Sleep and Dreaming in the Predictive Processing Framework

Alessio Bucci & Matteo Grasso

Sleep and dreaming are important daily phenomena that are receiving growing attention from both the scientific and the philosophical communities. The increasingly popular predictive brain framework within cognitive science aims to give a full account of all aspects of cognition. The aim of this paper is to critically assess the theoretical advantages of Predictive Processing (PP, as proposed by Clark 2013, Clark 2016; and Hohwy 2013) in defining sleep and dreaming.

After a brief introduction, we overview the state of the art at the intersection between dream research and PP (with particular reference to Hobson and Friston 2012; Hobson et al. 2014). In the following sections we focus on two theoretically promising aspects of the research program.

First, we consider the explanations of phenomenal consciousness during sleep (i.e. dreaming) and how it arises from the neural work of the brain. PP provides a good picture of the peculiarity of dreaming but it can't fully address the problem of how consciousness comes to be in the first place. We propose that Integrated Information Theory (IIT) (Oizumi et al. 2014; Tononi et al. 2016) is a good candidate for this role and we will show its advantages and points of contact with PP.

After introducing IIT, we deal with the evolutionary function of sleeping and dreaming. We illustrate that PP fits with contemporary researches on the important adaptive function of sleep and we discuss why IIT can account for sleep mentation (i.e. dreaming) in evolutionary terms (Albantakis et al. 2014).

In the final section, we discuss two future avenues for dream research that can fruitfully adopt the perspective offered by PP: (a) the role of bodily predictions in the constitution of the sleeping brain activity and the dreaming experience, and (b) the precise role of the difference stages of sleep (REM (Rapid eye movement), NREM (Non-rapid eye movement)) in the constitution and refinement of the predictive machinery.

Keywords

Bayesian brain | Consciousness | Dreaming | Embodiment | Evolution of sleep and dreaming | Hard problem | Integrated information theory | Predictive processing | Sleep | Synaptic pruning

Acknowledgements

The authors wish to thank Jennifer Windt and two anonymous reviewers for the precious criticisms and suggestions and Joe Dewhurst, Marco Viola, Benjamin Singer and Emma Hemmings for useful comments on an early draft of this paper. Special thanks also go to Thomas Metzinger, Wanja Wiese and the participants of the MIND23 workshop in Frankfurt for the invaluable feedback and critical discussion of our work.

1 Introduction

Dreaming is a fundamental aspect of our mental activity. We spend almost one-third of our life sleeping, and a good portion of that time dreaming. We often report dreaming experiences upon awakening, and they can have a huge impact on our daily life. Dreaming has also been an object of philosophical investigations, representing a conundrum for our theories on the nature of reality and for the accuracy and reliability of perception.

In spite of this, not so much has been written about dreaming in either the philosophical or the scientific literature until the 20th century. Since the discovery of REM sleep in the early 1950s (Aserinsky and Kleitman 1953), there has been in fact an increasing interest for the topic in psychology and philosophy. The research on REM sleep has sparked a proliferation of theories about dreaming, although it is now fairly established that dreams can happen at any stage of sleep (Nielsen 2000).

On the scientific side, authors are divided in their theoretical proposals on the topic, in particular regarding the *explananda* of their research. A first group of theories concerns the neural underpinning of dreaming: such as the famous AIM model (Hobson et al. 2000), the neuropsychoanalysis of dreaming (Solms 2000), and more cognitive-functional approaches (Domhoff 2001). A second, although not entirely distinct, set of theories concerns the functional role of sleeping and dreaming:

Bucci A. & Grasso M. (2017). Sleep and Dreaming in the Predictive Processing Framework.
In T. Metzinger & W. Wiese (Eds.). *Philosophy and Predictive Processing:* 6. Frankfurt am Main: MIND Group. doi: 10.15502/9783958573079

from the early proposal of a "reverse learning" theory (Crick and Mitchison 1983) to the more updated "synaptic pruning" hypothesis (Tononi and Cirelli 2014), and the complementary idea of sleep as memory consolidation (Stickgold et al. 2001; Perogamvros and Schwartz 2012). Many researchers are also concerned with finding the specific evolutionary role of dreaming, rather than of sleep in general. Interesting proposals have been formulated over the last decades, such as the "threat simulation theory" (Revonsuo 2000; Valli and Revonsuo 2009) and its more recent version "social simulation theory" (Revonsuo et al. 2015). Generally speaking though, the idea of an evolutionary advantage of dreaming per se has been received with scepticism (Flanagan 1995; Flanagan 2000). Finally, another approach is focused on the developmental and cognitive aspects of dreaming: the analysis of dream reports highlighted a progressive enrichment in structure and length of children's dreams (Foulkes 1999) as well as continuities between dream content and wakeful activities (Foulkes 1985; Domhoff 2011a).

This plethora of proposals highlights the difficulty in formulating a cohesive theory of dreaming. This is reflected by the early philosophical scepticism in regards to the topic (famously expressed by authors such as Malcolm 1959 and Dennett 1976). A recent comprehensive analysis of the field of dream research (Windt and Noreika 2011) pointed out the so-called "integration problem: the problem of how to integrate dreaming into broader theories of consciousness" (Windt and Noreika 2011, p. 1091) and, more generally speaking, of cognition.

Taking the integration problem as a springboard, we aim to give an account of dreaming through the lens of Predictive Processing (as proposed by Clark 2013, Clark 2016; Hohwy 2013) in order to provide a more comprehensive alternative to the present theoretical fragmentation. The paper will provide an overview of how sleep and dreaming are explained within the framework, focusing on pertaining issues and their possible solutions.

In order to do that, in the first section we will briefly recapitulate the main tenets of Predictive Processing and illustrate the state of the art of this approach in dream research. We will give a definition of dreaming in Predictive Processing terms and highlight the main theoretical advantages offered by the framework.

