Being Realist about Bayes, and the Predictive Processing Theory of Mind

June 8, 2018

Abstract

Some naturalistic philosophers of mind subscribing to the predictive processing theory of mind have adopted a realist attitude towards the results of Bayesian cognitive science. In this paper, we argue that this realist attitude is unwarranted. The Bayesian research program in cognitive science does not possess special epistemic virtues over alternative approaches for explaining mental phenomena involving uncertainty. In particular, the Bayesian approach is not simpler, more unifying, or more rational than alternatives. It is also contentious that the Bayesian approach is overall better supported by the empirical evidence. So, to develop philosophical theories of mind on the basis of a realist interpretation of results from Bayesian cognitive science is unwarranted. Naturalistic philosophers of mind should adopt instead an anti-realist attitude towards these results and remain agnostic as to whether Bayesian models are true. For continuing on with an exclusive praise of Bayes within debates about the predictive processing theory will impede progress in philosophical understanding of actual scientific practice as well as the architecture of the mind.

Keywords: Bayesian cognitive science, representing uncertainty, naturalistic philosophy of mind, scientific realism, underdetermination thesis

1 Introduction

Most contemporary philosophers of mind are methodological naturalists. They tend to believe that philosophical theories of mind should be evaluated by the same criteria that are employed in evaluating scientific theories in the cognitive and brain sciences. Philosophers of mind subscribing to methodological naturalism are committed to the ideas that the cognitive and brain sciences are the ultimate source of knowledge about minds, and that the construction of philosophical theories of mind should be empirically constrained.

A consequence of these commitments is that if there is reason to suspend belief on what a scientific theory in the cognitive sciences alleges regarding mental entities and processes, then methodological naturalists should not construct philosophical accounts of mind that assume the scientific theory is true. On the other hand, if the evidence is sufficiently strong to warrant belief in the truth of the theory, then methodological naturalists would have good reason to construct philosophical accounts that take the scientific theory at face value. In any case, naturalistic theories in the philosophy of mind should be coherent with the epistemically warranted views about the results of relevant scientific investigations.

It turns out that several philosophers of mind subscribing to methodological naturalism and who have formulated philosophical theories on the basis of results from Bayesian cognitive science tend to be scientific realists. Before we articulate this point in detail, let us clarify how we understand the terms 'scientific realism' and 'Bayesian' in what is to follow. We understand 'scientific realism' as a philosophical position concerning the epistemic status of scientific theories and models. In particular, a realist attitude towards Bayesian cognitive science amounts to believing that predictively successful Bayesian models deliver knowledge about both observable and unobservable aspects of the mind and/or brain. If the current evidence justifies this realist view, there is sufficient reason to believe that all entities and processes posited by predictively successful Bayesian models exist. Anti-realism, contrastively, amounts to an agnostic or sceptical attitude towards the existence of any unobservable entity or process posited by predictively successful Bayesian models. According to the anti-realism we have in mind, successful scientific models and theories should be understood as conceptual tools or instruments for achieving practical goals and engaging with the world; they should not be understood as descriptions of how things stand in the world.

Regarding the term 'Bayesian', we follow the common construal in philosophy and cognitive science where 'Bayesian' is meant to be a placeholder for a set of interrelated principles, methods, and problem-solving procedures, which are unified by three tenets. First: uncertainty should be captured by a real-valued function that measures degrees of belief. Second: degrees of belief, at any given time, ought to satisfy the axioms of probability theory. Third: degrees of belief, represented by determinate probabilities, ought to be updated in the light of new information, typically by the canonical rule of conditionalisation. However, the third tenet may be construed more broadly since other learning and inference rules have proved successful including Jeffrey conditionalisation (Zhao and Osherson 2010), the minimisation of a statistical distance such as the Kullback-Leibler divergence (Diaconis and Zabell 1982, Eva and Hartmann 2018), and formal procedures for approximating posterior probability distributions like variational free energy minimisation (Hinton and van Camp 1993) and Monte Carlo methods (Sanborn, Griffiths, and Navarro 2010).

Terminology aside, we now clarify the target of the paper guilty of unwarranted Bayesian realism, namely the so-called predictive processing theory of mind and brain (Clark 2013, 2016; Hohwy 2013). In short, the theory views the mind as being fundamentally engaged in prediction error minimisation – that is, minimising the mismatch between internally generated predictions of sensory inputs and actual sensory inputs generated externally. In the development of the theory, the predictive processing theory of mind has posited various theoretical entities and processes including neurally encoded "hierarchical probabilistic generative models", "predictions" and "prediction errors", and "precision weighting" of prediction errors. Insofar as predictive processing theorists adopt a positive epistemic stance towards these posits based on results in Bayesian cognitive science, the general argument we articulate below against such stance applies to these posits, which we argue are not yet worthy of the epistemic commitment.

It is also worth clarifying early on why our contribution in this paper is important for both philosophers and cognitive scientists. Bayesianism has become ever more prominent in the cognitive and brain sciences, as well as in naturalistic philosophy of mind that is concerned with predictive processing. Driven by mathematical advances in statistics and computer science, along with engineering successes in fields such as machine learning, Bayesian modelling has become incredibly useful for studying brain function and mental phenomena including perception, motor control, learning, decision-making, and reasoning (Chater, Tenenbaum and Yuille 2006; Knill and Richards 1996; Körding 2007; Rao, Lewicki, and Olhausen 2002; Oaksford and Chater 2007; Tenenbaum et al. 2011). Given Bayes' success, it is thus important to examine the current epistemic status of Bayes, and how this status would re-configure our understanding of the nature of predictive processing.

With these clarifications in place, we now point to textual evidence supporting the allegation that several philosophers of mind relying on results from Bayesian cognitive science tend to be scientific realists. In a seminal article developing the predictive processing theory, for instance, Clark (2013) claims that "[T]he computational framework of hierarchical predictive processing realises [...] a robustly Bayesian inferential strategy, and there is mounting neural and behavioural evidence that such a mechanism is somehow implemented in the brain" (Clark 2013, p. 189). Hohwy (2013), even more explicit, says:

> [T]here is converging evidence that the brain is a Bayesian mechanism. This evidence comes from our conception of perception, from empirical studies of perception and cognition, from computational theory, from epistemology, and increasingly from neuroanatomy and neuroimaging. The best explanation of the occurrence of this evidence is that the brain *is* a Bayesian mechanism that is, in fact, engaged in inference, belief, and decision. (Hohwy 2013, p. 25).

In introducing a special journal issue, Kirchoff (2017) echoes Hohwy and Clark on"Bayesian prediction processing" being committed to the sweeping, causal claim that "all psychological phenomena come about through the same process: minimisation of prediction error and precision estimation." This commitment would presumably be justified by evidence from Bayesian modelling in the cognitive sciences. Madary (2016, 95ff) confirms this suspicion by taking evidence from the "Bayesian predictive processing approach" as providing empirical support to the thesis that visual perception is an ongoing process of anticipation and fulfillment. While not a predictive processing theorist, Rescorla (2015, 2016) too holds a realist stance towards the results of Bayesian sensorimotor psychology in arguing for intentional realism concerning mental representation.[1]

Despite the many predictive processers who are apparently committed to realism, we find it is currently unwarranted as the appropriate epistemic attitude towards the results of Bayesian cognitive science. If philosophers of mind like Clark, Hohwy, Rescorla, and others are committed to methodological naturalism, then they should not believe predictively successful Bayesian models have actual counterparts in mental architectures or brains. Instead, their arguments about the architecture of the mind, mental representation, and the nature of mental phenomena, should reflect agnosticism or scepticism in the existence of the theoretical entities and processes posited by Bayesian models.

[1]Another example is Danks (2014) who argues for a version of "representation realism", according to which many cognitive representations in the human mind are graphical models. As Danks (2014, p. 5) explains, graphical models consist "of two components: (i) a graph encoding qualitative information about the structure (causal, informational, or other) among various factors; and (ii) a quantitative representation of the relation strengths". As graphical models may or may not involve Bayesian updating, they need not be Bayesian. Because Danks is careful not to equate graphical models with Bayesianism (see in particular Chapter 8.2.2 of his book), his representation realism is beyond the scope of our paper.

We suspect that most cognitive scientists employing Bayesian models are less concerned with the truth of any particular Bayesian model than with its degree of fit with measurement outcomes. Philosophers of mind working on predictive processing, however, have unfortunately put relatively little effort towards making the nature, aims, and scope of Bayesian modelling in cognitive science transparent. Philosophical work on predictive processing has typically taken Bayesian results at face value, and has tried to draw conclusions for traditional debates in the philosophy of mind such as the nature of conscious experience, the relationship between perception and action, and representationalism by assuming that"Bayesian predictive processing" is true.[2]

This lack of attention to actual scientific practice has bolstered Bayes' reputation in the philosophy of mind while covering up the controversies over the empirical and theoretical adequacy of Bayesianism (see, e.g., Bowers and Davis 2012a, 2012b; Chater et al. 2011; Griffiths et al. 2012; Jones and Love 2011; Marcus and Davis 2013; for recent philosophical treatments see Colombo and Seriè s 2012; Colombo and Hartmann 2017; Eberhardt and Danks 2011; Icard 2018; Zednik and Jäkel 2016). This lack of attention is consequential. If Bayesian models are not to be understood realistically, philosophical controversies about the status of representations in Bayesian models, or their bearing on externalist (or internalist) arguments about mental states, are pointless. If our argument succeeds in undermining Bayesian realism, the nature and epistemic status of arguments concerning the predictive processing theory should be reappraised.

In looking ahead, our argument will not appeal to an implausible version of underdetermination, wherein there are logically possible modelling approaches that are equally well-supported by available empirical evidence or rationally defensible as Bayesianism. Rather, our argument will start by reconstructing what we call the *argument from uncertainty for Bayesian cognitive science*. This argument provides cognitive scientists with a common reason for choosing Bayesianism, and begins from the observation that uncertainty is an in-eliminable feature of cognitive systems' interactions with the world. In order to survive and behave adaptively, biological cognitive systems must rely on knowledge derived from sparse, noisy, or ambiguous sensory data produced by a constantly changing environment. Because sensory data are often sparse, ambiguous, or corrupted by noise, cognitive systems would constantly face problems of inference and decision making under uncertainty. Unless these problems are effectively solved, reliable action, accurate perception, and adaptive learning would not be achievable.

Since uncertainty is an in-eliminable feature of cognitive systems' interactions with the world – so continues this argument – the explanatory framework cognitive scientists use to seek explanations of mental capacities and phenomena should account for how cognitive systems effectively deal with uncertainty. That is, the framework ought to include sound inferential and decision-making procedures that biological cognitive systems would deploy when interacting with an uncertain environment. Given the success of Bayesianism in other fields dealing with inference and decision-making under uncertainty, one may conclude that Bayes should also serve as a successful explanatory approach in the cognitive sciences. Upon reconstructing the *argument from uncertainty for Bayesian cognitive science*, we will clarify the kinds of reasons that are commonly cited to justify the choice to adopt Bayesianism as a favourite approach for cognitive and neural modelling.

After putting these reasons into focus, we will explicate under which conditions the choice to adopt a Bayesian framework is justified. Over time, the issue will be resolved empirically; and in fact, some debates in cognitive science are currently focused on the

[2]For a sample of this expanding literature, see the papers in Metzinger and Wiese (2017).

question of whether there is some non-Bayesian model that is equally good at fitting behavioural or neural data. But in the meantime, the reasons for supporting the choice to adopt Bayesianism fails to include its non-empirical (or superempirical) virtues. Our contribution fills this gap in the literature, and continues on by examining whether the argument from uncertainty provides enough support for choosing Bayes as the favourite approach to understanding mind and brain. If not, and if the empirical evidence fails to clearly favour Bayesianism over alternatives, then, based on both empirical considerations and a comparison of the non-empirical virtues enjoyed by Bayes and its alternatives, a realist attitude towards the results of Bayesian modelling is unwarranted.

In developing this argument, we first clarify what the argument from uncertainty is supposed to establish (Section 2) and what role it plays in scientific practice (Section 3). Then, we consider whether there is any serious, actually conceived, alternative to Bayesianism for representing uncertainty and explaining how biological cognitive systems can effectively manage it. We then go on to argue that if plausible alternatives are available but systematically neglected, then the strength of the argument from uncertainty is significantly weakened (Section 4). A key feature of our argument at that point is that there is no convincing reason to believe that Bayes enjoys special epistemic virtues in comparison to its alternatives.

