or at least our physics does not provide us with an account of its properties. This world remains wholly inaccessible to our epistemic endeavours. For the QBist, experience provides the epistemic basis for our cognitive engagement with the world, while quantum mechanics is applied to anything beyond

⁶⁴ That QBism doesn't reduce to solipsism, understood as the position that only I myself am fundamentally real is evident once we realize that the world *as we construct it from our experience* contains minds other than our own. As we represent the world, distinct minds are included at the level of representation. And if QBism does not speak about the world *as it is in itself* it obviously cannot say that the world understood in this way only contains my mind. Since there is no third conception of the world, the world according to QBism cannot be understood along solipsist lines.

⁶⁵ 2008: 592.

⁶⁶ von Baeyer 2016: 221.

⁶⁷ Brown 2019: 80.

this basis.⁶⁸ If quantum mechanics, as the QBist understands it, cannot tell us anything about the world ‘out there’, and if experience, by its very nature as an ‘internal personal’ phenomenon cannot either, it is clear that there is no epistemic route whatsoever⁶⁹ to a world behind the experiential representation.

3d.2 Limitationism

Other takes on QBism are much more aligned with a limitationist interpretation, appearing to be uneasy with the idea of a ‘world external to ourselves’ laid out in an objective, observer-independent manner, even if we could never determine the way in which it is so laid out. This is particularly evident in the case of QBism’s response to the problem of Schrödinger’s cat, which appears to deny the cogency of the existence (not simply the epistemic inaccessibility) of a way things are in themselves. If the right response to the question whether the cat is alive or dead is to say that “unperformed experiments have no results”⁷⁰ the claim appears to be not only that we can never know what is the case in the room before the experimenter opens the door, but that there is no point in assuming there to be a fact to the matter. As such, there is no perspective, whether this is the transcendent perspective of a human being or the perspective of the state of nature as such which settles what is ‘really’ going on in the room independent of specific acts of

⁶⁸ Fuchs et al 2014: 751.

⁶⁹ Unless we assume that this world can be accessed by pure reasoning, or by something like artistic or mystical intuition. The number of philosophers who want to rely on these when arguing against idealist accounts of the world is presumably small.

⁷⁰ Peres 1978: 746, von Baeyer 2016: 142-143, Hoffman 2105b: 1553.

observation.⁷¹ References to the nature or existence of a world behind representations are illegitimate⁷² for an account that should only be concerned with the systematisation of representations.

Further support for the limitationist picture is provided by QBism's take on the Heisenberg cut, the dividing line between the quantum and the classical world. This is sometimes referred to as the 'shifty split', since this dividing line (like prediction error minimisation theory's Markov blanket) can move in either direction, without any clear point beyond which further motion is impossible. The same entity can be seen as a part of the classical world (for example an apparatus in a laboratory used for measuring quantum phenomena) or as part of the quantum world (when considered in terms of the structure of the microscopic entities that make up the apparatus). There is no ultimate fact about which side of the cut the apparatus is to be properly placed on.

QBism interprets this cut as the division between the world and the agent's experience of the world, adding that every observer has their own cut.⁷³ Yet if the line between the 'inner' world of experience and the 'outer' world of the experienced differs for every subject it immediately follows that no sense can be made of an objective world behind all experience as it is in itself, whether such a world is ineffable or not. If what is on one side of the cut for Alice is on the other side for Bob, and if we believe we can still speak about a common world of objects Alice and

⁷¹ Glick 2021: 4.

⁷² Peres 1978: 746.

⁷³ Mermin 2012: 8.

Bob represent (rather than each one having only their own experiences, and the objects constructed on the basis of these experiences, which are somehow⁷⁴ supposed to be coordinated) this world cannot be one to which the predicate 'out there' can be applied. (As every instance of 'in here' is relativized to a subject, every instance of 'out there' must be too, so that there is 'out there for Alice', 'out there for Bob' etc). It may be the case that *as a matter of fact* some entity or entities end up at one side of the cut for every one of a given group of observers, but that would hardly reinstate the notion of an objective world behind all experience, since such a world is conceived to be precisely observer-independent, while what happens to be part of the 'outer' world relative to some observers evidently does not have this property in an independent manner, but only in a relative manner. By adding further observers to the group something 'outer' might subsequently end up as 'inner'.