In the second section, we will illustrate these theoretical advantages while tackling a weak spot in Predictive Processing: the explanation of phenomenal consciousness, with specific reference to the hard problem of consciousness (as formulated by Chalmers 1996). We will introduce Integrated Information Theory (hereafter IIT, Oizumi et al. 2014; Tononi et al. 2016) as an example of a theoretical proposal that deals with the correlations between the phenomenal aspects of dreaming and the neural work of the sleeping brain. We will show that Predictive Processing and Integrated Information Theory share important analogies and theoretical points of contact, and therefore could be natural allies in providing a more detailed picture of how and why we dream. Afterwards, we will examine in more detail the evolutionary function of sleep according to Predictive Processing and Integrated Information Theory and show why the latter identifies an evolutionary role for dreaming too.

In the final section, we will focus on two topics in dream research that can be promisingly investigated within the Predictive Processing framework in future studies. First, the role of the body in regards to dream formation, which can be accounted for by the predictive architecture proposed by Predictive Processing. Second, the specific role of the different sleep stages (REM and NREM) in the refinement of the predictive machinery: in order to resolve some ambiguity presented by the scientific literature, we will propose a two-step mechanism of refinement operating during sleep.

2 Dreaming in the Predictive Brain

2.1 What is Predictive Processing?

Predictive Processing (hereafter PP; see Clark 2013, Clark 2016; Hohwy 2013) is an emerging framework in cognitive science, rooted in a vast and diverse scientific and philosophical literature (for a summary, see Friston 2010; see also Clark 2013, pp. 181-186, for the historical antecedents). The main tenet of the framework is that brains are predictive machines with a hierarchical structure, continuously in the business of predicting their own internal states in relation to the external sensory input. This result is achieved through a combination of top-down flows of predictions and bottom-up flows of error signals. Predictions here are based on hypotheses construed on a generative model[1] of the world, which tracks by means of Bayesian inferences the worldly causes behind the sensory input (for a detailed account of Bayesian statistics in this context see Hohwy 2013, ch. 1; see also Clark 2016, pp. 301-303). This inferential process creates expectations (or priors) which guide prediction at each level of the cognitive hierarchy.

The hierarchical structure is the following: priors are organised from bottom levels (which track fast time-scale, perceptual details) to top levels (which track slow time-scale, abstract regularities) (Hohwy 2013, pp. 27-28). Predictions are streamed top-down (and laterally) and matched with the bottom-up sensory information. That first matching generates an amount of prediction error that indicates how much the current prediction differs from the input. The prediction error is then streamed upward (forward) in the architecture, repeating the matching process at each level, through the mutual informational exchange between error units and representation units (the latter being the carrier for the top-down predictions).

The goal of the whole system is to minimize the amount of prediction error (and the overall level of surprisal to the system), i.e. to generate successful predictions of its own states, ultimately corresponding to successful inferences about the world (Clark 2013, p. 186; see also Hohwy 2013, pp. 51-53).

According to PP, therefore, brains are sophisticated neural networks that rely on statistical inferences to produce the best prediction of the incoming sensory input and of their own internal states. The uncertainty of the signal from the environment (its reliability) is handled through a mechanism of assignment of expected precision to incoming input and gain regulation of the error units (Hohwy 2013, pp. 64-66; Clark 2016, pp. 53-59). In other words, when the precision of the signal coming from the sensory input is judged as low, the gain on error units is also low and the brain relies more on its previously acquired priors. In other cases in which the signal is considered more precise, the gain is high and the brain relies more on the inputs. This process, in a nutshell, recapitulates the role of attention in the hierarchical architectures described by PP: "Attention […] names the means or process by which an organism increases the gain (the weighting, hence the forward-flowing impact) on those prediction error units estimated to provide the most reliable sensory information relative to some current task, threat, or opportunity" (Clark 2016, pp. 59-60).

As stated above, the goal of the brain is to minimize the amount of prediction error generated in the system. According to PP there are two ways of achieving this goal. The first is to explain away the prediction error by deploying better predictions that fit the upcoming signal, i.e. perception. A second and complementary strategy is to modify the stream of sensory data so that it matches the predictions better. This is, in PP terms, action: actively seeking to match the predictions by interacting with the environment and sampling it through bodily movements, in order to produce or evoke the sensory consequences expected by the brain. But how is the mechanism of action implemented in the first place? In PP, proprioceptive predictions play a central role in determining actual bodily movement. Motor control, as Clark (Clark 2016, p. 121) puts it, is "subjunctive": given a prediction of a non-actual proprioceptive state, the body will move accordingly in order to minimise prediction error.

According to the framework[2] then, perception and action are two recurring and complementary strategies adopted by the brain to minimise prediction error. Their combination and cyclical succession — labelled "active inference" (Friston et al. 2011) — seem to be the very basis of our interaction

---

[1] A generative model "[…] aims to capture the statistical structure of some set of observed inputs by inferring a causal matrix able to give rise to that very structure" (Clark 2016, p. 21).

[2] Clark (Clark 2013, Clark 2016), in this regard, proposes the label "action-oriented predictive processing" to differentiate from other approaches, like the one proposed by Hohwy , which do not place as much emphasis on the mechanism of active inference.

with the world (Hohwy 2013, pp. 90-92; Clark 2016, pp. 120-124). According to the action-oriented formulation of PP the interaction with the environment is crucial in determining the specific quality and accuracy of our percepts. This becomes particularly relevant for the PP explanation of dreaming, as we shall see below.

2.2 The State of the Art: Neurobiology of Sleep According to Predictive Processing

Recent works by Hobson and Friston (Hobson and Friston 2012; Hobson and Friston 2014) summarise the evidence in support of a PP explanation of dreaming and the labour of the sleeping brain.

In a nutshell, the idea is that "[…] the brain is essentially doing the same thing in sleep and waking; with one key difference — there is no sensory input during sleep. However, the recurrent hierarchical message passing is still in process; with continually changing expectations and hierarchical predictions that constitute dream content." (Hobson and Friston 2014, p. 8). In other words, the very same cognitive architecture that drives the perception-action loop when awake is still active during sleep, but devoid of the important role of environmental perceptual input and motor feedback.