In the conclusion of the paper, we discuss a practical reason for choosing a Bayesian approach. We realise that many of the Bayesian cognitive scientists we cite are in principle open to non-Bayesian methods for modelling certain problems. Yet, these scientists hardly ever discuss the virtues and problems of the alternatives. One reason why alternatives might get neglected is that the Bayesian approach, in comparison to these alternatives, currently affords cognitive scientists with a richer body of conceptual and empirical results as well as computational tools that have been well-developed and employed in neighbouring fields of machine learning and statistics. So, the popularity of Bayesianism in cognitive science, similar to other scientific fields, is not necessarily grounded by the pursuit of truth, empirical evidence, or by super-empirical virtues like simplicity and unifying power. Cognitive scientists may choose Bayes because, pragmatically, there are a lot of tools and methods from neighbouring fields to draw from (Gigerenzer 1991). Upon recognising this sociological point, the claims put forward by naturalistic philosophers of mind like "the brain is literally Bayesian" (Hohwy 2015, p. 17) or that current sensorimotor psychology strongly supports a realist attitude towards the theoretical posits of Bayesian models (Rescorla 2015, 2016) are not well-supported scientifically. We briefly discuss how this conclusion is consequential for framing current disputes about the predictive processing theory of mind; we'll suggest that predictive processing may be best understood "aesthetically", as offering an invitation to see humans and the place of mind in the world from a new vantage point (cf., Prinz forthcoming).

2 From Uncertainty to Bayesian Brains

The *argument from uncertainty* aims to provide a good reason for thinking that Bayesianism is the best way to explain mental phenomena involving uncertainty. The argument includes two steps. The first step substantiates the claim that biological cognitive systems must effectively deal with uncertainty in order to survive and thrive (i.e., in order to interact adaptively with their environment). The second step tries to establish that Bayesianism is the best approach for explaining how cognitive systems effectively deal with uncertainty.

2.1 Uncertainty, Underdetermination, and Noise

Bayesian cognitive scientists typically introduce their studies by pointing out that the mind must constantly grapple with uncertainty. For example, Knill and Pouget (2004) introduce their discussion of the Bayesian brain hypothesis by claiming that "humans and other animals operate in a world of sensory uncertainty" (Knill and Pouget 2004, p. 712). Ma et al. (2006) motivate their study on how populations of neurons might perform Bayesian computations by saying that "virtually all computations performed by the nervous system are subject to uncertainty" (Ma et al. 2006, p. 1432). Oaksford and Chater (2007) advocate a Bayesian approach to human cognition suggesting that "human reasoning is well-adapted to the uncertain character of everyday reasoning to integrating and applying vast amounts of world knowledge concerning a partially known and fast-changing environment" (Oaksford and Chater 2007, p. 67).

Orbán and Wolpert (2011) focus on Bayesian approaches to sensorimotor control and motivate their focus by saying that "uncertainty is ubiquitous in our sensorimotor interactions, arising from factors such as sensory and motor noise and ambiguity about the environment" (Orbán and Wolpert 2011, p. 1). Tenenbaum et al. (2011) point out that "we build rich causal models, make strong generalisations, and construct powerful abstractions, whereas the input data are sparse, noisy, and ambiguous – in every way far too limited" (Tenenbaum et al. 2011, p. 1279). Vilares and Körding (2011) review several lines of research in Bayesian cognitive neuroscience emphasising that "uncertainty is relevant in most situations in which humans need to make decisions and will thus affect the problems to be solved by the brain" (Vilares and Körding 2011, p. 22). Pouget et al. (2013, p. 1170) write that "uncertainty is an intrinsic part of neural computation, whether for sensory processing, motor control or cognitive reasoning."

Provided the apparently broad consensus that cognitive systems inevitability face uncertainty, there are two ways to interpret these cognitive scientists' claims. According to one reading, the point is merely expository. Claims like the ones just quoted get readers to see why the Bayesian approach is worth considering when one is interested in understanding how cognitive systems can deal with uncertainty. According to another reading, the point is justificatory. The claims merely function as a way to convince readers that the choice to pursue the Bayesian approach is justified.[3]

Granted, cognitive systems must grapple with uncertainty, but an inference from 'biological agents are faced with states of uncertainty' to 'Bayes is the best' is invalid on either reading. If the point is merely expository, then it is puzzling that no alternatives to the Bayesian approach are ever mentioned in the literature. Bayesianism is one possible approach for representing and handling inference under uncertainty, but it is by no means

[3]We should note that this justificatory interpretation has few if any explicit proponents in the cognitive sciences. We suspect pretty much every cognitive scientist agrees that more direct empirical confirmation than any current argument could possibly provide is needed for us to be justified in believing that mental processes are Bayesian. Some philosophers, on the other hand, have come close to suggesting this justificatory interpretation. Hohwy (2013), for example, presents Bayesian priors and likelihoods as furnishing minds with "required additional constraints" (p. 49) with normative import for dealing with uncertainty. Gadziejewski (2016) motivates predictive processing by claiming that "the brain deals with uncertainty by implementing or realising (approximate) Bayesian reasoning." (p. 561) Williams (2018) advertises Bayesianism as offering "attractive explanations of [...] how brains overcome the noise and ambiguity in their sensory evidence." (pp. 4–5)

the only one. On the other hand, if the point is justificatory, then there is an implicit assumption that Bayesianism is somehow more empirically adequate, or enjoys greater (or more) non-empirical virtues over alternative frameworks. However, such an assumption can be easily undermined, as it is far from obvious that Bayesianism is simpler, more unifying, or more rational than alternative approaches. We will have much more to say on this issue later on in Section 4.

2.2 The Explanatory Power of Bayes

Biological cognitive systems access the world through their senses, which are viewed as sources of uncertain information about the state obtaining in the world at any given time. In statistical terminology, we may refer to the states of the world as 'environmental parameters' or 'hidden states', and the sensory information as 'sensory data' or 'evidence'. The problem biological cognitive systems face at any given time is inferring which hidden state in their environment generated their sensory data. Unless biological cognitive systems can effectively solve this problem, such systems become incapable of adaptive action and perception, and reliable learning.

However, the values of environmental parameters are underdetermined by the available sensory data – that is, there are multiple, different states in the world that fit the sensory data received by a biological cognitive system at any given time. Because many different environmental states may be consistent with the same piece of sensory data received, processing sensory data alone is not sufficient to determine which state in the world caused it. Hence, sensory data underdetermine their environmental causes, which manifests a state of uncertainty within the system. In Bayesian cognitive science, the term 'uncertainty' is used broadly to characterise this problem of underdetermination that biological cognitive systems must constantly solve.

As an example, the sensory data generated by a convex object under normal lighting circumstances underdetermines its external cause. In such environment, there are at least two possible states that can fit the available sensory data: the object in the world that caused the data is convex and the light illuminating the object comes from overhead; or the object is concave and the illuminating light comes from below. In order to perceive the world as being in one specific state, as opposed to being in a superposition of two or more different states, cognitive systems must find some method to solve this problem.

Furthermore, uncertainty may also arise from noise, where the source is internal or external to biological cognitive systems. As noisy signals contain meaningless information, noise modifies the original meaning of a signal and extends the cognitive system's freedom of choice in decoding that meaning. This is an undesirable freedom, to the extent that the adaptive behaviour the system can produce requires an appropriate degree of fidelity between original and decoded signals. While noise poses a challenge for biological systems estimating environmental parameters, "noise permeates every level of the nervous system, from the perception of sensory signals to the generation of motor responses" (Faisal et al. 2008, p. 292).

Within a biological cognitive systems' signal processing, there are three sources of noise. The first source of noise lies in the thermodynamic or quantal transduction of the energy comprised by sensory signals into electrical signals. "For example, all forms of chemical sensing (including smell and gustation) are affected by thermodynamic noise because molecules arrive at the receptor at random rates owing to diffusion and because receptor proteins are limited in their ability to accurately count the number of signalling molecules" (Knill et al. 1996, p. 4). The second source of noise lies in biophysical features of ion channels, of synaptic transmission, of network interactions and random processes governing neural activations. These biophysical features introduce noise at the level of cellular signalling. A third source of noise lies in the transduction of signals carried by motor neurons into mechanical forces in muscle fibers. This transduction introduces noise in the signals supporting motor control, and can make motor behaviour highly variable even in the same types of circumstances when the same motor goal is pursued. To perform motor commands reliably and behave optimally, biological systems must find some strategy to handle the noise introduced at different levels of neural processing.

We have described how uncertainty tends to manifest within a biological cognitive system either from noise (i.e., random disturbances corrupting the sensory signals and processes of the system) or underdetermination of percepts, along with other cognitive states, by sensory data. Whether caused by noise or surfacing from ambiguity, uncertainty arises from environment and behaviour. The environment (including the body) constantly changes; and even if the environment stays fixed, a cognitive system's behaviour shows an in-eliminable degree of variability. For example, if you reach for an object in darkness, your visual and motor systems will lack relevant information about the location of the object. Your uncertainty about its location will be reflected by a lack of accuracy in any one reaching trial. If you try to reach for that object over and over again, you'll observe a large variability in your movement over reaching trials. Likewise, when a visual stimulus is held constant, your visual perceptions of the stimulus will still manage to vary over time. In order for biological systems to have accurate perceptions and to display reliable motor behaviour, they must find some way to tame this variability.

Biological cognitive systems would effectively deal with sensory and motor uncertainty if they maintain, "at each stage of local computation, a representation of all possible values of the parameters being computed [in accordance to Bayes' rule] along with associated probabilities" (Knill and Pouget 2004, p. 713). This idea is borne out by recent advancements in machine learning, where Bayesian methods are often used to solve problems of underdetermination and to mitigate detrimental effects of noise (Ghahramani 2015). As beneficiaries of the computer scientists' and machine learners' labor, cognitive scientists may employ the same methods to seek explanations of central aspects of cognition, brain and behaviour.

2.3 Bayes and Uncertainty: A Natural Marriage?

The second step in the argument from uncertainty tries to establish that Bayesianism is the best for seeking explanations of how biological cognitive agents grapple with uncertainty, focusing specifically on non-empirical (or super-empirical) virtues. That is, the argument tries to show that Bayes is able to explain most simply, most generally, and most rationally how biological cognitive agents solve the problem of underdetermination and handle the effects of noise. If this step is successful, then we would be justified to claim that the Bayesian approach is the best for discovering and assessing explanations of cognitive, neural and behavioural phenomena. The thought that Bayes is in some sense "the best" for explaining how a system grapples with uncertainty is widely assumed by both philosophers of mind and cognitive scientists. Unfortunately, the thought is hardly clarified and thus left ambiguous.

On one dimension, the Bayesian approach provides a better *language* for representing uncertainty than alternatives. If true, Bayes should be the preferred way of modelling uncertainty and inference in cognitive science. On another dimension, systems that implement Bayesian algorithms deal with uncertainty in the most *rational* way. Bayesianism, then, should be the preferred way if we seek to provide a "rational analysis", guiding explanations of behaviour in terms of an adaptive response to specific problems posed by the environment. While the "optimal language" feature concerns the *representational* virtues of Bayesianism, the "rationality" feature concerns its *normative* character. Both virtues pertain to non-empirical properties of Bayes. We will discuss the relevant evidence and literature for both of these virtues in Section 3 and evaluate Bayes against alternatives in terms of relative non-empirical virtues in Section 4.6. For now, we specify how Bayesian cognitive scientists represent uncertainty in their models, and how they understand the rational character of their models.