In the resulting picture everything we can reasonably analyse, speak or think about, or use as the data from which to construct our scientific theory is a representation. If the Heisenberg cut is relativised to perceivers, assuming the existence of a world behind representations as it is in itself, independent of any perceiver leads to a contradiction. As far as what belongs to our epistemic compass constitutes our world, there is no world for us beyond our representations. And as far as QBism rejects to notion of a god's eye point of view of the world, there is no world beyond our representations for anyone else either.

⁷⁴ Mermin (2014a: 20) suggests that this coordination takes place through language, though he unfortunately does not make clear how language is supposed to move across sets of subject-related experiences. See also Mermin 2014b: 422.

Both the noumenalist and the limitationalist way of spelling out the non-correspondence principle have specific advantages. Noumenalism fulfils a desire for a foundation on which the world of our experience rests. That certain actions increase fitness, that some of our predictions are accurate, that a given quantum measurement has a specific result is due to something, which is the noumenal reality behind our representations, even though we are unable to say anything about how it can fulfil this role.

Limitationalism, on the other hand, suggests itself to those who are either sceptical of foundationalist projects, or regard the inclusion of ineffable entities into our ontology as problematic (or both). One may ask why one should believe that the world of appearance has any foundation at all,⁷⁵ or why an entity about which nothing can be said is not an explanatorily idle wheel that should not be included amongst our ontological commitments.⁷⁶

Independent of whether we side with the noumenalist or the limitationalist reading of the non-correspondence principle, QBism's adoption of this principle, in connection with representationalism, and at least some

⁷⁵ Westerhoff 2020, chapter 3.

⁷⁶ A position we have not discussed here that might be able to combine the advantages of both noumenalism and limitationism is a view that agrees with the noumenalist that we should assert the existence of a world behind the representations, but concedes to the limitationalist that such a world can be only made sense of *as one of the representations*. (See Westerhoff 2016, 2020). According to this position, though we habitually refer to a 'world out there' and to entities behind our representations, such references do not differ in type from references to representations like tables and chairs.

pointers in the direction of the principle of formation demonstrate it to be in general agreement with central idealist notions.⁷⁷

4. Conclusion

The above discussion has shown how all of the three scientific theories considered, the interface theory, prediction error minimization, and QBism incorporate the three key idealist principles we identified at the outset: representation, formation, and non-correspondence. What is particularly interesting is that all three spell out these principles in terms of specific, conceptually precise constructions. *Representation* is handled by the interface theory in terms of a mapping of the set of world-states to the set of experiences.⁷⁸ Prediction error minimization theory spells it out in terms of a Markov blanket instantiating an evidentiary boundary,⁷⁹ while QBism does so in terms of the Heisenberg cut understood as the distinction between the experiences of each perceiver, and whatever are the objects of these experiences.⁸⁰ In providing examples of *formation* the interface theory introduces the idea of an error-correcting code to explain the sources of the spatial structure of our experience, while prediction error minimization refers to hyperpriors, and QBism to coherence constraints such as the Born rule and the Dutch book theorem. Finally, *non-*

⁷⁷ Brown (2019: 75, 78-81) refers to the “variant of Berkeleyian idealism which suffuses QBism”. For an attempt to link QBism with an ontology of idealist monism see Mohrhoff 2021.

⁷⁸ Hoffman et al 2015a: 1482.

⁷⁹ Hohwy 2016: 264.

⁸⁰ Mermin 2012, von Baeyer 2016: 152-155.

correspondence is introduced by the interface theory via the FBT theorem, while prediction error minimization is prevented from holding a robust view of a structural correspondence between our beliefs and the probabilistic-causal structure of the 'hidden causes' by its subjectivism about probabilities, as well as by the fact that conceptual resources like the concept of causality can only be deployed within the confines of the Markov blanket, not beyond it. Finally, QBism's focus on the systematisations of an agent's experiences implies that the question of a structural correspondence between these and some extra-experiential realm cannot even be coherently formulated against the background of its own understanding of what quantum physics is trying to achieve.