Hobson and Friston base their view on the combination of the famous AIM model (Hobson et al. 2000) and the free energy principle (Friston 2010)[3]. The AIM model makes use of a multidimensional state-space for keeping track of the brain's changes in activation (A), input-output gating (I) and neurochemical modulation (M). The shifts in parameters are mirrored by shifts in subjective experience. In particular, two positions in the state-space are relevant for their discussion: wake (characterised by high activation, externally driven processing and prevalently aminergic modulation) and REM sleep (characterised by high activation, internally driven processing and cholinergic modulation). The peculiar condition of the brain during REM sleep determines the formation of dreams and their sometimes bizarre and perceptually unstable narrative. During REM sleep, in fact, the change in neurochemical modulation affects the input gating so to draw the attention of the brain from the sensory periphery to internally generated activations.

In PP terms this means that the stream of bottom-up sensory information is attenuated through the assignment of low precision, and the brain has to rely mainly on internally generated predictions (based on progressively more abstract, middle-to-high level priors) to carry out the task of minimizing prediction error (Clark 2016, pp. 98-102). To complete the picture, the specific activations of the brain during REM sleep are different from waking (Hobson and Friston 2014, pp. 9-10): activations of the primary visual and non-visual sensory cortices, the thalamocortical sensory system and basal forebrain explain the perceptual-like character of dreams. Particularly relevant in this context is the lack of activation of areas of the prefrontal cortex (deputed to executive functions), which would explain the diminished meta-cognitive awareness during dreams.

The motor cortex is still active and presumably deploying motor and proprioceptive predictions to engage in active inference; however, motor commands are inhibited at the pontine level, resulting in REM atonia and an effective paralysis of the body[4]. As a consequence, proprioceptive predictions can never be fully satisfied by proprioceptive feedback (as would happen during wake, when the strategy of active inference would elicit actual movements in the environment), forcing the brain to jump from one prediction to another, determining the inconsistent nature of dream narrative. Of particular relevance in this context is the role of ponto-geniculate-occipital (PGO) waves, which originate in the brainstem and proceed all the way up to the visual cortex. Their presence is correlated with anticipation and elicitation of ocular movements in both wake and REM sleep (Hobson and Friston 2012, pp. 85-90; Hobson and Friston 2014, p. 8). During wake, PGO waves peak in response to a change in peripheral vision (increase of surprisal), bringing about new predictions which need to be matched with the sensory stream in order to minimise prediction error. This in turn drives attention towards

---

[3] The free energy principle is a much more general theory that encompasses the PP proposal (see Clark 2016, pp. 305-306, for more details).

[4] In fact, subjects affected by REM sleep behaviour disorder lack the blockade of motor output and "act" their dreams while asleep.

the (visual) sensory periphery and a motor prediction is issued from the motor cortex, which results in a saccade towards the origin of the input. This is a typical example of the perception-action loop in wake, a strategy that usually is effective in minimising prediction error. However, during REM sleep, the precision assigned to the bottom-up sensory stream of information is very low. PGO waves remain present though, as do ocular movements[5]. The system therefore has to make sense of randomly generated activations of several areas as well as the actual deployment of ocular movements to fit visual predictions without the aid of the fine-grained environmental feedback. In our understanding of PP, these conditions force the brain to rely only on its available middle-to-high level priors (which have a more abstract nature) for the formation of the perceptual scene (the *dreamscape*), since the low-level priors are flagged (via precision weighting) as unreliable. This means that the dreamscape will be populated by objects that lack the fine-grained perceptual details and depth provided from the external environment during wake. Furthermore, these objects will be more likely to present bizarre and mixed features and the stability and continuity of the perceptual scene will be partially compromised. In the next section we will analyse the phenomenal aspects of dreaming in more details; for now, we argue that PP can account for the peculiar character of the phenomenal aspects of dreaming by linking it to the differences in the neural work during sleep. However, as we shall see later, this is not sufficient to explain why we dream in the first place.

It must be noted that the picture presented by Hobson & Friston is controversial insofar as it takes into account REM stages only to explain dreaming. However, it is now widely accepted in the dream research community that dream mentation can happen during NREM stages as well as during transitional stages such as hypnagogia (onset of sleep) and hypnopompia (onset of wakefulness) (see Windt 2015, pp. 50-56 and 530-550, for a detailed critique of transitional states), although they differ from REM dreams because they tend to be more similar to static images and they lack a narrative development. Moreover, recent articles (Domhoff 2011b; Fox et al. 2013) that connect dreaming to other cognitive phenomena instantiated by the default-mode network (DMN) put pressure on the neurological description provided above. A detailed analysis of the neurological details goes beyond the purpose of this paper: for example, the peculiar differences in content and vividness between NREM and REM sleep would deserve a separate discussion that accounts for the different occurrences of PGO waves. It is worth noticing, though, that the general architecture described by PP would be compatible with a theory of dreaming formation involving more widespread brain activations, as long as it doesn't contradict the main tenet of a predictive perception-action loop[6] ongoing in the brain.

2.3 Dreaming in the Bayesian Brain: Theoretical Advantages of Predictive Processing

From the general description of PP and the analysis of the scientific literature explaining dreaming through the framework, we can now propose a clearer definition of what dreaming is in this context.

> **Dreaming** $=_{\mathrm{Df}}$ A process of hypothesis testing through perception-action loops under the constrained, altered neurophysiological conditions of sleep.

The mechanism of prediction error minimization is always in place - but the conditions under which the mechanism operates are different. Therefore, it is arguable that the brain tries to instantiate the loop with the external environment, but fails to do so effectively. Without the important feedback of the external environment, the brain "runs wild" from one prediction to another, in accordance with the probability distribution among priors expressed by the generative model, while trying to make

---

[5] Extra-ocular muscles are one of the few groups of muscles barely affected by REM atonia.

[6] The presence of a perception-action loop in dreams has not been widely discussed in the literature so far. A provisional suggestion is that the dreamscape plays part of the role of the external environment, insofar as the dreamer "acts" in the dream world.

sense of the (mostly) internally generated stream of information. As Windt (Windt 2015, p. 603) nicely puts it, dreaming is "a process of mental improvisation".