Bayesian cognitive scientists *represent* uncertainty through probability.[4] Cognitive systems are assumed to entertain degrees of "belief" over a hypothesis space $\mathcal{A}$. Degrees of belief concern what in the world could have caused the sensory data E currently available to the system. Each belief is associated with a prior probability $P(\mathrm{H})$, which represents the weight borne by the belief that H on the processes carried out by the system on its sensory data. Probabilities are also assigned to (E, H) pairs in the form of a *generative model* that specifies a joint probability distribution over sensory data and hypotheses about states in the world generating those data. Generative models define likelihood functions specifying how probable it is that the system would receive the current data E, given a hypothesised state H about the world, *viz.* $P(\mathrm{E}|\mathrm{H})$. With a generative model, a likelihood $P(\mathrm{E}|\mathrm{H})$, the current data E, and prior knowledge $P(\mathrm{H})$, the system computes the posterior conditional probability $P(\mathrm{H}|\mathrm{E})$, thereby reallocating probabilities across the hypothesis space in accordance with some learning rule, for instance with straight conditionalisation, Jeffrey conditionalisation, distance minimisation, or free-energy minimisation.[5]

Bayesian cognitive scientists generally see themselves as providing *rational* analyses of cognitive functions, where they make a specific hypothesis about the types of problems the mind would face, try to find optimal solutions to these problems, and use these solutions to guide research on how the mind actually solves those problems (e.g., Griffiths, Vul, and Sanborn 2012; Rahnev and Denison 2018; for recent philosophical treatments, see Icard 2018; Zednik and Jäkel 2016). Bayesian rational analyses are typically understood within David Marr's levels of analysis framework (Marr and Poggio 1977). These levels include the computational, the algorithmic, and the implementation level. The computational level specifies the problem to be solved in terms of an input-output mapping. In the case of Bayesian modelling in cognitive science, this is typically a problem of inference under uncertainty. If the task is one of extracting some property of a noisy sensory stimulus, the input-output mapping that defines the computational problem is a function mapping the noisy sensory data to an estimate of the stimulus that caused that data. The class of rules for generating the output and their associated representational posits are defined at the algorithmic level. At the level of implementation, the focus is on how probabilistic representations and Bayesian algorithms can be realised in neural circuits and activities

[4]For simplicity, we will assume that the mathematical object is a finitely additive probability mass function.

[5]Learning rules govern belief update, but they do not specify how the beliefs entertained by the system are used to produce a decision, action, or some other behavioural phenomenon. How the posterior is used to produce a decision requires defining a loss function, which specifies the relative cost of making a certain decision based on a certain belief. To determine the most rational decision available at a given time, the system needs to compute the estimated loss for any given decision and belief.

(Pouget et al. 2013; Ma and Jazayeri 2014). In fact, many Bayesian cognitive scientists are interested in all of Marr's levels of analysis (Griffiths, Vul and Sanborn 2012), and for many naturalistic philosophers of mind working on predictive processing, "the real interest [of Bayesian models] comes from the stronger notion that human beings might actually use the apparatus of probability theory to make their decisions, explicitly (if not consciously) representing prior probabilities, and updating their beliefs in an optimal, normatively sound fashion based on the mathematics of probability theory" (Marcus and Davis 2013, p. 2358).

3 Representing Uncertainty and Explaining Rationally

In the previous section, we distinguished between two ways one might understand the claim that Bayesianism is the best approach for explaining how cognitive systems grapple with uncertainty. One way is in terms of its representational virtues, while the other is in terms of its normative virtues. Both of these virtues rely on non-empirical (or superempirical) features of Bayesianism. Hohwy, Roepstorff, and Friston's (2008) explanation of binocular rivalry offers a good illustration of how predictive processers may appeal to these virtues.

Binocular rivalry is the alternating percept that often results when incompatible images are presented dichoptically. Hohwy et al. (2008) have proposed a Bayesian, predictive processing mechanism as the "best" for explaining binocular rivalry (p. 688). They assume that the brain is a hierarchically organised probabilistic machine, which would model and infer the environmental causes of its perceptual inputs. Hohwy and colleagues' explanation specifies two conditions under which we should expect binocular rivalry. The first condition is that the visual system assigns both high prior probability and high likelihood to no single model of the environmental causes of current visual input. During binocular rivalry, the visual system selects the percept associated with the model that is assigned the highest prior probability. The second condition is that, when one of the two images is selected by the visual system, the sensory input produced by the other image results in an "unexplained but explainable prediction error signal [which] induces instability in perceptual dynamics that can give rise to perceptual alternations" (p. 687). While to the best of our knowledge no specific hypothesis based on this explanation has been experimentally tested, Hohwy and colleagues suggest the explanation they offer is "unifying" (p. 688), "parsimonious" (p. 690), but also "more principled than alternative non-epistemological accounts" (p. 694). In short, Hohwy et al. (2008) demonstrate how an appeal to the representational and normative virtues of Bayes can be used to conclude that a Bayesian approach to explaining certain mental phenomena is the best.

Unfortunately, these virtues alone do not provide philosophers of mind and cognitive scientists with a compelling reason to prefer the Bayesian approach over alternatives as one's basis for discovering and assessing explanations of mental phenomena. In the light of Hohwy et al.'s (2008) case and with the discussion in the previous section in hand, let's address the following questions: what are exactly the properties of Bayesianism that contribute to its representational power? And what are the properties of Bayesianism that contribute to its capacity to define computational problems and to provide "rational" (or "principled") solutions to such problems?

3.1 Representing Uncertainty

Explanatory frameworks possess certain non-empirical, epistemic properties that increase representational power. Two such properties are simplicity and unificatory power. If a framework $\mathcal{F}$ possesses all the properties or each of the properties to greater extent in comparison to an alternative framework $\mathcal{F}^*$, then there is reason, based on non-empirical virtues, to prefer $\mathcal{F}$ to $\mathcal{F}^*$ as one's working framework for scientific explanation.

In the proceeding discussion, simplicity may roughly be understood as some measure of the number and conciseness of the framework's basic principles and representational posits, and unification may roughly be understood as some measure of the number of different kinds of phenomena or systems that the framework can be applied to. Taken together, the simplicity of a framework brings with it pragmatic advantages like being more perspicuous and easier to use and to manipulate, and unification brings with it epistemic advantages related to explanation and confirmation (see e.g., Sober 2003).

In thinking about the Bayesian approach in cognitive science along these lines, the approach achieves some degree of simplicity. The set of principles guiding the approach is fairly minimal: (i) an agent's beliefs that differ in strength are modelled by real numbers; (ii) at any given time, an agent's beliefs obey the axioms of probability; (iii) over time, an agent updates their beliefs according to a rule of conditionalisation.[6] These principles allow cognitive scientists to formulate research questions compactly and precisely (Chater et al. 2006, p. 287). The language also has much unifying power (Tenenbaum et al., 2011, p. 1285). In fact, it offers a common, encompassing, and flexible mathematical language for studying a wide variety of phenomena and systems.

However, even though the Bayesian language appears to be simple, many Bayesian models of real-world, high-dimensional tasks are hard to formulate and manipulate. One challenge concerns computing the posterior distribution, which is intractable (or "computationally complex") for most real-world problems and calls for approximations and heuristics that might themselves be intractable (Kwisthout, Wareham, and van Rooij 2011). Another challenge is choosing a suitable model and prior. A suitable model should not limit the form of probability distributions (e.g., always normal) or functions (e.g., always linear), which are part of the solution to a cognitive task. Priors should not rule out plausible candidate hypotheses by assigning them zero probability, nor should they spread uniform mass over all possible hypotheses. Upon resolving these modelling challenges, one ends up with Bayesian models that are significantly complicated and hard to manipulate.[7]

As for unification, although Bayesianism has been used to fit an impressive range of data from a diverse variety of cognitive and behavioural tasks, this kind of unificatory power does not obviously have explanatory or confirmatory import (Colombo and Hartmann 2017; Eberhardt and Danks 2011). Similar to Bayesianism, the language of Lagrangian field theory can be used for studying many kinds of systems. For example,

[6]Note that there are several schemes for representing precise probabilities such as a full probability distribution, or a scheme requiring the representation of at least two moments (e.g., variance and skewness) of the sensory distribution. While most Bayesian models assume a full probability distribution, this assumption is not supported by empirical evidence (Rahnev 2017).

[7]The notion of "simplicity" regarding these modelling challenges is understood, respectively, as a property of a model itself and as a property of the activity of modelling. More precisely, the "simplicity" ("tractability" or "computational complexity") of a model itself concerns the time a Turing machine requires to return an output for each input. The "simplicity" of the activity of modelling may be explicated as the degree of ease with which a model can be built and manipulated to obtain a desired result (see Colombo 2015, Section 3).

it can be applied both to the behaviour of a system of gravitational masses and that of an electric circuit. This fact, however, does not warrant the conclusion that we have a common explanation of the behaviour of both systems. More generally, while the representational flexibility of Bayesian models contributes to its unifying power, it might prompt cognitive scientists to select priors and likelihood functions post-hoc to merely accommodate but not explain behavioural and neural data (for this kind of criticism see, Bowers and Davis 2012a).

3.2 Explaining Rationally

Bayesianism has undoubtedly much appeal to cognitive scientists because it provides them with a clear normative standard on how adaptive agents should combine and weigh different degrees of belief, how they should update their degrees of belief upon receiving novel information, and how they should make decisions and inferences under uncertainty.

The normative appeal of Bayesianism initially arose from synchronic and diachronic Dutch book arguments justifying its tenets: degrees of belief are (i) probabilistic and (ii) updated via conditionalisation. Dutch book arguments aim to establish that it is practically irrational for an agent to have degrees of belief that violate the probability calculus and/or rule of conditionalisation. After all, either of these violations would expose an agent to accepting a set of bets that guarantees a net loss.

While Dutch book arguments are purely pragmatic, the rationality of Bayesianism has more recently been grounded in accuracy considerations, where accuracy is construed as "closeness to truth" (Joyce 1998). The argument relies on a proper scoring rule for measuring the inaccuracy of a belief function at a possible world w that gradually penalises the function as it becomes more distant from the ideal belief function at w. Leitgeb and Pettigrew (2010a, b) have shown that for some proper scoring rule, $\mathcal{S}$, an agent that has probabilistic degrees of belief and who updates by conditionalisation will have less inaccurate degrees of beliefs compared to an agent with non-probabilistic degrees of belief and who updates in some other way. On the assumption that accuracy is a fundamental feature of epistemic rationality, then Bayesianism is epistemically rational.

Each of these justifications is not uncontroversial, however. Dutch book arguments imply that the utility for a bundle of bets is the summed utilities associated with the individual bets, but this is not true if utility is nonlinear in money. Of course, one might counter by arguing that utility *should* be linear in money. However, there is little empirical basis for mandating this requirement since nonlinear utilities are common in ordinary decision making as a result of risk aversion or diminishing marginal utility.[8] As for the accuracy-based justification, there is much controversy surrounding the central assumption – that is, accuracy is a fundamental feature of being epistemically rational. Littlejohn (2015) points out that some accurate beliefs are worthless, yet the accuracy argument seems to neglect the fact. Besides worthless accuracy, some have argued that proponents have not established feasible epistemic interpretations for the objects of the mathematical framework (see e.g. Carr 2017). Without a compelling story, the accuracy argument is driven purely by fancy mathematics.[9]

[8]Besides the linear vs. nonlinear utility contention with Dutch book arguments, we refer the reader to objections found in Hájek (2008), which also presents a "Czech Book Argument" with the conclusion that agents ought to violate the probability calculus.

[9]Gelman (2008) offers a concise overview of arguments against Bayesian statistics. He focuses on two classes of objections: first, Bayes is overly-flexible, since Bayesianism does not constrain which of many

So, even though Bayesianism appears to be – as Hohwy and collaborators (2008) put it – "more principled than alternative non-epistemological accounts", its normative force is in fact controversial.

4 Tracking Uncertainty: A Zoo of Theories

Early on, we alluded to a problem for proponents of the predictive processing theory and philosophers of mind who take results from Bayesian cognitive science at face value. The problem is a neglect of different yet promising approaches that represent several of the mental phenomena captured by current Bayesian models.

Ignorance is bliss, of course, since naturalistic philosophers may avoid addressing the controversies ignited by sceptics and non-Bayesians. The neglect has also contributed to motivating an unwarranted scientific realism regarding Bayesian models. Bringing the problem into focus, we contend that it should not be taken for granted that states of uncertainty realised by cognitive systems are best (or most naturally) represented and explained within a Bayesian framework. This is because a systematic comparison remains to be seen of the relative epistemic virtues (both empirical and non-empirical) of Bayesianism and alternative approaches to representing and handling uncertainty. Thus, we find the view that 'Bayes is best' has unjustifiably gained the endorsement of many philosophers of mind through systematic neglect of other plausible alternatives.