The incorporation of the idealist principles mentioned in the scientific theories we discussed is interesting for at least three reasons. The first concerns empirical *support* for idealism. *If* we regard our science as a guide to metaphysics and *if* any of the three theories of perception, neuroscience, or the interpretation of quantum physics is taken to be best game in town (two conditionals that, to be sure, require a significant amount of additional support), we should also accept the idealist principles that the theories incorporate. Second, the psychology of perception, the role of inference in the functioning of the human mind, and interpretations of the mathematical formalism for describing quantum events at least *prima facie* do not seem to have too much to do with one another. In this respect the way the three theories described *converge* on the idealist principles mentioned is interesting, and might suggest that they have a somewhat deeper significance than simply being one of many

background assumption one could pick to make sense of some empirical data. Finally, the study of idealism through the lens of scientific theories is not primarily interesting because of any prospective ‘scientific proofs’ of idealism this might unearth, but because it provides a better way of *integrating* the questions idealism poses into our large-scale attempts of making sense of the world and suggesting routes for their development. First, investigating whether to accept only some of the three principles, but not all leads to a clearer understanding of the conceptual options available. One might have scientific reasons for endorsing a form of representationalism, without also assuming that the structure of our mind plays a large role in determining the structure of our experience. Or one might take this on board, but refrain from a crucial step in the idealist direction by maintaining that we still manage to achieve a fairly accurate representation of structural features of a mind-independent world.⁸¹ Second, it is interesting to see how the different technical devices uses to spell out the key idealist principles mentioned above might be related to one another, how, for example, the interface theories ‘interface’ is related to the notion of a Markov blanket,⁸² or how Bayesian ideas about probabilities shape both prediction error minimization and QBism.⁸³ Finally, studying the conceptual underpinnings of the three empirical theories discussed allows us to bring the questions that have driven philosophical and theological forms of idealism into the contemporary discussion in a scientifically literate way. Showing how idealist concepts

⁸¹ In this case the notion of an isomorphic representation of structure constitutes the cut between the representation and the represented, demarcating, in the words of Hermann Weyl “the self-evident insurmountable boundary of cognition” (1949: 26).

⁸² For some remarks on this see Fields et al 2017: 268.

⁸³ See e.g. Fields et al 2022.

can be made scientifically precise, or at least provided with scientifically precise analogues (in the form of constructions like the FBT theorem, Markov blankets, hyperpriors etc) allows us to see idealism not simply as a perhaps flamboyant but somewhat quaint exhibit in the philosophical history museum. Instead, we can regard it as the demarcation of a field of questions concerning the nature of representation, the role our minds play in shaping our view of the world, and the nature of the world as it is in itself on which a lively, empirically informed debate of some our most fundamental questions can take place.

Conflict of interest

The author has no conflict of interest to declare.

Bibliography

Michael Anderson, Tony Chemero: "The problem with brain GUTs: Conflation of different senses of "prediction" threatens metaphysical disaster", *Behavioral and Brain Sciences*, 36:3, 2013, 204-205.

Michael Anderson: "Of Bayes and bullets: an embodied, situated, targeting-based account of predictive processing", in Thomas Metzinger, Wanja Wiese (eds): *Philosophy and Predictive Processing 4*. Frankfurt am Main: MIND Group, 2017.

Hans Christian von Baeyer: QBism. The Future of Quantum Physics, Harvard University Press, Cambridge, MA, London, 2016.

Phillip Ball: “Consciously quantum”, New Scientist, 8th November 2017

Majid D. Beni: “Commentary: The predictive processing paradigm has roots in Kant”, Frontiers in Systems Neuroscience 11:98, 2018, 1-3.

David Berman: The Essential Berkeley and Neo-Berkeley, Bloomsbury Academic, London, 2022.

Harvey Brown: “The reality of the wavefunction: Old arguments and new”, in Alberto Cordero (ed): Philosophers Look at Quantum Mechanics, Springer International Publishing, Cham, 2019, 63-86.