The PP explanation of dreaming has a clear theoretical advantage, in response to the integration problem mentioned above, insofar as it is by definition inclusive: it encompasses all forms of cognition under the same architecture and ongoing mechanism. PP therefore blurs the line between cognitive phenomena that were traditionally conceptualised as distinct: imagination, mind-wandering, dreaming, hallucinations, standard waking perception are all generated by the same predictive engine under different circumstances. In fact, as Clark puts it, "perceivers like us, if this is correct, are inevitably potential dreamers and imaginers too" (Clark 2012, p. 764). PP poses a necessary link between a set of potential cognitive phenomena that will arise from the specific Bayesian hierarchical architecture described above, distinguishing itself as a particularly parsimonious framework, while at the same time retaining the ability to explain the specific character of each of those phenomena.

In the next section, we will examine two more theoretical advantages of PP. The first is that it provides a clear insight on the correlations between neural states and subjective experience in dreaming. This ties in directly with the explanation of consciousness. Why do we dream in the first place? We will introduce Integrated Information Theory (Oizumi et al. 2014; Tononi et al. 2016) to address this question, showing the benefits of a comparison with PP. Secondly, PP is also a good framework in which to understand the evolutionary role of sleep and dreaming. To show this, we will provide an explanation of the advantageous mechanism of generative model optimisation operated during sleep that is compatible with the latest empirical evidence offered by sleep research. With the aid of IIT, we will also tackle the issue whether dreaming has an evolutionary role *per se*.

3 The Phenomenal Character of Dreaming

3.1 What Does it Feel like to Be Dreaming?

As seen above, during REM sleep the brain works under different conditions compared to wake. These conditions determine the resulting phenomenal character. The selective deactivation of large portions of the prefrontal cortex and the subsequent diminished meta-cognitive awareness contribute to the immersive nature of the dreaming experience. This picture led researchers to compare dreaming to a form of intensified mind-wandering (Fox et al. 2013, pp. 10-11). Indeed, in both cases the attention is drawn from the external input to internally generated stimuli, creating a form of (partial) seclusion from the environment. In PP terms, we may construe this as the assignment of low precision to the stream of information coming from the sensory periphery. It makes sense, phenomenally speaking, to compare mind-wandering to dreaming: while our mind strays from the present tasks, we feel immersed in our own thoughts to the point that we lose contact with our surroundings and only direct, sudden or life-threatening stimuli bring us "back to reality". During REM sleep, the physiological changes conspire to seclude us even more, raising the threshold for external stimuli to pass into the system.

This is far, however, from stating that the sleeping brain is totally disconnected from its environment. There are many cases in the literature that report integration of external stimuli into dreams. A study on the effect of somatosensory stimulation on dreaming (Sauvageau et al. 1998) provides an interesting example. For the study, the participants were monitored in a sleep lab and the stimulation was administered through the inflation of a blood pressure cuff fitted above the knees. Here is the dream report of one of the subjects upon awakening: "I was in our school gym bleachers. I decided to go join some gymnasts on the floor. It was really crowded with people; I've never seen so many. I was making my way through the crowd all out of breath and there was this big woman with a scarf. *The scarf got hooked on my leg and I couldn't get it off.* I could feel it there; it didn't hurt, but it bothered me that I couldn't take it off." (Sauvageau et al. 1998, p. 11, italics in original). From the PP perspective, this would be a clear case of deployment of new predictions (the scarf wrapped around the leg) to match a stimulus that has passed the high threshold imposed by sleep (the pressure applied by the cuff). However, the functioning of the predictive hierarchy is still disrupted, from which follows the (wrong) attribution of the stimulus to the scarf. Please note that, although the origin of the stimulus is somehow explained away by the sleeping brain, it can be argued that a certain degree of error is still present, hence the emotional reaction of the dreamer (she is bothered that she could not take off the scarf).

The occasional dream oddities resulting from the integration of external stimuli are a case of a more general phenomenal feature: bizarreness (Scarone et al. 2008; Noreika et al. 2010). Although this feature is probably not as frequent as Hobson's AIM model (which compares dreaming to a form of psychosis) would claim (see Domhoff 2007, for a critique), PP can also account for it. The brain relies on middle-to-high level priors which are more abstract in nature and it can't match the resulting hypotheses with the informationally-rich stream of external sensory input. This generates a disruption in the "binding process" of dream content (Revonsuo and Tarkko 2002), resulting in oddities such as mixed features of objects and dream characters (like dreaming of a duck-winged man), and contextual displacement (i.e. people or objects appearing where they are not supposed to in a standard waking environment). It is worth noticing, however, that the study reports the presence of bizarreness in only about 50% of the reports analysed, and of these cases "only 37% concerned the internal features of the representation itself (structure and outlook, familiarity, semantic knowledge, temporal continuity)" (Revonsuo and Tarkko 2002, p. 14), the remaining cases being related to contextual bizarreness. In PP terms this could be interpreted, all in all, as a sign of the strong reliability of the generative model upon which predictions are based, albeit in impaired conditions.

Another salient phenomenological feature of REM dreams is their narrative development (Hobson et al. 2000; Nir and Tononi 2010). The narrative structure is overall (bizarreness aside) continuous and similar to waking experience, a fact reflected by the ability to report dreams in a narrative fashion. In PP terms this could be directly linked to the way organisms construct their own ongoing subjective experience - that is by deploying active inference. The sleeping brain does not simply process scattered, random stimuli, but a constant flow of endogenous activations that it makes sense of through the aforementioned perception-action loop. It is not clear, however, how much of the cognitive architecture has to be in place and functioning to guarantee the presence of a cohesive narrative. NREM dreams are also widely reported, and although sometimes different in their content (more static, conceptual, less vivid), they can also be narrative in nature. Hypnagogic hallucination seems to be a better case of dream-like imagery without a narrative (Nir and Tononi 2010, p. 94). Studying these cases might provide a good way to pinpoint the exact neuronal circuits that implement the active inference strategy in contrast to cases in which the brain simply "stands still" on the internally generated input.