In this section, we address this deficiency in the philosophical literature by discussing five formal approaches for representing uncertainty including Dempster-Shafer theory, imprecise probability, possibility theory, ranking theory, and quantum probability theory. Along the way, we highlight their logical relationships with Bayes and one another.[10] Of course, what follows is not intended to be an exhaustive literature review (see Halpern 2003; Huber 2014). Nevertheless, the highlighted merits of each theory considered suffice to undermine the assumption that Bayesianism is the *simplest*, *most unifying*, or *most rational* approach to explaining uncertainty-involving cognitive phenomena.

Later in this section, we identify some important perceptual and cognitive phenomena that Bayesianism fails to adequately predict but are predicted by some of the alternative approaches we review. These observations count as strikes against Bayes, further undermining the assumption that Bayesianism is the best way to represent and explain how cognitive systems manage uncertainty. We also think that predictive processers should take note of the different theoretical entities posited by these alternative theories, which, from a scientific realism perspective, would have significant consequences on debates in philosophy of mind currently relying on results from Bayesian cognitive science such as the nature of mental representation and internalism vs. externalism. We conclude the section with a recommendation that naturalistic philosophers of mind should not be Bayesian realists given the empirical successes of these other models that posit different theoretical entities.

possible models should apply in any given task; second, hierarchical and empirical Bayes is implausibly over-sold as an all-purpose, automatic statistical engine.

[10]Although each framework we consider provides a representation of uncertainty within the scope of probability theory broadly construed, there are also implicit, non-probabilistic approaches to uncertainty in cognitive science, and particularly in computational cognitive neuroscience (see e.g., Drugowitsch and Pouget, 2012). We do not discuss such theories here, but making the reader aware of them at least expands the set of alternatives to consider. By expanding the space of possibilities, the case against the argument from uncertainty becomes more convincing.

	$\{\}$	$\{\omega_1\}$	$\{\omega_2\}$	$\{\omega_3\}$	$\{\omega_1, \omega_2\}$	$\{\omega_1, \omega_3\}$	$\{\omega_2, \omega_3\}$	$\{\omega_1, \omega_2, \omega_3\}$
Mass	0	0.1	0.2	0.1	0.1	0.1	0.3	0.1
Belief	0	0.1	0.2	0.1	0.4	0.3	0.6	1
Plausibility	0	0.4	0.7	0.6	0.9	0.8	0.9	1

Table 1: Dempster-Shafer Example

4.1 The Dempster-Shafer Framework

Instead of modelling a state of uncertainty with a probability function, the Dempster-Shafer (D-S) theory of evidence generalises Bayes by representing "degrees of belief" through a pair of non-additive functions and accommodates learning through Dempster's rule for aggregating evidence instead of simple conditionalisation (Shafer 1992). For those with strong commitments to Bayes, however, we find that all is not lost since Bayes is limiting case. The advantage of D-S, though, as a matter of logic, is that D-S does not imply Bayes, and consequently we are afforded more expressive modelling power.

In the Dempster-Shafer theory of evidence, there are three functions used in modelling a state of uncertainty: a *mass* function, a *belief* function, and a *plausibility* function. Let W be a finite set of states. A mass function m is a mapping of subsets from the power set or *frame of discernment* $\wp(W)$ (the set of all subsets including $\oslash$) to the unit interval $[0, 1]$ where $m(\oslash) = 0$ and the sum of the masses for all $\mathrm{X} \subseteq W$ is 1. The D-S belief function, Bel, and the plausibility function, Pl, define a lower and upper bound, respectively, representing the levels of *support* and *lack of evidence against* each element $\mathrm{X} \in \wp(W)$. The lower bound, Bel, is defined as the sum of the masses for all the subsets of a *set of interest*, $\mathrm{X} \in \wp(W)$. The upper bound, Pl, is defined as the sum of the masses for all the subsets that intersect the set of interest. Consider the simple example given in Table 1 involving a finite set of states $W = \{\omega_1, \omega_2, \omega_3\}$.

We see that $\sum_{\mathrm{X} \subseteq W} m(\mathrm{X}) = Bel(W) = Pl(W) = 1$, thus entailing a complete lack of uncertainty with respect to the sure event W. For each proper subset of W, however, the same cannot be said given the above example. Levels of uncertainty associated with each proper subset are realised by $Bel(\cdot)$ and $Pl(\cdot)$. Take the subset $\{\omega_1, \omega_3\}$, for example, which we will label X. The sum of the masses assigned to $\{\}, \{\omega_1\}$, $\{\omega_3\}$, and $\{\omega_1, \omega_3\}$ is 0.3, which is the lower bound, $Bel(\mathrm{X})$. The sum of all the masses of subsets that intersect the set of interest is the upper bound, $Pl(\mathrm{X}) = 0.8$. The pair generates a "belief interval", $[0.3, 0.8]$, which, in addition to uncertainty, captures *partial ignorance* that is illustrated by the difference between Bel and Pl where $Bel(\mathrm{X}) \leq Pl(\mathrm{X})$ always.

Moreover, a further distinctive feature of D-S theory is that the union of disjoint events are believed at least as strongly as the sum of the lower bounds given for each event instead of there being strict equality between the probability of the union and the sum of the individual probabilities. This implies that Bel is *superadditive*, i.e. $Bel(\mathrm{X} \cup \mathrm{Y}) \geq Bel(\mathrm{X}) + Bel(\mathrm{Y})$ for all disjoint elements X, $\mathrm{Y} \in \wp(W)$. The conjugate Pl, on the other hand, is *subadditive*, i.e. $Pl(\mathrm{X} \cup \mathrm{Y}) \leq Pl(\mathrm{X}) + Pl(\mathrm{Y})$ for all disjoint elements X, Y $\in \wp(W)$. Notice, then, that the mathematical objects representing "degrees of belief" in the D-S theory are non-additive, distinct from the classical Bayesian theory.

Regarding inference, D-S theory employs Dempster's rule of combination for aggregating mass functions associated with information from multiple, independent sources:

$$(m_1 \otimes m_2)(X) = \frac{1}{1-K} \cdot \sum_{\{\mathrm{X}_1, \mathrm{X}_2 : \mathrm{X}_1 \cap \mathrm{X}_2 = \mathrm{X}\}} m_1(\mathrm{X}_1)\, m_2(\mathrm{X}_2),$$

where $K = \sum_{\{\mathrm{X}_1, \mathrm{X}_2 : \mathrm{X}_1 \cap \mathrm{X}_2 = \oslash\}} m_1(\mathrm{X}_1)\, m_2(\mathrm{X}_2)$, $\mathrm{X} \neq \oslash$, and $m(\oslash) = 0$. This rule corresponds to a normalised joint operation in which information is combined by favouring the agreement between the sources and ignoring all the conflicting evidence (Dempster 1968).

4.2 The Imprecise Probability Framework

A common model in imprecise probability theory (IP) is a non-empty set of probability functions $\mathcal{P}$ where each probability function $P \in \mathcal{P}$ is defined on an algebra $\mathcal{A}$ generated by means of a finite set of states W (Augustin et al. 2014; Levi 1974; Walley 1991). For the sheltered Bayesians, fear not since Bayes implies IP. But similar to the D-S theory of evidence, IP does not imply Bayes. So we are able to once again obtain more expressive power by using IP while maintaining a connection to Bayes.

Moreover, there are some important details to note about sets of probabilities. For one, they are usually accompanied by bounds where the lower bound is defined by a *lower probability* $\underline{\mathrm{P}}(\mathrm{X}) = \inf\{P(\mathrm{X}) : P \in \mathcal{P}\}$ for all $\mathrm{X} \in \mathcal{A}$. The upper bound is defined by an *upper probability* $\overline{\mathrm{P}}(\mathrm{X}) = \sup\{P(\mathrm{X}) : P \in \mathcal{P}\}$. Lower and upper probabilities are conjugates such that $\underline{\mathrm{P}}(\mathrm{X}) = 1 - \overline{\mathrm{P}}(\mathrm{X}^c)$ and $\overline{\mathrm{P}}(\mathrm{X}) = 1 - \underline{\mathrm{P}}(\mathrm{X}^c)$. Since the instantiation of one necessarily induces the other, it is only required to specify either a lower $\underline{\mathrm{P}}$ or an upper $\overline{\mathrm{P}}$ (the same goes for *Bel* and *Pl* in DS-theory). Proponents of IP usually prefer the language of (coherent) lower probabilities and thus introduce $\underline{P}$ in the model.

Like in D-S theory, imprecise probability is distinct from classical probability in that the lower $\underline{\mathrm{P}}$ and upper $\overline{\mathrm{P}}$ functionals are non-additive. Instead, $\underline{\mathrm{P}}$ is superadditive and $\overline{\mathrm{P}}$ is subadditive: $\underline{\mathrm{P}}(\mathrm{X} \cup \mathrm{Y}) \geq \underline{\mathrm{P}}(\mathrm{X}) + \underline{\mathrm{P}}(\mathrm{Y})$ and $\overline{\mathrm{P}}(\mathrm{X} \cup \mathrm{Y}) \leq \overline{\mathrm{P}}(\mathrm{X}) + \overline{\mathrm{P}}(\mathrm{Y})$, for all $\mathrm{X}, \mathrm{Y} \in \mathcal{A}$. And if we assume that $\mathcal{P}$ is closed under convex combinations – that is, for any $\lambda \in [0, 1]$ and $P_1, P_2 \in \mathcal{P}$, $\lambda P_1 + (1 - \lambda) P_2 \in \mathcal{P}$ – then uncertainty is represented by non-additive, interval-valued, imprecise probabilities: $\mathcal{P}(\mathrm{X}) = [a, b]$, for all $\mathrm{X} \in \mathcal{A}$. On the surface, imprecise probability and D-S theory look very much alike.

By contrast, though, IP is even more general than Dempster-Shafer given that every D-S belief interval is an imprecise probability, but not every imprecise probability is a D-S belief interval (Huber 2016). Furthermore, IP collapses into Bayesianism with less restriction. To demonstrate, suppose that $\mathcal{P}$ is a singleton set for all events in a respective algebra. Then, the lower and upper probabilities are realised by P, i.e., $\mathrm{P} = \underline{\mathrm{P}} = \overline{\mathrm{P}}$. What we learn from this fact is that $\mathcal{P}$ can in principle always be Bayesian, but whether *Bel* is Bayesian or not will ultimately depend on the masses assigned to subsets.

Furthermore, IP more closely resembles Bayes in its inference procedure. Inference proceeds by way of conditioning each individually precise $P \in \mathcal{P}$ on new information E, assuming $P(\mathrm{E}) > 0$. The result is a set of conditional probabilities $\mathcal{P}(.|\mathrm{E})$ bounded by lower and upper conditional probabilities, $\underline{\mathrm{P}}(.|\mathrm{E})$ and $\overline{\mathrm{P}}(.|\mathrm{E})$. In the instance that $\mathcal{P}(.)$ and $\mathcal{P}(\mathrm{E})$ are singleton sets, $\mathcal{P}(.|\mathrm{E})$ is also a singleton set, i.e. $\mathcal{P}(.|\mathrm{E}) = \{P(.|\mathrm{E})\}$. In the latter instance, imprecise probability collapses into Bayesianism with less restriction than D-S since $\mathcal{P}(.|E)$ can in principle always be Bayesian whereas Dempster's rule is a more limiting case. We conclude that IP retains many of the benefits of Bayesianism while also enjoying increased expressive power, similar to Dempster-Shafer.

4.3 The Possibility Framework

Possibility theory was inspired by ideas in fuzzy logic aiming at accommodating vagueness (Zadeh 1975). Using possibility theory for the purpose of measuring degrees of uncertainty rather than degrees of truth, a possibility measure, Π, models "the knowledge of an agent (about the actual state of affairs) distinguishing what is plausible from what is less plausible, what is the normal course of things from what is not, what is surprising from what is expected" (Dubois and Prade 2007). Although the language seems much different, representations of uncertainty closely resembles a D-S representation. So once again, we find another generalisation of Bayes that has increased expressive power.