Jelle Bruineberg, Krzysztof Dołęga, Joe Dewhurst, Manuel Baltieri: “The emperor’s new Markov blankets:”, Behavioral and Brain Sciences 45: e183, 2022, 1-76.

Fritjof Capra: The Tao of Physics, Bantam Books, New York, 1975.

Andy Clark: “Dreaming the whole cat: generative models, predictive processing, and the enactivist conception of perceptual experience ”, Mind 121:483, 2012, 753-771.

Andy Clark: "Whatever next? Predictive brains, situated agents, and the future of cognitive science", *Behavioral and Brain Sciences* 36, 2013, 181-253.

Andy Clark: "Busting out: predictive brains, embodied minds, and the puzzle of the evidentiary veil", *Noûs* 51, 2016, 727-753.

Andy Clark: "How to knit your own Markov blanket: Resisting the second law with metamorphic minds", Thomas Metzinger, Wanja Wiese (eds): *Philosophy and Predictive Processing 4*. Frankfurt am Main: MIND Group, 2017.

Michael Devitt: *Realism and Truth*, 2nd edition, Princeton University Press, Princeton, New Jersey, 1997.

Jeremy Dunham, Iain Hamilton Grant, Sean Watson: *Idealism. The History of a Philosophy*, Acumen, Durham, 2011.

Regina E. Fabry: "Predictive processing and cognitive development". Thomas Metzinger, Wanja Wiese (eds): *Philosophy and Predictive Processing 4*. Frankfurt am Main: MIND Group, 2017.

Regina E. Fabry: "Transcending the evidentiary boundary: Prediction error minimization, embodied interaction, and explanatory pluralism", *Philosophical Psychology* 30:4, 2017b, 1-20.

Joshua Farris, Benedikt Paul Göcke (eds): The Routledge Handbook of Idealism and Immaterialism: A Historical and Philosophical Study, Routledge, London, New York, 2022.

Jacob Feldman: "Tuning your priors to the world", Topics in Cognitive Science 5: 1, 2013, 13-34.

Chris Fields, Donald D. Hoffman, Chetan Prakash, Robert Prentner: "Eigenforms, Interfaces and Holographic Encoding. Toward an Evolutionary Account of Objects and Spacetime", Constructivist Foundations 12:3, 2017, 265-274.

Chris Fields, Karl Friston, James F. Glazebrook, Michael Levin: "A free energy principle for generic quantum systems", Progress in Biophysics & Molecular Biology 23, 2022.

John Foster: The Case for Idealism, Routledge & Kegan Paul, London, 1982.

John Foster: A World for Us. The Case for Phenomenalistic Idealism, Oxford University Press, Oxford, 2008.

Christopher A. Fuchs, "QBism, the perimeter of Quantum Bayesianism", 2010, arxiv:1003.5209

Christopher A. Fuchs: On participatory realism. In Ian Durham, Dean Rickles (eds): *Information and Interaction: Eddington, Wheeler, and the Limits of Knowledge*, Springer International Publishing, Cham, 2017, 113-134.

Christopher A. Fuchs, N. David Mermin, Rüdiger Schack: “An introduction to QBism with an application to the locality of quantum mechanics”, *American Journal of Physics* 82:8, 2014, 749-754.

Jay L. Garfield, Nalini Bhushan: *Minds Without Fear. Philosophy in the Indian Renaissance*, Oxford University Press, New York, 2017.

Paweł Gładziejewski: “Predictive coding and representationalism”, *Synthese* 193, 2016, 559-582.

David Glick: “QBism and the limits of scientific realism”, *European Journal for Philosophy of Science* 2021, 11:53, 1-19.

Tyron Goldschmidt, Kenneth L. Pearce (eds): *Idealism. New Essays in Metaphysics*, Oxford University Press, Oxford, 2017.

Amit Goswami: “The idealistic interpretation of quantum mechanics”, *Physics Essays* 2, 1989, 385-400.

Donald D. Hoffman: “Conscious realism and the mind-body problem”,
Mind & Matter 6:1, 2008, 87–121.