PP's explanation of dreaming has the potential to accommodate well the vast empirical literature on dream phenomenology and its neural substrate. However, there is still no indication in regards to why dreaming should be a form of conscious experience at all.

3.2 Why Doesn't PP fully Explain Consciousness?

Since the formulation of the "hard-problem" of consciousness (Nagel 1974; Chalmers 1996), the question of "why it feels like this, or like anything at all, to be something" has puzzled philosophers and scientists alike. In particular, Chalmers has proposed and discussed various arguments against physicalism, either casting doubts on the (nomological, logical, or metaphysical) supervenience of phenomenal properties on physical properties, or arguing against their identity with them, given the impossibility of deducing all truths about phenomenal facts from the complete knowledge of truths about physical facts (Chalmers 1996). The problem concerning phenomenal consciousness becomes even more evident in the case of dreaming: while it might make some sense, at least evolutionarily speaking, to be able to have subjective experiences during wake in order to better cope with the environment, it is not immediately evident why we should experience the internally generated world of dreams instead of simply shutting down (phenomenally speaking) for a few hours per night.

Tentative solutions to the hard problem have been advanced in the PP literature. Hobson & Friston (Hobson and Friston 2014) and Hobson et al. (Hobson et al. 2014) equate consciousness to a form of active inference, while appealing to a "Cartesian theatre" metaphor to account for the connection between the labour of the brain and the subjective phenomenology. It is not clear, however, in what sense their proposal would protect the PP framework from the classical zombie objection (a system might have all the functional and behavioural properties of a conscious system, but no internal subjective experience)[7].

Hohwy proposes that "conscious perception is determined by the hypotheses about the world that best predicts input and thereby gets the highest posterior probability." However, he continues, "[…] this is not intended as a proposal that can explain why perceptual states are phenomenally conscious rather than not." (Hohwy 2013, pp. 201-202). In other words, Hohwy aims to explain precisely which specific representational content (among the many predictions elaborated by the brain) generates a subjective phenomenal experience, how the latter is generated and under which conditions. Interestingly, he later proposes to connect this to the Global Neuronal Workspace theory developed by Baars and Dehaene (Hohwy 2013, pp. 211-214). This proposal seems to suggest that only a small part of the information processed by the hierarchical architecture flows into consciousness, or in other words that consciousness is composed of/emerges from a series of subpersonal, subconscious processes.

On the same line, Clark suggests that Predictive Processing might be on the right track to begin to solve the hard problem (Clark 2016, p. 239). By explaining all the components of conscious experience, such as the sense of self, the sense of presence, agency, emotions, as well as the perceptual milieu and other cognitive features (imagination, dreaming and the like) under the same predictive architecture, PP is in a sense dissolving the hard problem piece by piece. However, Clark notes: "True believers in the hard problem will say that all we can make progress with using these new-fangled resources is the familiar project of explaining patterns of response and judgment, and not the very existence of experience itself." (Clark 2016, p. 324). This PP solution would work only for detractors of the hard problem (Dennett 2013). We want to remain agnostic on this point, and propose an alternative approach that tries to face the hard problem directly: Integration Information Theory (IIT).

3.3 Integrated Information Theory: What it Is

IIT (Tononi 2008; Tononi 2012; Oizumi et al. 2014; Tononi and Koch 2015; Tononi 2015; Tononi et al. 2016) attempts to account for consciousness by linking the phenomenological evidence we get from our own experience to the evidence provided in the last decades by cognitive neuroscience. In fact, IIT maintains that in order to solve the hard problem of consciousness a change of perspective is needed: "[A]s long as one starts from the brain and asks how it could possibly give rise to experience [...] the problem may be not only hard, but almost impossible to solve. But things may be less hard if one takes the opposite approach: start from consciousness itself, by identifying its essential properties, and then ask what kinds of physical mechanisms could possibly account for them. This is the approach taken by integrated information theory (IIT)" (Tononi and Koch 2015, p. 5).

The theory starts by identifying the properties of conscious experience (which are described as "axioms") and establishes connections with properties of the physical system that supports them ("postulates"). Each axiom about phenomenal experience has a corresponding postulate about the physical substrate (Oizumi et al. 2014; Tononi et al. 2016). The first axiom is *intrinsic existence*: experience exists, it is actual, undeniable, self-evident, and intrinsic, namely independent of external observers or

---

[7] For a critical discussion of these ideas, see Dołęga & Dewhurst (Dołega and Dewhurst 2015) and response from Hobson & Friston (Hobson and Friston 2016).

objects. This axiom constitutes the starting point of IIT and corresponds to the Cartesian assumption that conscious experience is a given, a self-evident and indubitable truth. The corresponding postulate claims that the system supporting such experience must exist “intrinsically”, namely must have cause-effect power upon itself[8].

The second axiom, *composition*, claims that the structure of experience is composed of multiple (higher-order) “phenomenological” distinctions, namely different aspects of each individual experience (such as the perception of various objects seen in the visual field, their shape, parts, colour, extension, etc.). The corresponding postulate claims that the system must be composed of sets of elements with cause-effect power within the system, forming a structure of mechanisms of different order that corresponds to the structure of phenomenal experience.

The third axiom, *information*, states that conscious experience is specific and that every conscious state is informative, inasmuch as it identifies specific sets of phenomenological distinctions and differentiates from (rules out) other possible experiences. The corresponding postulate states that the system must specify a cause-effect structure, roughly the repertoire of activation states of the mechanisms that compose the system, which characterizes the cause-effect profile of such states and differentiates it from other possible ones.

The *integration* axiom claims that conscious experience is unified and irreducible to its component phenomenological distinctions (i.e. experience comes as a whole, the experience of a blue book is not reducible to the experience of a colourless book plus the experience of the colour blue, nor they can be experienced separately). The postulate states that the cause-effect structure specified by the system must be unified and intrinsically irreducible to the one specified by non-interdependent sub-systems obtained by unidirectional partitions (namely a partition that is unable to affect and be affected by the activity of other parts of the system). The degree of intrinsic irreducibility is measured as integrated information (Φ), which quantifies the changes in the cause-effect structure of a system when the system is partitioned.