To see what the theory consists of, we begin by defining a *possibility distribution* π on a finite set of states W. The function π maps a state $w \in W$ to a real number in the unit interval $[0, 1]$ where $\pi(w) = 1$ for at least one $w \in W$. From a possibility distribution, we can construct a possibility measure $\Pi : \mathcal{A} \to \mathbb{R}$ that assigns 0 to $\oslash$ and 1 to W. For any $\mathrm{X} \in \mathcal{A}$, the possibility measure is defined as $\Pi(\mathrm{X}) = \sup_{w \in \mathrm{X}} \pi(w)$. Π provides a degree to which an event is possible where 1 is maximum possibility and 0 is minimum possibility. Like D-S and IP, a possibility measure induces a conjugate measure $\mathcal{N} : \mathcal{A} \to \mathbb{R}$ called *necessity*. The necessity measure is defined as $\mathcal{N}(\mathrm{X}) = \inf_{w \in \mathrm{X}} \pi(w)$ for all $\mathrm{X} \in \mathcal{A}$. A necessity measure $\mathcal{N}$ provides a degree to which an event is necessary.

Distinct from additive probability functions, a possibility measure, Π, has a unique property of "maxitivity". The maxitivity property says that if X and Y are disjoint sets, then $\Pi(\mathrm{X} \cup \mathrm{Y}) = \max(\Pi(\mathrm{X}), \Pi(\mathrm{Y}))$. This means that the union of disjoint sets is at least as possible as the maximally possible disjoint set, yet the union is no more possible than such set – hence, subadditivity. While $\Pi(\mathrm{X})$ is an upper bound with respect to uncertainty toward X, $\mathcal{N}(\mathrm{X})$ is the lower bound where $\mathcal{N}(\mathrm{X}) = 1 - \Pi(X^c)$. Consequently, we obtain a dual property, $\mathcal{N}(\mathrm{X} \cap Y) = \min(\mathcal{N}(\mathrm{X}), \mathcal{N}(\mathrm{Y}))$.

Regarding inference in possibility theory, if $\Pi(\mathrm{Y}) > 0$ and the set X is non-empty, then one way to account for new information is as follows:

$$\Pi(\mathrm{X}|\mathrm{Y}) = \begin{cases} 1 & \text{if } \Pi(\mathrm{X} \cap \mathrm{Y}) = \Pi(\mathrm{Y}); \\ \Pi(\mathrm{X} \cap \mathrm{Y}) & \text{if } \Pi(\mathrm{X} \cap \mathrm{Y}) < \Pi(\mathrm{Y}). \end{cases}$$

The difference between conditional possibility and conditional probability is that $\Pi(\mathrm{X} \cap \mathrm{Y})$ cannot be the product $\Pi(\mathrm{X}|Y)\,\Pi(\mathrm{Y})$ in an ordinal setting, so $\times$ is replaced by min.

With the bigger picture in mind here, Π and $\mathcal{N}$ are similar to Pl and Bel, respectively. In fact, if a mass function m on a finite frame of discernment is consonant by assigning positive mass only to an increasing sequence of sets, then a plausibility function Pl, relative to m, is a possibility measure (see Theorem 2.5.4 in Halpern 2003). With Bel_m being the conjugate of Pl_m, it follows that Bel_m is a necessity measure. Thus, possibility theory is a limiting case of D-S and also IP.

4.4 The Ranking Framework

Ranking functions can be viewed as measures of how surprising it would be if an event were to occur (or if some hypothesis is true). Formally, a ranking function $\mathcal{K} : 2^W \to \mathbb{N} \cup \{\infty\}$ models the degree of disbelief or surprise assigned to a subset of a finite space W (Spohn 2012). We say that a subset X is surprising or disbelieved just in case its rank is positive, i.e. $\mathcal{K}(\mathrm{X}) > 0$. Subsets that are disbelieved are ranked gradually by the natural numbers to a maximum of ∞. The higher rank, the higher degree of surprise. Intuitively, the empty set should be disbelieved to the highest degree.

On the other hand, if any subset X of W is not at all surprising, then that subset is assigned a rank of 0. Intuitively, W should never be disbelieved for it is not surprising that one of the states in W will obtain. However, unsurprisingness does not necessarily imply that if X is assigned a rank of 0, then X is believed. A subset X is said to be believed just in case its complement is disbelieved, i.e. $\mathcal{K}(\mathrm{X}^c) > 0$. Otherwise, a rank of 0 assigned to X and X^c would seem to suggest suspension of judgment since one has yet to come to disbelieve one of the disjoint sets.

Regarding inference in the ranking theory, a conditional model of uncertainty may be defined like so: $\mathcal{K}(\mathrm{X}|\mathrm{Y}) = \mathcal{K}(\mathrm{X} \cap \mathrm{Y}) - \mathcal{K}(\mathrm{Y})$. By using conditional ranks, the main rules for updating on new information correspond to Bayesian conditionalisation through an analogue to Bayes' rule: $\mathcal{K}(\mathrm{X}|\mathrm{Y}) = \mathcal{K}(\mathrm{Y}|\mathrm{X}) + \mathcal{K}(\mathrm{X}) - \mathcal{K}(\mathrm{Y})$.

As the theory fits into the larger picture, ranking functions are closely related to possibility measures and can be transformed in the following way: $\Pi_{\mathcal{K}}(\mathrm{X}) = 1/(1+\mathcal{K}(\mathrm{X}))$ and $\Pi_{\mathcal{K}}(\mathrm{X}) = 0$ if $\mathcal{K}(\mathrm{X}) = \infty$. Given the latter transformation, the degree of surprise is over $[0, 1]$ instead of $\mathbb{N} \cup \{\infty\}$. When uncertainty in ranking theory is defined over the unit interval, we obtain a generalisation of Bayes, and its increased expressive power affords the modeler with more flexibility than Bayesianism.

4.5 The Quantum Probability Framework

Quantum probability theory is a geometric model of uncertainty. It uses fragments of the language of mathematical probability, but outcomes are distinctively represented as subspaces of varying dimensionality in a multidimensional Hilbert space, which is a vector space used to represent all possible outcomes for questions asked about a system. Unit vectors correspond to possible states of the system and embody knowledge about the states of the system under consideration.

Probabilities of outcomes are determined by projecting the state vector onto different sub-spaces and computing the squared length of the projection. The determination of probabilities is context and order-dependent since individual states can be superposition states and composite systems can be entangled. Thus, while in the Bayesian framework $P(\mathrm{X} \cap \mathrm{Y}) = P(\mathrm{Y} \cap \mathrm{X})$, the commutative property in quantum probability does not always hold. More generally, unlike in Bayesianism, quantum probability does not obey the law of total probability (see Rédei and Summers (2007) for a nice introduction).

Incompatibility in quantum probability theory entails that it is impossible to concurrently assign a truth-value to two hypotheses. Psychologically, two incompatible hypotheses in this sense can be processed only serially because the processing of one hypothesis interferes with the other. Given hypotheses A and B, for example, if A is true at a certain time, then B can be neither true nor false at that time. Conjunctions between incompatible hypotheses are then defined in a sequential way as "A and then B."

One advantage of quantum probability is that it permits explanations of cognitive systems in a superposition of different states. Superposition can give rise to a specific kind of uncertainty that is dependent on the fuzziness and ambiguity of information and that characterises the ambivalence of many of our normal judgments. Additionally, entanglement tracks the interdependencies between different parts of complex cognitive systems. In entangled systems, it is not possible to define a joint probability distribution from the probability distributions of variables corresponding to different constituent parts – changes in one constituent part entails instantaneous changes in another part.

Courtesy of interference, superposition and entanglement, we are able to explain the conjunction fallacy, non-compositional conceptual semantics, order effects in perception, and violations of the sure thing principle (Busemeyer and Bruza 2012).

4.6 Heir to the Throne?

Each of the theories of uncertainty we have briefly reviewed has some epistemic advantage over Bayesianism as well as limitations too. The Dempster-Shafer theory, for instance, has an advantage of representing states of complete ignorance without precise degrees of belief: 0 mass everywhere except for the sure event. Evidence and beliefs are both formalised as Shafer belief functions. Furthermore, combining evidence with Dempster's rule has the desirability of relaxing strong independence assumptions. Upon gathering new evidence, beliefs should be determined by combining the vacuous belief function with the total evidence. Thus, when used to model problems of sensory integration, the D-S approach can generate a measure of conflict between belief functions, which can be used to determine the degree of coherence between distinct sources of sensory information, and to reject unreliable sources. All of this suggests that Bayes might not be the most *unifying* or explanatory theory after all.

A problem for D-S theory, however, is that inference is even more computationally expensive than Bayesian inference. The inefficiency stems from evidence being represented by a belief function that is induced by a mass function on the frame of discernment, instead of a probability distribution over a partition. Combining evidence by Dempster's rule increases computational complexity as the number of possible states increases. In an attempt alleviate the complexity issue, though, Shafer and Tversky (1985) emphasised that "[t]he usefulness of one of these formal languages [i.e., the Bayesian and the D-S language] for a specific problem may depend both on the problem and on the skill of the user... A person may find one language better for one problem and another language better for another" (Shafer and Tversky 1985, p. 311). Despite their theoretical differences, virtues, and drawbacks, D-S and Bayes are not incompatible theories from a formal perspective. Some precise probability distributions are limiting cases of D-S belief functions and some Bayesian conditional probability distributions are limiting cases of applying Dempster's rule. As mentioned earlier, D-S theory is a generalisation of Bayesianism. Thus, any evidence confirming Bayes confirms D-S theory also.

With respect to the IP theory, it clearly has more expressive power than Bayesianism and therefore can capture more uncertainty-involving phenomena. For one, ignorance is better represented in terms of intervals rather than sharp, probability distributions. Another advantage is that the IP framework opens the door to a whole host of decision rules, which have normative implications as well as an ability to explain certain behaviour like ambiguity aversion. Since Bayes fails on both counts, Bayesianism is neither the most *unifying* nor *rational* framework. Of course, IP is not without flaws. For one, it has a computational inefficiency problem involving greater complexity when updating convex sets of probabilities. The theory also encounters trouble when it comes to updating "trivial states of uncertainty" or the non-informative prior, $[0, 1]$, which has prompted some to rule out vacuous priors since they give rise to vacuous posteriors (Walley 1991). But in doing so, the theory becomes restricted to representing partial ignorance and loses the capability of representing complete ignorance. In order to recover such representation, IP theory needs to be extended by a non-Bayesian updating rule. Despite their theoretical differences, virtues, and drawbacks, however, imprecise probability is a generalisation of Bayes. So again, any evidence that confirms Bayes also confirms IP.

The possibility approach has a computational advantage over probability as "maxitivity" makes possibility measures compositional, *viz.* $\Pi(X \cup Y)$ is determined by the maximum of $\Pi(X)$ and $\Pi(Y)$. Minimal computation indicates that possibility theory is at least *simpler* than Bayesianism. Within the larger picture, there are similarities between possibility theory and Dempster-Shafer theory in which a Pl function can be a possibility measure. However, possibility need not be restricted to a D-S interpretation. In general, possibility theory "can be seen either as a coarse, non-numerical version of probability theory, or as a framework for reasoning with extreme probabilities, or yet as a simple approach to reasoning with imprecise probabilities" (Dubois and Prade 2007). The upshot of the possibility approach is its usefulness in assessing vague statements like 'Bob is tall' or 'the shirt is blueish'. In fact, possibility theory has also been used in cognitive science for modelling default and non-monotonic reasoning (e.g., Benferhat, Bonnefon and da Silva Neves 2005). Given its application to vagueness, possibility theory offers cognitive scientists a more *unified* modelling framework for explaining reasoning under uncertainty with fuzzy concepts (Smithson and Verkuilen 2006). But since fuzzy approaches to uncertainty such as possibility theory are not isomorphic to probability theory, it could be suggested that Cox's theorem rules out possibility theory as a *rational* means of quantifying uncertainty (but see Colyvan (2008) on whether probabilism is the only 'coherent' approach to uncertainty).

Ranking theory has an intimate connection to possibility theory. But, distinct from Bayesianism and the other approaches we have considered, proponents point out that ranking theory accommodates the everyday, categorical notion of belief (and disbelief), not just quantitative degrees of belief. On these grounds, they claim that the ranking theoretic approach has advantages over probabilistic approaches because it allows for everything that we can do with quantitative measures and also to tackle traditional problems in epistemology that center around the traditional tripartite concept of belief (Spohn 2012). Ranking theory can then be thought of as more *unifying* than Bayesianism. While ranking functions seem intractable because of the computational expense involved similar to D-S and IP, Goldszmidt and Pearl (1992) have formulated a Spohn-like framework that makes learning actually tractable. More recently, Häming and Peters (2011) show that a ranking function can be used "as a filter on possible actions a reinforcement learning [. . .] agent may take" (p. 226). However, although ranking theory has received some attention, especially in the AI community (e.g., Kern-Isberner and Eichorn 2014), its applications in experimental psychology are currently limited in comparison to Bayesian approaches, but efforts are increasing in fields such as conditional and non-monotonic reasoning (e.g., Skovgaard-Olsen 2016).