Donald D. Hoffman, Manish Singh, Chetan Prakash: “The interface theory
of perception”, Psychonomic Bulletin & Review 22:6, 2015a, 1480–1506.

Donald D Hoffman, Manish Singh, Chetan Prakash: “Probing the interface
theory of perception: reply to commentaries”, Psychonomic Bulletin &
Review 22:6, 2015b, 1551–76.

Donald Hoffman, Chetan Prakash: “Objects of consciousness”, Frontiers of
Psychology, 5:577, 2014, 1-22.

Donald D. Hoffman: The Case Against Reality. Why Evolution Hid the Truth
from our Eyes, W.W. Norton & Company, New York, London, 2019.

Jakob Hohwy: The Predictive Mind, Oxford University Press, Oxford, 2013.

Jakob Hohwy: “The self-evidencing brain”, Nôus 50:2, 2016, 259–285.

Jakob Hohwy: “How to entrain your evil demon”, Thomas Metzinger,
Wanja Wiese (eds): Philosophy and Predictive Processing 4. Frankfurt am
Main: MIND Group, 2017.

Jakob Hohwy: "New directions in predictive processing", *Mind & Language*,
Mind & Language 35, 2020, 209–223.

Bernardo Kastrup: *Science Ideated*, iff Books, Winchester, Washington,
2021.

Franz von Kutschera: *Die Wege des Idealismus*, Mentis, Paderborn, 2006.

Larry Laudan: "Explaining the success of science: beyond epistemic
realism and relativism", in A.I. Tauber (ed.): *Science and the Quest for
Reality*, Palgrave Macmillan, London, 1997, 137–161.

Olivier Massin: "Realism's kick." in Christoph Limbeck-Lilienau, Friedrich
Stadler (eds): *The Philosophy of Perception. Proceedings of the 40th
International Ludwig Wittgenstein Symposium*, de Gruyter, Berlin, 2019,
39–56.

John Makeham: *Transforming Consciousness. Yogācāra Thought in Modern
China*, Oxford University Press, Oxford, 2014.

Richard Menary, Alexander James Gillett: "Are Markov blankets real and
does it matter?" in Dina Mendonça, Manuel Curado, Steven S. Gouveia
(eds):

The Philosophy and Science of Predictive Processing, Bloomsbury Academic, London, 2021, 39-58.

N. David Mermin" Quantum mechanics: Fixing the shifty split", Physics Today 65: 7, 2012, 8-10.

N. David Mermin: "QBism in the New Scientist" 2014a, <https://arxiv.org/abs/1406.1573v1>.

N. David Mermin: "QBism puts the scientist back into science", Nature 507: 7493, 26 March 2014b, 421-423.

Ulrich Mohrhoff: "QBism: A Critical Appraisal" 2020, <https://arxiv.org/abs/1409.3312>

Ulrich Mohrhoff: "A QBist ontology", 2021, <https://arxiv.org/abs/2005.14584>

Markus P. Müller: "Mind before matter: reversing the arrow of fundamentality", 2018, <https://arxiv.org/abs/1812.08594>

Markus P. Müller: "Law without law: from observer states to physics via algorithmic information theory", 2020, <https://arxiv.org/abs/1712.01826>

Travis Norsen: “Quantum solipsism and non-locality”, in Mary Bell and Shan Gao

(eds): *Quantum Nonlocality and Reality: 50 Years of Bell’s Theorem*, Cambridge University Press, Cambridge, 2016, 204-237.

Kenneth Norwich: *Information, Sensation, and Perception*, Academic Press, San Diego, 1993.

Gerard O’Brien, John Opie: “Notes toward a structuralist theory of mental representation”, in Hugh Clapin, Phillip Staines, Peter Slezak (eds): *Representation in Mind: New Approaches to Mental Representation*, Elsevier, Amsterdam 2004, 1-20.

Thomas Parr, Giovanni Pezzulo, Karl J. Friston: *Active Inference. The Free Energy Principle in Mind, Brain, and Behavior*, MIT Press, Cambridge, MA, London 2022.

Chetan Prakash, Kyle D. Stephens, Donald D. Hoffman, Manish Singh, Chris Fields: “Fitness beats truth in the evolution of perception”, *Acta Biotheoretica* 69:3, 2021: 319-341.