Finally, the *exclusion* axiom claims that consciousness has a definite spatio-temporal grain, and that there is only one conscious subject at a time (which cannot have parts or be part of a bigger subject). The corresponding postulate claims that the cause-effect structure of a system must be definite: it is specified over the set of elements that is maximally irreducible from its intrinsic perspective, hence having the highest level of integrated information (Φ^{MAX}) among all its sub-systems or the systems that comprise it.

Starting from these axioms and postulates, IIT describes the phenomenology of consciousness as constituted by the sum of informational relationships among activation states of the system. The cause-effect repertoire, or *qualia space*, is defined as a high-dimensional space with one axis for each possible past and future state of the system in which a structure of *concepts* can be represented (Tononi 2012). Qualia are sets of informational relationships in high Φ-level generating systems, or maximally irreducible cause-effect repertoires, called “concepts”, generated by a complex of elements. Φ here represents the measure for intrinsic irreducibility called *integrated information*. Φ also quantifies how the cause-effect structure changes when the system is partitioned. In fact, the greatest role is played by what Tononi calls maximally irreducible conceptual structures (MICS). A conceptual structure is a constellation of points in concept space, where each axis is a possible past/future state of the set of elements, and each point is a concept specifying differences that make a difference within the set.

Besides axioms and postulates, IIT posits a central identity: every experience is identical with a conceptual structure that is maximally irreducible intrinsically, namely with a MICS (Tononi et al. 2016). In particular, the “quality” of the experience - its content of phenomenal distinctions - is spec-

---

[8] Causal power is a condition for existence, and cause-effect power upon itself is a condition for the existence of a system independent of external observers. Tononi writes: “in order to exist, [a system] must have cause-effect power, as there is no point in assuming that something exists if nothing can make a difference to it, or if it cannot make a difference to anything. Moreover, to exist from its own intrinsic perspective, independent of external observers, a system of elements in a state must have cause-effect power upon itself, independent of extrinsic factors.” (Tononi 2015, p. 4164).

ified by the form of the conceptual structure (by the concepts and their relationship in cause-effect space), whereas the "quantity" of the experience - its level - is given by its irreducibility (Φ^{MAX}), i.e. the quantity of integrated information of the MICS. IIT therefore posits an identity between integrated information in a system and conscious phenomenal experience. This equals to say that, if a system satisfies the requirements for having a non-null-quantity of Φ, it will be conscious by definition.

3.4 Integrated Information Theory: How it Helps

IIT makes a number of predictions concerning the neural processes fundamental for consciousness and their impairment in pathological and altered (i.e. non-standard waking) conditions (Casali et al. 2013). In particular, perturbational studies using TMS (Transcranial magnetic stimulation) as a method of breaking down cortical connectivity and reactivity (and therefore, the level of Φ^{MAX} in the brain) have shown that in REM sleep (Massimini et al. 2009; Massimini et al. 2010) and some instances of NREM sleep (Nieminen et al. 2016) the level of integrated information in the brain is high enough for exhibiting conscious experience (dreams, reported upon awakening). On the contrary, extensive cortical connectivity breakdowns (as the ones happening during slow-wave sleep and anaesthesia) impair the system to the point where no global conscious experience is possible. What this means is that waking-like consciousness does not appear in such cases, but few sub-complexes could still lead to limited phenomenal experience[9]. This fits nicely with PP insofar as only certain areas of the brain might be sufficient in order to produce conscious experience, but at the same time the variations in the stream of information available to minimise prediction error determines variations in phenomenal experience, due to reliance on different priors. It also accounts for NREM dreams in regards to the different phenomenal content described above: more sparse activations during NREM stages would result in segregated information, but the local maxima might still be sufficient for static, non-narrative perceptual-like mentation.

Moreover, IIT states that the brain is organized like a device for "interpreting" spatial and temporal correlations representative of the causal structure of the environment, in the light of its memories (stored in connections). This causal structure is incorporated in the connectivity of the system via natural selection and learning mechanisms. IIT calls *matching* the measure of how well the integrated conceptual structure generated by an adapted complex fits or "matches" the cause-effect structure of its environment (Oizumi et al. 2014; Tononi 2012). In other words, IIT tries to account for the adaptation of the cognitive structures and behaviour of biological organisms (and artificial systems) to the environment. This is reminiscent of the PP description of the role of the generative model (what IIT calls the "integrated conceptual structure"): the process of increasing the matching value can be seen as equivalent to the refinement of the generative model[10]. This is achieved by a progressive reorganisation of priors (concepts or structure of concepts, in IIT terminology) via learning and, as we shall see below, sleep.

Given the assumption of the identity between consciousness and integrated information, IIT predicts that conscious experience must correspond to the high levels of Φ measured during REM sleep (Massimini et al. 2010) and NREM sleep (Nieminen et al. 2016), we suggest that PP might benefit

---

[9] This view implies that when the level of Φ of the whole system becomes lower than the level of Φ of some of its parts, the main conscious subject ceases to exist, but other subjects (corresponding to the new sub-complexes with maxima of Φ) come into existence. The issue concerning this consequence of the theory deserves deeper analysis, which goes beyond the scope of this paper, but its discussion should nonetheless be of top priority if IIT were to be a complete theory of consciousness.

[10] As suggested by Metzinger (personal communication), there is a long-standing debate on the possibility that adaptive mechanisms will lead to self-deception rather than an accurate modelling of the world. It is therefore debatable to what extent the mechanisms described by both PP and IIT will lead to a good mapping of the causal structure of the world. A full discussion of the topic goes beyond the purpose of this paper; it will suffice to say that the convergence of descriptions of mechanisms between PP and IIT remains valid independently from the actual product of these mechanisms (whether it would be accurate or adaptive but self-deceptive). For further reading on the topic, see Von Hippel and Trivers 2011; for a discussion of self-deception in the light of PP, see Pliushch 2017.

from assuming a similar non-eliminativist position in regards to phenomenal aspects of consciousness and the hard problem.