Quantum probability theory is uniquely based on axioms that give the theory the advantage of accounting for fuzzy and ambiguous information, but they also allow for an agent to be "Dutch-booked". As we noted, Dutch book arguments do not provide a decisive reason for the superiority of Bayesianism, but they do bear on the rationality of theories of uncertainty. But even though quantum probability theory "is perhaps a framework for bounded rationality and not as rational as in principle possible" (Pothos and Busemeyer 2014, p. 2), courtesy of its unique properties, including superposition, entanglement, incompatibility, and interference, it accommodates empirical results related to order and context effects that are not plausibly explained within a Bayesian framework (Pothos and Busemeyer 2013). Such capability indicates that quantum probability is more *unifying*. An example of this will be detailed in the following subsection.

With the exception of quantum probability, the other approaches are consistent with Bayesianism. However, all the approaches considered posit different theoretical constructs. Because they posit different theoretical constructs, one's epistemic attitude towards the results from any of these modelling approaches will be consequential with respect to our understanding of the nature of mental states and processes. After all, the different theoretical constructs aim at picking out different structures and processes within a cognitive architecture. So, if predictive processers are justified to adopt a realist attitude towards Bayesian posits, then they must show belief in these posits is better supported in terms of non-empirical virtues and empirical adequacy. We have shown the Bayesianism does not possess special non-empirical virtues over the alternatives. In the next subsection, we will focus on the issue of empirical adequacy.

4.7 Empirical Doubts

Observations made thus far undermine the assumption that Bayesianism is the best, on the basis of non-empirical considerations, for representing and explaining uncertainty, and that it is the most rational approach. However, virtually all Bayesian cognitive scientists and naturalistic philosophers of mind interested in Bayes omit the competitors in their work and proceed by taking for granted the superiority of Bayesianism. This behaviour is methodologically problematic, however, and will vitiate debates in philosophy of mind concerning, for instance, representationalism and externalism. In fact, several findings in perception, judgment, and decision making have been shown to be better explained by some of the theories outlined above.

Starting with perceptual effects, the experimental results of Conte et al. (2009) on quantum-like interference effects in the perceptual domain clearly demonstrate the failure of a Bayesian explanation. In their experiment, participants were presented with ambiguous images, which could be perceived in two mutually exclusive ways. One group of participants was presented with a single image A and asked to make a binary choice between $\mathrm{A} = a$ or $\mathrm{A} = q$ on the basis of the way in which they perceived the image at the instance of observation. Another group of participants was presented with two ambiguous images, B and A. After each presentation, participations had to make a binary choice between $\mathrm{B} = b$ and $\mathrm{B} = r$, and between $\mathrm{A} = a$ and $\mathrm{A} = q$.

By the law of total probability, a Bayesian theory predicts that the probability a participant chooses A = a in any of the trials is: $P(\mathrm{A} = a) = P(\mathrm{B} = b)\, P(\mathrm{A} = a|\mathrm{B} = b) + P(\mathrm{B} = r)\, P(\mathrm{A} = a|\mathrm{B} = r)$. The findings of Conte and colleagues were inconsistent with this prediction.[11] However, the results were consistent with a quantum probability prediction that participants' choices would be affected by quantum like interference where the context generated by making the first perceptual choice interfered with the second so that the participants' choices showed order effects, implying non-commutativity. Since such interference effects are ubiquitous in psychology (e.g., Kvam, Pleskac, Yu, and

[11] Bayesians will be quick to point out that context effects can straightforwardly be handled in Bayesian models by adjusting the likelihoods so that they no longer assume independent and identically distributed samples. While this move underwrites the flexibility of Bayes, it also highlights the risk of ad-hocery that may come with this flexibility

Busemeyer 2015), but incompatible with Bayesian predictions, the quantum probability theory better accounts for some psychological phenomena. Although the evidence does not fully vindicate quantum probability, it does undermine the view that Bayesianism is the most empirically adequate explanatory framework for perceptual phenomena.

In the domain of sensorimotor psychology, on which Rescorla (2016) builds his case for intentional realism, some phenomena do not seem to be obviously explained by Bayesianism. Anderson, O'Vari, and Barth (2011), for example, reported forms of visual interpolation generating strong illusory percepts. Because these percepts are highly improbable according to Bayesian models for contour synthesis, the authors describe their findings as "non-Bayesian". Another sensorimotor phenomenon that is not obviously explained by Bayesian models of sensory cue integration is the size-weight illusion. In this illusion, smaller objects are perceived as heavier than larger objects of the same weight, even though the prior expectation is that smaller objects are lighter. As the perception of heaviness discounts this prior expectation, the phenomenon has been characterised as "anti-Bayesian"(Brayanov and Smith 2010). We find that while Bayesian approaches to sensorimotor and perceptual performance have been successful in fitting a variety of phenomena, scarce attention has been paid in both the cognitive sciences and philosophy of mind to the extensive literature documenting non-Bayesian performance in perceptual and sensorimotor tasks (Rahnev and Denison 2018).

Beyond perceptual effects, Bayesianism is not an accurate predictor of human judgment. Anyone who has come across the literature on cognitive biases has likely encountered the "Linda the Bank Teller" study by Tversky & Kahneman (1983). In it, experimental participants were given a short description of "Linda" and asked which is more probable: 'Linda is a bank teller' (B) or 'Linda is a bank teller and is active in the feminist movement' (B $\cap$ F). Bayesianism seems to predict that respondents will say that B is more probable because $P(\mathrm{B} \cap \mathrm{F}) \leq P(\mathrm{B})$. But those familiar with the study will report that Bayesianism is not a very good predictor. For most respondents said (B $\cap$ F). The observed effect has become known as the infamous "conjunction fallacy". For a discussion, see Bovens & Hartmann (2003) and Hartmann & Meijs (2012).

Further empirical evidence against Bayes lies with the assumption that agents are coherent. Since Bayesianism conforms to classical probability, the sum of probabilities over a set of disjoint events is one, and the justifications for why the additivity axiom should be held as a normative standard have been explored above. But as Offerman et al. (2009) demonstrate, ordinary agents sometimes form incoherent distributions leading to what they call *additivity bias*. While this fact was known beforehand, they sought to explain the bias as a result of risk attitudes affecting judgment. After constructing a correction measure for risk attitudes, they found additivity bias to persist. So, additivity bias exists independently of risk attitudes. In their paper, they claim that non-additive models, like IP, are better able to explain the subjects' judgments.

Decision-making is another area where Bayes is not generally an adequate framework for modelling and explaining preferences. Ellsberg's paradox (1961) is an exemplary case where Bayesianism reveals its shortcomings. When presented with two urns, one containing fifty red balls and fifty black balls and the other containing one-hundred red and black balls in unknown proportions, subjects were indifferent to choosing a bet on drawing a red ball to a bet on drawing a black ball from either urn. When presented with bets across urns, e.g. red from the known proportioned urn or red from the unknown proportioned urn, the results were interesting. Bayesianism predicts that an agent would be indifferent, but the results were inconsistent with the prediction. Most prefer red from the known proportioned urn to red from the unknown proportioned urn. The explanation is that ordinary individuals tend to be averse to ambiguity, and their preferences can be better accommodated by IP and MaxiMin Expected Utility (Gilboa & Schmeidler 1989).

The psychological findings we have just reviewed, along with several other behavioural and psychological studies (see e.g., Rahnev and Denison 2018 on perceptual decision-making; Tversky and Koehler 1994 on judgment under uncertainty) limit the degree of empirical adequacy of specific Bayesian models. While these anomalies may be explained within a Bayesian approach by models with alternative priors, likelihood functions, cost functions or decision rules, Bayesian cognitive scientists often lack knowledge of their experimental participants' actual likelihood functions, priors, and cost functions.

At this point, one may draw attention to the fact that, despite these anomalies and uncertainties, Bayesian cognitive scientists, and predictive processors alike, make suggestions about the possible neural implementation of Bayesian posits (Friston 2009; Pouget et al. 2013; Ma and Jazayeri 2014); instead, advocates of the alternative approaches do not. So, one may conclude that, unlike its alternatives, Bayesianism is more empirically fruitful, as it spans all Marr's levels.

This conclusion is too quick, however. In fact, concrete suggestions have also been made about how algorithms based on D–S theory of evidence might be implemented in neurorobotic architectures for integrating different sources of sensory information (e.g., Murphy 1996). For quantum probability, some have gone so far as to propose not only that the brain directly implements quantum computations, but also that quantum computation might illuminate how the brain produces consciousness (e.g., Hameroff 2007). While we are not aware of any application at the neural level of algorithms based on imprecise probability, possibility theory or ranking functions, we don't see any reason why these approaches cannot be brought to bear on questions about how brains handle uncertainty. It is also important to point out that, though Bayesian cognitive scientists are considering various hypotheses about how brains might realise the entities and processes Bayesian models posit, the evidence in favour of any particular implementation of Bayesian inference is currently inconclusive.

In the face of the evidential uncertainties and of the empirical anomalies we've highlighted, it is surprising that several naturalistic philosophers of mind are willing to make blanket assertions such as the mind or brain is Bayesian. For these assertions to be convincing, a better job needs to be done in addressing Bayes' failures, and consequently the empirical success of alternative approaches.

4.8 A Bayesian Argument

In closing this section, we would like to give a Bayesian argument to the effect that the cognitive science, and especially the philosophical communities should not adopt a realistic attitude towards Bayes too quickly. To do so, we consider a theory H which accounts for the evidence E. Introducing binary propositional variables H and E with the values H ("The hypothesis is true") and ¬H ("The hypothesis is false') and E ("The evidence obtains") and ¬E ("The evidence does not obtain"), the Bayesian Network depicted in Figure 1 describes the probabilistic relation between H and E.[12]

We assume that H entails E, i.e. that the evidence is a deductive consequence of the

[12] See Bovens and Hartmann (2003) for an introduction to the application of Bayesian Network methods in epistemology and philosophy of science.

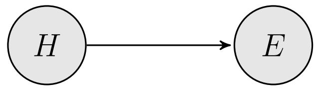

Figure 1: The Bayesian Network with the variables H and E.

theory. Hence, we set

$$P(\mathrm{H}) = h \quad , \quad P(\mathrm{E}|\mathrm{H}) = 1 \quad , \quad P(\mathrm{E}|\neg\mathrm{H}) = \alpha \tag{1}$$

with $0 < \alpha < 1$. It is then easy to show that the posterior probability of H, i.e. the probability of H after learning that E is true, is given by[13]

$$P(\mathrm{H}|\mathrm{E}) = \frac{h}{h + \alpha\, \bar{h}} > h \tag{2}$$

Hence, E confirms H. But how much? As one sees from eq. (2), $P(\mathrm{H}|\mathrm{E})$ is a decreasing function of α. For $\alpha \to 0$, i.e. if we consider it to be impossible that an alternative theory (which is contained in the "catch-all" $\neg$H) accounts for the evidence, then $P(\mathrm{H}|\mathrm{E}) \to 1$. For $\alpha \to 1$, i.e. if we are convinced that an alternative theory accounts for the evidence, then $P(\mathrm{H}|\mathrm{E}) \to h$. Hence, if we consider it quite likely that an alternative theory accounts for the evidence, i.e. if we set $\alpha \approx 1$, then $P(\mathrm{H}|\mathrm{E}) \approx h$ and we won't get much confirmation for H after observing E.