Asher Peres: “Unperformed experiments have no results”, *American Journal of Physics* 46, 1978, 745-747.

Michał Piekarski: "Commentary: Brain, Mind, World: Predictive Coding, Neo-Kantianism, and Transcendental Idealism", *Frontiers in Psychology* 8:2077, 2017, 1-3.

Chetan Prakash: "On Invention of Structure in the World: Interfaces and Conscious Agents", *Foundations of Science* 25, 2020, 121-134.

Chetan Prakash, Chris Fields, Donald D. Hoffman, Robert Prentner, Manish Singh: "Fact, Fiction, and Fitness", *Entropy* 22:514, 2020, 1-23.

Dale Purves, Yaniv Morgenstern, William T. Wojtach: "Perception and reality: why a wholly empirical paradigm is needed to understand vision", *Frontiers in Systems Neuroscience*, 9:156, 2015, 1-10.

Christophe de Ray: "An evolutionary sceptical challenge to scientific realism", *Erkenntnis* 87, 2022, 969-989.

Nicholas Rescher: "Noumenal causality" in Lewis White Beck (ed): *Kant's Theory of Knowledge*, D. Reidel, Dordrecht, 1974, 175-183.

John Shand: "Predictive mind, cognition, and chess", *Analysis* 74:2, 2014, 244-249.

Peter Simons: “Road safety. Why agency and intentionality were made for each other (and how this refutes idealism)”, in Jesús Padilla-Galvez, Margit Gaffal (eds): *Intentionality and Action*, de Gruyter, Berlin 2017, 23–34.

Peter Simons: “The long and winding road. Folly and feedback in metaphysics”, *Grazer Philosophische Studien* 98, 2021, 75-89.

Timothy Sprigge: *The Vindication of Absolute Idealism*, Edinburgh University Press, Edinburgh, 1983.

Paul Snowdown: “How to interpret ‘direct perception’”, in Time Crane (ed.): *The Contents of Experience*, Cambridge University Press, Cambridge, 1992, 48-78.

Matthew Soteriou: *Disjunctivism*, Routledge, London, New York, 2016.

L.R. Swanson: “The predictive processing paradigm has roots in Kant”, *Frontiers in Systems Neuroscience* 10:79, 2016.

John Taber: “Philosophical Reflections on the *sahopalambhaniyama* argument”, in Birgit Kellner, Patrick McAllister, Horst Lasic, Sara McClintock (eds): *Reverberations of Dharmakīrti’s Philosophy. Proceedings of the Fifth International Dharmakīrti Conference Heidelberg August 26 to 30, 2014*, Austrian Academy of Sciences Press, Vienna 2020, 441-462.

Christopher Gordon Timpson: "Quantum Bayesianism: A study", *Studies in History and Philosophy of Modern Physics* 39, 2008, 579-609.

Alexander Wendt: *Quantum Mind and Social Science*, Cambridge University Press, Cambridge, New York, 2015.

Jan Westerhoff: "What it means to live in a virtual world generated by our brain", *Erkenntnis*, 2016, 81:3, 507-528.

Jan Westerhoff: *The Non-Existence of the Real World*, Oxford University Press, Oxford, 2020.

Kenneth R. Westphal: "Noumenal causality reconsidered: affection, agency, and meaning in Kant", *Canadian Journal of Philosophy* 27:2, 1997, 209- 245.

Hermann Weyl: *Philosophy of Mathematics and Natural Science*, Princeton University Press, Princeton, 1949.

Wanja Wiese, Thomas Metzinger: "Vanilla PP for philosophers: a primer on predictive processing", in Thomas Metzinger, Wanja Wiese (eds): *Philosophy and Predictive Processing* 4. Frankfurt am Main: MIND Group, 2017.

Dan Zahavi: "Mind, brain, world: predictive coding, neo-Kantianism, and transcendental idealism", *Husserl Studies* 34, 2018, 47-61.

Gary Zukav: *The Dancing Wu Li Masters*, William Morrow and Company, New York, 1979.