It could be argued that IIT has several problems on its own. There is no space here for a thorough critique of the theory, but a relevant objection in this context is that it does not really solve the hard problem. In fact, the solution proposed by the theory comes from the assumption of the fundamental identity between consciousness and integrated information (more precisely, maximally irreducible conceptual structures). This solution constitutes a third way of responding to the hard problem, different from both classic dualism and reductive/eliminativist physicalism. As seen above, different versions of PP have different degrees of resistance to the hard problem and consequently different degrees of compatibility with IIT. The analysis offered in this paper justifies the provisional assumption that a connection between the two theories is feasible and potentially fruitful. Despite a more detailed analysis is needed, the similarities in the conceptual vocabulary used by IIT and PP allow for a direct comparison between the two theories.

In the next section, we will turn to another theoretical advantage of PP, that is the explanation of the evolutionary role of sleep. With the support of IIT, we will later argue for an evolutionary role of dreaming as well.

4 The Evolutionary Role of Sleep and Dreaming

4.1 The "Synaptic Homoeostasis Hypothesis"

The evolutionary role of sleep and dreaming has been an object of discussion among the scientific community for a long time. It is not evidently clear why an organism should, for at least some time a day, stop doing evolutionarily advantageous activities (like foraging for food) and become potentially vulnerable to predators. However, there is an increasing amount of evidence that points in the direction of a cognitive benefit deriving from sleep.

A particularly relevant proposal (independent from but compatible with IIT) has been advanced by Tononi & Cirelli (Tononi and Cirelli 2014) in a recent review of the state of the art in neurobiology: the "synaptic homoeostasis hypotheses" (SHY). SHY claims that sleep is a way to improve overall synaptic organisation and restore energetic equilibrium (homoeostasis) in the brain. During wake, our brains constantly form new synaptic pathways and neuronal connections, strengthening them in response to the stimuli from the environment. However, the brain's resources are limited and neural plasticity takes its toll: over time the energetic expense for synaptic maintenance, combined to synaptic saturation and decreased signal-to-noise ratio will become disadvantageous. This explains the presence of sleep: according to SHY, during slow wave sleep (i.e. NREM stages 3 and 4) the combination of synaptic depotentiation (triggered by cholinergic neuromodulation) and spontaneous activation throughout the brain contributes to an overall downgrading and optimisation of the synapses and, as a consequence, the restoration of synaptic homoeostasis. This process is called "activity-dependent down-selection" (Tononi and Cirelli 2014, p. 15). During wake, the brain searches for potential statistical regularities in the environment (repetitive and suspicious coincidence in the sensory input) and potentiates synapses accordingly. The stronger the regularities, the stronger the synaptic bonds. However, noise too could drive synaptic formation, thus leading to maladaptive connections and over-fitting of the model in the long run, since the waking brain will continue to form new synaptic connections in order to grasp all the possible regularities in the environment. During slow wave sleep, the brain operates a general connective depotentiation through low frequency diffused activations in neuromodulatory conditions that promote synaptic depression (Tononi and Cirelli 2014, p. 19). The strongest connections (i.e. the one representing the strongest regularities learned while awake) will be less affected, resulting therefore in a relative reinforcement and subsequent increase in signal-to-noise ratio. To give a simplified example of this, think of two different sets of connections, the first (representing strong regularities) with value *n* and the second (representing weaker ones) with value *n-2*. If the downgrading operates on a -3 factor, over time the weaker connections will be deleted, leaving only the stronger (more adaptive) ones in place and more room for new connections (learning) upon awakening.

This process contribute to a better consolidation of previous useful information while integrating it in long-time learning schemes, and systematically protecting it from noise interference (by forgetting noise-related connections). Finally, the "gist extraction" of important regularities from the environment is ultimately improved (Tononi and Cirelli 2014, pp. 21-23).

4.2 Over-Fitting Avoidance and Optimisation of the Generative Model

The mechanism described by SHY bears striking similarities to the PP account of the evolutionary role of sleep. Hobson & Friston (Hobson and Friston 2012), in response to the apparently non-adaptive loss of thermoregulation during REM stages, propose that this is nonetheless a necessary step in order to reduce the general complexity of the generative model and avoid over-fitting. This idea has roots in the "wake-sleep algorithm" proposed by Hinton (Hinton et al. 1995). The mechanism is simple: during wake the system learns new things (optimising posterior beliefs), but over time its generative model of the world becomes overly-complex, incapable of distinguishing between meaningful signal and noise. A second phase is therefore needed: "During sleep, the brain's model is insulated from further sensory testing but can still be improved by simplification and streamlining. […] Sleep may thus allow the brain to engage in synaptic pruning so as to improve (make more powerful and generalizable) […] the generative model" (Clark 2016, p. 101). This "synaptic pruning" (i.e. removal of redundant, weak connections) is precisely what SHY describes as happening during NREM sleep.

We will discuss the different roles of REM and NREM stages of sleep in the optimisation of the generative model in the next section. For now, let's turn back to dreaming.

4.3 Are Dreams Adaptive?

We showed that PP provides an elegant and rich explanation for the adaptive role of sleeping. But what about sleep mentation (like REM dreaming or hypnagogic imagery)? Does it have an evolutionary role or is it just a "spandrel of sleep" (Flanagan 1995)? In the light of the discussion above, there are two main reasons to believe that dreaming has an adaptive value.

First, provided that assuming a convergence between PP and IIT is fruitful, the identity postulated between integrated information and consciousness suffices to establish that sleep mentation is a form of conscious experience, insofar as the adaptive neural mechanism ongoing during sleep implies high levels of Φ. More generally, it seems that in this context the question about an evolutionary role of dreaming per se is ill-conceived: given that the level of integrated information in a system depends on the structure of the system itself, a structure that has a clear evolutionary advantage in sleep, it makes poor sense to distinguish between the two. The question of what is the evolutionary role of dreaming (if there is any) is then linked to the more general question of what is the evolutionary role of consciousness.