This situation changes is we consider several independent pieces of evidence $E_1, \ldots, E_n$. Assuming $E_i \perp\!\!\!\perp E_j | H$ for $i \neq j = 1, \ldots, n$ and setting $P(\mathrm{H}) = h$, $P(\mathrm{E_i}|\mathrm{H}) = 1$ and $P(\mathrm{E_i}|\neg\mathrm{H}) = \alpha$ for $i = 1, \ldots, n$, we obtain

$$P(\mathrm{H}|\mathrm{E_1}, \ldots, \mathrm{E_n}) = \frac{h}{h + \alpha_{eff}\, \bar{h}} \,. \tag{3}$$

For large n, $\alpha_{eff} := \alpha^n \approx 0$ and hence $P(\mathrm{H}|\mathrm{E}) \approx 1$. Given that Bayesian cognitive science accounts for many different phenomena, this seems to justify taking it very seriously. Note, however, that we made two important assumptions. First, we assumed that the different pieces of evidence $\mathrm{E_1}, \ldots, \mathrm{E_n}$ are independent (given H). This is controversial and needs to be justified on a case by case basis. Second, we assumed that all pieces of evidence are a deductive consequence of H, i.e. we assumed that $P(\mathrm{E_i}|\mathrm{H}) = 1$ for all $i = 1, \ldots, n$. This is controversial as H may make E_i only highly likely, and so we should set $P(\mathrm{E_i}|\mathrm{H})$ to a value smaller than 1. Assigning a value smaller than 1 to $P(\mathrm{E_i}|\mathrm{H})$ is also supported by the observation that E_i typically does not follow from H alone, but from H and some additional auxiliary assumptions. (This is the famous Duhem Quine Problem.) Hence, $P(\mathrm{E_i}|\mathrm{H}) < 1$ (for all $i = 1, \ldots, n$) which effectively lowers the posterior probability $P(\mathrm{H}|\mathrm{E_1}, \ldots, \mathrm{E_n})$.

To accept H, we would also like to make sure that the posterior probability of H is fairly high. As eq. (2) shows, the value of $P(\mathrm{H}|\mathrm{E})$ also depends on the prior probability of H (i.e. on h) and neglecting it would mean to commit the base-rate fallacy.

So let us now explore what we can say about the prior probability of H. We will argue that it depends on our beliefs about the existence of alternative theories that explain the evidence. To proceed with our analysis, we additionally introduce the binary

[13]Here and throughout we use the abbreviation $\bar{x} := 1 - x$.

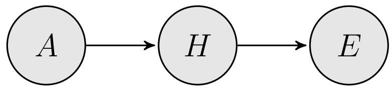

Figure 2: The Bayesian Network with the variables H, E and A.

propositional variable A with the values A := "There is an alternative explanation for E" and ¬A accordingly and study the Bayesian Network depicted in Figure 2. We set

$$P(\mathrm{A}) = a \quad , \quad P(\mathrm{H}|\mathrm{A}) = \beta \tag{4}$$

with $0 < a, \beta < 1$. β will be large if we believe that H is part of a better explanation (provided there is one) or if we believe that there can be multiple equally acceptable explanations for E. β will be small if we believe that an alternative explanation will be better and eventually replace H. β will also be small if one beliefs that either H or some alternative is right and that there can be only one explanation for E. Hence, if there is an alternative, this alternative might well be the true theory and hence one assigns a small value to β. Given that there are several more or less unexplored alternative theoretical frameworks in cognitive science (as argued above), it seems rational to assign a rather low value to the parameter β.

We also set

$$P(\mathrm{H}|\neg\mathrm{A}) = 1 \tag{5}$$

as there must be (or so we assume) an explanation for E. If A is false and there is no alternative explanation for E, then H has to be true.

With this, we calculate

$$P(\mathrm{H}) = h = a\beta + \bar{a}. \tag{6}$$

Hence, if one has good reasons to believe that a is fairly large (i.e. if, as in the case of Bayesian cognitive science, alternatives to H are known and we can assume that they provide alternative explanations of E) and if β is fairly small (as we argued), then the "prior" h is relatively small and hence the posterior probability $P(\mathrm{H}|\mathrm{E})$ is relatively small.

To sum up, we have given a Bayesian argument to the effect that we must be very careful and not accept too quickly claims that brains are Bayesian mechanisms or that we should adopt a realist stance towards Bayesian models in general.

5 Conclusion: Against Bayesian Realism

If there is good reason to doubt that, currently, the Bayesian approach provides us with the best explanations of many cognitive phenomena, then there is good reason to remain agnostic about the truth of Bayesian models of cognitive phenomena and behaviour, contrary to what has been claimed in the philosophical literature (e.g., Hohwy 2013; Clark 2016; Rescorla 2016), as well as in fragments of the cognitive science literature (e.g., Knill and Pouget 2004; Ma et al. 2006; Friston 2009).

Facts about the institutional organisation of contemporary scientific inquiry bolster this agnosticism, providing us with some explanation of why alternatives to Bayesianism have been neglected. As pointed out by Stanford (2015), the institutional apparatus of contemporary scientific inquiry has "served to reduce not only the incentives but also the freedom scientists have to pursue research that challenges existing theoretical orthodoxy or seeks to develop fundamental theoretical innovation." While this conservatism has fostered specialisation in the sciences, it has also shielded reputable theories and frameworks from comparison with relevant, under-considered alternatives. Jettisoning conservatism by reversing the neglect of available alternatives, their relative non-empirical virtues, and their relative empirical advantages over Bayes would pose a serious challenge for a realist stance towards the results of Bayesian cognitive science.

But as things currently stand, Bayesianism remains the most popular approach for representing and managing uncertainty. The tools which a Bayesian cognitive scientist is currently afforded to address problems of uncertain inference are more abundant in comparison to alternatives, and continue to be refined in neighbouring fields. By comparison, Bayesian approaches prevail over the Dempster-Shafer theory, IP, possibility theory, ranking theory, and quantum probability theory in disciplines ranging from statistics to machine learning, AI, and economics (Poirier 2006). And the popularity of Bayesian modelling continues to grow in the cognitive sciences, too, as evidenced by an increase in the number of articles, conference papers, and workshops dedicated to Bayesian modelling of cognition and its foundations (Kwisthout et al. 2011, note 1).

Despite the fact that Bayesianism does not enjoy special epistemic virtues in comparison to alternatives, the choice to employ Bayesian method by cognitive scientists may be explained in terms sociological factors connected with the reward structure of scientific institutions, which is biased towards conservatism (Stanford 2015). These sociological factors may have led researchers to approach their research questions with a Bayesian eye that created a neglect of alternatives. Seeing that more and more researchers have addressed their questions within the Bayesian framework, a division of cognitive labour has been fostered in the fields of cognitive science, epistemology, and philosophy of mind. Sophisticated tools have been developed and exploited to approach problems at a higher level of specialisation in both machine learning and human cognition (Gershman, Horvitz and Tenenbaum 2015). But if this higher degree of specialisation continues within research communities with an incentive structure that strongly favours conservatism, then exploring and developing novel or alternative theoretical frameworks will ultimately face a much more difficult path. As this will impact the trajectory of cognitive science, we believe – like Gigerenzer (1991) – that it is important to take a step back and evaluate whether the net result is the best way to advance our understanding of how minds work.

Upon conceding to the point that the choice of the Bayesian approach in cognitive science is currently based on pragmatic and sociological considerations instead of epistemic considerations, and if there are plausible, actually conceived alternatives to Bayes that are systematically neglected, then naturalistic philosophers of mind ought to acknowledge that scientific realism is not the appropriate epistemic attitude towards the outputs of Bayesian cognitive science. If this is correct, then naturalistic theories of mind such as Hohwy's (2013) or Clark's (2015) predictive processing theory, which build on results from Bayesian cognitive science, cannot make the substantial empirical claims that brains are Bayesian mechanisms, that mental states are fundamentally "predictions" or that mental activity is fundamentally prediction-error minimisation. These theories, and the debates they generated in the philosophy of mind should instead reflect an instrumental agnosticism about whether or not minds are "really" Bayesian.

An endorsement of anti-realism towards Bayes should re-configure our understanding of debates about predictive processing in a number of ways. In particular, controversies concerning the nature of mental representations within predictive processing should all be re-interpreted in a fictionalist way (cf., Sprevak 2013). According to the kind of fictionalism we have in mind, we should be agnostic concerning claims about representations within predictive processing debates; these claims do not aim at truth, but pretending they aim at truth is worthwhile for various purposes. One obvious purpose of fiction is aesthetic appeal. Perhaps, like Prinz (forthcoming) suggests, the predictive processing theory is, in fact, best understood not as aiming at making substantial empirical claims about mind, but at providing us with a grand unifying picture of mind and its place in the world. Because this picture invites us to see mind as a pro-active predictive engine instead of a passive feed-forward, input-dominated machine, we may well find it rewarding. From this perspective, a grand unifying theory like predictive processing is a useful tool not only for summarising recent research findings from various research programs – including Bayesian cognitive science, reinforcement learning and deep learning – where the notion of "prediction" is centerpiece in one way or another, but also for highlighting the relevance of such notions as "uncertainty", "prediction" and "prediction error minimisation" to our image of humanity. So, one reason why the predictive processing theory has been so attractive for philosophers of mind is really its aesthetic appeal.

References

[1] Anderson, B.L., O'Vari, J., & Barth, H. (2011). Non-bayesian contour synthesis. *Current Biology, 21*(6), 492–496.

[2] Augustin, T., Coolen, F., de Cooman, G. & Troffaes, M.C.M. (Eds.) (2014). *Introduction to Imprecise Probabilities*. Somerset: Wiley.

[3] Benferhat, S., Bonnefon, J.F., & da Silva Neves, R. (2005). An overview of possibilistic handling of default reasoning, with experimental studies. *Synthese, 146*(1-2), 53–70.

[4] Bovens, L. & Hartmann, S. (2003). *Bayesian Epistemology*. Oxford: Oxford University Press.

[5] Bowers, J.S. & Davis, C.J. (2012a). Bayesian just-so stories in psychology and neuroscience. *Psychological Bulletin, 138,* 389–414.

[6] Bowers, J.S. & Davis, C.J. (2012b). Is that what Bayesians believe? Reply to Griffiths, Chater, Norris, and Pouget (2012). *Psychological Bulletin, 138,* 423–426.

[7] Brayanov, J.B. & Smith, M.A. (2010). Bayesian and anti-Bayesian biases in sensory integration for action and perception in the size-weight illusion. *Journal of Neurophysiology, 103*(3), 1518–1531.

[8] Busemeyer, J. R. & Bruza, P.D. (2012). *Quantum Models of Cognition and Decision*. Cambridge: Cambridge University Press.

[9] Carr, J. R. (2017). Epistemic utility theory and the aim of belief. *Philosophy and Phenomenological Research, 95*(3), 511-534.

[10] Chater, N., Goodman, N.D., Griffiths, T. L., Kemp, C., Oaksford, M., & Tenenbaum, J.B. (2011). The imaginary fundamentalists: The unshocking truth about Bayesian cognitive science. *behavioural and Brain Sciences, 34,* 194–196.

[11] Chater, N., Tenenbaum, J.B. & Yuille, A. (2006). Probabilistic models of cognition: conceptual foundations. *Trends in Cognitive Sciences, 10(7),* 287–291.

[12] Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *behavioural and Brain Sciences, 36*(3), 181–204.

[13] Clark, A. (2016). *Surfing Uncertainty: Prediction, Action, and the Embodied Mind.* Oxford: Oxford University Press.

[14] Colombo, M. (2015). For a few neurons more: Tractability and neurally-informed economic modelling. *The British Journal for Philosophy of Science, 66*, 713-736.

[15] Colombo, M. & Hartmann, S. (2017). Bayesian cognitive science, unification, and explanation. *The British Journal for the Philosophy of Science,* axv036.

[16] Colombo, M. & Seriès, P. (2012). Bayes in the brain–on Bayesian modelling in neuroscience. *The British Journal for the Philosophy of Science, 63(3),* 697–723.

[17] Colyvan, M. (2008). Is probability the only coherent approach to uncertainty? *Risk Analysis, 28,* 645–52.

[18] Conte, E., Khrennikov, A.Y., Todarello, O., Federici, A., Mendolicchio, L. & Zbilut, J.P. (2009). Mental states follow quantum mechanics during perception and cognition of ambiguous figures. *Open Systems and Information Dynamics, 16,* 1–17.

[19] Cox, R.T. (1946). Probability, frequency, and reasonable expectation. *American Journal of Physics, 14,* 1–13.

[20] Danks, D. (2014). *Unifying the Mind: Cognitive Representations as Graphical Models.* MIT Press.

[21] Diaconis, P. & Zabell, S. (1982). Updating subjective probability. *Journal of the American Statistical Association, 77,* 822–830.

[22] Dempster, A. (1968). A generalisation of bayesian inference. *Journal of the Royal Statistical Society, 30,* 205–247.

[23] Drugowitsch, J. & Pouget, A. (2012). Probabilistic vs. non-probabilistic approaches to the neurobiology of perceptual decision-making. *Current Opinion in Neurobiology, 22,* 963–969.

[24] Dubois, D. & Prade, H. (2007). Possibility theory. *Scholarpedia, 2,* 2074.

[25] Eberhardt, F. & Danks, D. (2011). Confirmation in the cognitive sciences: the problematic case of Bayesian models. *Mind and Machines, 21,* 389–410.

[26] Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *The Quarterly Journal of Economics, 75,* 643–669.

[27] Eva, B. & Hartmann, S. (2018). Bayesian argumentation and the value of logical validity. To appear in *Psychological Review.*

[28] Faisal, A.A., Selen, L.P. J. & Wolpert, D.M. (2008). Noise in the nervous system. *Nature Reviews Neuroscience, 9,* 292–303.

[29] Friston, K. (2009). The free-energy principle: A rough guide to the brain? *Trends in Cognitive Sciences, 13,* 293–301.

[30] Gelman, A. (2008). Objections to Bayesian statistics. *Bayesian Analysis, 3*(3), 445-449.

[31] Gershman, S.J., Horvitz, E. J. & Tenenbaum, J.B. (2015). Computational rationality: A converging paradigm for intelligence in brains, minds, and machines. *Science, 349,* 273–278.

[32] Ghahramani, Z. (2015). Probabilistic machine learning and artificial intelligence. *Nature, 521,* 452–459.

[33] Gigerenzer, G. (1991). From tools to theories: a heuristic of discovery in cognitive psychology. *Psychological Review, 98,* 254–267.

[34] Gilboa, I. & Schmeidler, D. (1989). Maximin expected utility with non-unique prior. *Journal of Mathematical Economics, 18,* 141–153.

[35] Goldszmidt, M. & Pearl, J. (1992). Reasoning with qualitative probabilities can be tractable. In D. Dubois, M.P. Wellman, B. D'Ambrosio, & P. Smets (Eds.), *Uncertainty in Artificial Intelligence* (pp. 112–120).

[36] Griffiths, T.L., Chater, N., Norris, D. & Pouget, A. (2012). How the Bayesians got their beliefs (and what those beliefs actually are): Comments on Bower and Davis (2012). *Psychological Bulletin, 138,* 415–422.

[37] Griffiths, T.L., Vul, E., & Sanborn, A.N. (2012). Bridging levels of analysis for probabilistic models of cognition. *Current Directions in Psychological Science, 21,* 263–268.

[38] Hájek, A. (2008). Dutch book arguments. *The Handbook of Rational and Social Choice.* Oxford: Oxford University Press. 173–196.

[39] Halpern, J. (2003). *Reasoning About Uncertainty.* Cambridge: MIT Press.

[40] Häming, K. & Peters, G. (2011). Ranking functions in large state spaces. In *Artificial Intelligence Applications and Innovations.* Berlin: Springer, 219–228.

[41] Hameroff, S. R. (2007). The brain is both neurocomputer and quantum computer. *Cognitive Science, 31*(6), 1035-1045.

[42] Hartmann, S. & Meijs, W. (2012). Walter the banker: the conjunction fallacy reconsidered. *Synthese, 184,* 73–87.

[43] Hohwy, J. (2013). *The Predictive Mind.* Oxford: Oxford University Press.

[44] Hohwy, J. (2015). The neural organ explains the mind. In T. Metzinger & J. M. Windt (Eds). Open MIND: 19(T). Frankfurt am Main: MIND Group. doi: 10.15502/9783958570016.

[45] Hohwy, J. (2016). The self evidencing brain. *Noûs, 50*(2), 259–285.

[46] Huber, F. (2014). Formal representations of belief. *The Stanford Encyclopedia of Philosophy,* Edward N. Zalta (ed.), <http://plato.stanford.edu/archives/spr2014/entries/formal-belief/>.

[47] Icard, T.F. (2018). Bayes, bounds, and rational analysis. *Philosophy of Science, 85*(1), 79–101.

[48] Jones, M.I. & Love, B.C. (2011). Bayesian fundamentalism or enlightenment? On the explanatory status and theoretical contributions of Bayesian models of cognition. *behavioural and Brain Sciences, 34,* 169–188.

[49] Joyce, J. (1998). A nonpragmatic vindication of probabilism. *Philosophy of Science, 65,* 573–603.

[50] Kern-Isberner, G. & Eichhorn, C. (2014). Structural inference from conditional knowledge bases. *Studia Logica, 102,* 751–769.

[51] Kvam, P. D., Pleskac, T. J., Yu, S. & Busemeyer, J. R. (2015). Interference effects of choice on confidence: quantum characteristics of evidence accumulation. *Proceedings of the National Academy of Sciences, 112,* 10645–10650.

[52] Knill, D.C. & Pouget, A. (2004). The Bayesian brain: The role of uncertainty in neural coding and computation. *Trends in Neurosciences, 27,* 712–719.

[53] Knill, D.C. & Richards, W. (1996). *Perception as Bayesian Inference*. Cambridge: Cambridge University Press.

[54] Kwisthout, J., Wareham, T. & van Rooij, I. (2011). Bayesian intractability is not an ailment that approximation can cure. *Cognitive Science, 35,* 779–784.

[55] Leitgeb, H. & Pettigrew, R. (2010a). An objective justification of Bayesianism I: Measuring inaccuracy. *Philosophy of Science, 77,* 201–235.

[56] Leitgeb, H. & Pettigrew, R. (2010b). An objective justification of Bayesianism II: The consequences of minimising inaccuracy. *Philosophy of Science, 77,* 236–272.

[57] Levi, I. (1974). On indeterminate probabilities. *The Journal of Philosophy, 71,* 391–418.

[58] Littlejohn, C. (2015). Who cares what you accurately believe?. *Philosophical Perspectives, 29*(1), 217–248.

[59] Ma, W.J. & Jazayeri, M. (2014). Neural coding of uncertainty and probability. *Annual Review of Neuroscience, 37,* 205–220.

[60] Ma, W.J., Beck, J.M., Latham, P.E. & Pouget, A. (2006). Bayesian inference with probabilistic population codes. *Nature Neuroscience, 9,* 1432–1438.

[61] Marr, D. & Poggio, T. (1977). From understanding computation to understanding neural circuitry. *Neurosciences Research Progress Bulletin, 15*, 470–488.

[62] Metzinger, T. & Wiese, W. (Eds.) (2017). Philosophy and predictive processing. Frankfurt am Main: MIND Group. doi: 10.15502/9783958573024.

[63] Marcus, G.F. & Davis, E. (2013). How robust are probabilistic models of higher-level cognition? *Psychological Science, 24,* 2351–2360.

[64] Murphy, R. R. (1996). Biological and cognitive foundations of intelligent sensor fusion. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, 26*(1), 42-51.

[65] Oaksford, M. & Chater, N. (2007). *Bayesian Rationality: The Probabilistic Approach to Human Reasoning.* Oxford: Oxford University Press.

[66] Offerman, T., Sonnemans, J., Van de Kuilen, G. & Wakker, P. P. (2009). A truth serum for non-bayesians: Correcting proper scoring rules for risk attitudes. *The Review of Economic Studies, 76*(4), 1461–1489.

[67] Orbán, G. & Wolpert, D.M. (2011). Representations of uncertainty in sensorimotor control. *Current Opinion in Neurobiology, 21,* 1–7.

[68] Poirier, D.J. (2006). The growth of Bayesian methods in statistics and economics since 1970. *Bayesian Analysis, 1,* 969–980.

[69] Pothos, E. M. & Busemeyer, J. R. (2013). Can quantum probability provide a new direction for cognitive modelling? *behavioural and Brain Sciences, 36,* 255–327.

[70] Pouget, A., Beck, J.M., Ma W.J. & Latham, P.E. (2013). Probabilistic brains: knowns and unknowns. *Nature Neuroscience, 16,* 1170–1178.

[71] Prinz, J. (forthcoming). Ways of mindmaking. In M. Colombo, L. Irvine, & M. Stapleton (Eds.), *Andy Clark and His Critics.* Oxford: Oxford University Press.

[72] Rao, R.P.N., Olshausen, B.A. & Lewicki, M.S. (2002). *Probabilistic Models of the Brain: Perception and Neural Function.* Cambridge: MIT Press.

[73] Rahnev, D. (2017). The case against full probability distributions in perceptual decision making. bioRxiv, 108944.

[74] Rahnev, D., & Denison, R. (2018). Suboptimality in perception. bioRxiv, 060194.

[75] Rédei, M. & Summers, S.J. (2007). Quantum probability theory. *Studies in the History and Philosophy of Modern Physics, 38,* 390–417.

[76] Rescorla, M. (2015). Bayesian perceptual psychology. in Mohan Matthen (ed.), *The Oxford Handbook of the Philosophy of Perception.* Oxford: Oxford University Press.

[77] Rescorla, M. (2016). Bayesian sensorimotor psychology. *Mind and Language, 31,* 3–36.

[78] Sanborn, A.N., Griffiths, T. L. & Navarro, D.J. (2010). Rational approximations to rational models: alternative algorithms for category learning. *Psychological Review, 117*, 1144.

[79] Shafer, G. (1992). The Dempster-Shafer theory. *Encyclopedia of Artificial Intelligence, 2nd edition,* New York: Wiley, 330–331.

[80] Shafer, G. & Tversky, A. (1985). Languages and designs for probability judgment. *Cognitive Science, 9,* 309–339.

[81] Skovgaard-Olsen, N. (2016). Ranking theory and conditional reasoning. *Cognitive Science, 40,* 848–880.

[82] Smithson, M. & Verkuilen, J. (2006). *Fuzzy Set Theory: Applications in the Social Sciences*. Thousand Oaks: Sage.

[83] Sober, E. (2003). Two uses of unification. in F. Stadler (ed.), *The Vienna Circle and Logical Empiricism – Vienna Circle Institute Yearbook 2002*. Dordrecht: Kluwer, 205–216.

[84] Spohn, W. (2012). *The Laws of Belief: Ranking Theory and its Philosophical Applications*. Oxford: Oxford University Press.

[85] Sprevak, M. (2013). Fictionalism about neural representations. *The Monist, 96*(4), 539-560.

[86] Stanford, P.K. (2015). Unconceived alternatives and conservatism in science: the impact of professionalization, peer-review, and big science. *Synthese,* 1–18.

[87] Vilares, I. & Kording, K. (2011). Bayesian models: the structure of the world, uncertainty, behaviour, and the brain. *Annals of the New York Academy of Sciences, 1224,* 22–39.

[88] Tenenbaum, J.B., Kemp, C., Griffiths, T.L. & Goodman, N.D. (2011). How to grow a mind: statistics, structure and abstraction. *Science, 331,* 1279–1285.

[89] Troffaes, M.C.M & Hable. R. (2014). Computation. In T. Augustin, F. Coolen, G. de Cooman, M.C.M. Troffaes (Eds.), *Introduction to Imprecise Probabilities* (pp. 329-337). Somerset: Wiley.

[90] Tversky, A. & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review, 90* (4), 293–315.

[91] Walley, P. (1991). *Statistical Reasoning with Imprecise Probabilities*. London: Chapman & Hall.

[92] Wilson N. (2000). Algorithms for Dempster-Shafer theory. In J. Kohlas and S. Moral (Eds.), *Handbook of Defeasible Reasoning and Uncertainty Management Systems*, vol 5. Springer: Dordrecht.

[93] Zadeh, L. (1975). Fuzzy logic and approximate reasoning. *Synthese, 30,* 407–428.

[94] Zednik, C. & Jäkel, F. (2016). Bayesian reverse-engineering considered as a research strategy for cognitive science. *Synthese, 193*(12), 3951-3985.

[95] Zhao, J. & Osherson, D. (2010). Updating beliefs in light of uncertain evidence: Descriptive assessment of Jeffrey's rule. *Thinking & Reasoning, 16,* 288–307.