Second, and in support to the previous point, "IIT predicts that adaptation to an environment should lead to an increase in matching and thereby to an increase in consciousness." (Tononi 2015). This means that an increase in matching, or in PP terms the optimisation of the generative model, is strictly connected to a wider and richer phenomenal repertoire (a higher value of Φ). Studies on adaptive logic-gate networks, or animats (Albantakis et al. 2014), have shown that over the course of their adaptation, integrated information increases with respect to task fitness and matches the complexity of the environment. Although this form of task-dependent fitness was tested exclusively online, it is plausible that in much more complicated neural networks like our brains the sleep phase contributes to the same ultimate result: increase in matching. Given that dreaming is the inevitable phenomenal aspect of this procedure (for during both REM and NREM dreams the level of Φ is high enough, see

§3.4), it makes no sense to ask for an evolutionary role disconnected from the neural processes that give raise to it.

In sections 3 and 4 we have shown in depth what we think are two important theoretical advantages of adopting PP in the study of dreaming, discussing the possible issues arising. In the next section we will outline two open problems in dream research and we will suggest that they can be fruitfully re-examined through the lens of Predictive Processing,

5 Two Avenues of Research for the Future

5.1 The Dreaming Body: Is Dreaming an Exclusively Internal/Off-Line Simulation?

Dreaming has often been referred to in the literature as a form of cranial-bound, off-line, disembodied experience. Windt (Windt 2015, pp. 350-354) neatly sums up this position as the "functional-disembodiment hypothesis", according to which the sleeping-body inputs and motor outputs are disconnected from the phenomenal experience presented in dreams. This hypothesis could be interpreted as assimilating dreaming either to a radical form of cranial envatment or to a particularly vivid case of imagination. Indeed, the very depiction of dreaming by Hobson & Friston presented in the first section of this paper seems to fall in line with it. This position, although widespread, has limited scope since it does not account for phenomena like the integration of external stimuli and, more generally, the role of the complex dynamics in the body while asleep.

However, we want to suggest that this position does not represent PP in its entirety. In fact, a growing body of literature within the framework seems to indicate that bodily predictions are almost always in place. As suggested by Seth (Seth 2013), interoceptive predictions (i.e. predictions about the states of our own internal organs, muscular and visceral sensations, hunger, pain, breathing, etc.) constitute the basis for our sense of conscious presence in the world and the building blocks for the formation of a rich, embodied conscious experience. Arguably, proprioceptive predictions (i.e. predictions about the position of our body in space) might play a role too, in particular for what concerns vestibular positioning and motion (Dharani 2005). Even emotions, explained as arising from a combination of interoceptive predictions (as described above) and cognitive ones (how we are "supposed to feel" about our own body reactions) (Clark 2016, pp. 231-235), might play a role inasmuch as they constitute a conspicuous part of dreaming phenomenology.

Bodily predictions, according to PP, are therefore an important part of the cognitive hierarchy and the study of their neural instantiation should provide a better insight of why dream phenomenology has a "diminished" embodied flavour. Interestingly, a recent study (Windt et al. 2014) used lucid dreams as a condition to explore the subjective difference between self-tickling (usually ineffective) and being tickled by another (usually intense). During wake condition, this pattern of subjective difference was respected. The results in the lucid dreaming condition indicate however that a form of sensory attenuation is in place: in spite of the dreamer "commanding" other dream characters to tickle her, the subjective feeling is almost indistinguishable from self-tickling in intensity. Interpreted through the lens of PP (Clark 2016, pp. 112-114), the impossibility of self-tickling is due to the inclusion of motor commands issued for the tickling itself into predictions of the sensory output, thus leading to a general sensory attenuation. The results of the lucid dreaming experiment indicate that bodily predictions are still in place during sleep. However, "a strong conviction driving these effects in lucid dreams might be that to the extent that one is able to control an agent, this agent cannot be fully distinct from oneself" (Windt et al. 2014, p. 7). The authors here refer explicitly to the presence of hyperpriors that guide body-ownership attribution — and those hyperpriors, in the case of dreaming, would be among the ones leading predictions, in the absence of reliable sensory input.

Following this, we do not want to suggest that dreams are fully embodied experience, but rather support the proposal that “dreams are weakly functionally embodied states” (Windt 2015, p. 383): bodily predictions, especially in the form of high-level priors, are still present during dreaming, therefore it would be inappropriate to think about it as a form of disembodied mentation. More generally speaking, PP helps rethinking dreaming as yet another particular state in which the cognitive architecture operates: one that shuts off the environment but only up to a certain threshold, and that doesn’t cancel out the role of the body in the formation of the dreamscape. Future empirical research are needed in order to establish exactly how “weak” this functional embodiment is.

5.2 Synaptic Pruning or Synaptic Strengthening? The Exact Role of Sleep Stages in the Optimisation of the Generative Model

Recall from section 4.2 that according to both SHY and PP, sleep serves a role of optimisation of the generative model via global activity-dependent down-selection. However, the supporters of SHY claim that this synaptic pruning is mainly operated during slow wave sleep, while Hobson & Friston (Hobson and Friston 2012) focus their attention on REM sleep. Given the current state of the evidence provided, it is still unclear when exactly the optimisation is conducted. Interestingly, a recent controversy (Heller 2014 and response by Cirelli and Tononi 2015) puts pressure on the general picture drawn by SHY: some studies suggest that sleep does involve synaptic potentiation and strengthening as mechanisms of memory consolidation. There is indeed a broad literature about the role of both NREM and REM sleep in memory evolution, recombination and integration in mental schemes (Stickgold and Walker 2013). The reward system also seems to play an important part in this respect, explaining the highly emotional character of REM dreams (Perogamvros and Schwartz 2012; Perogamvros and Schwartz 2013).

More empirical research is needed to set the debate. Our tentative proposal is to consider the different stages as performing two different but complementary functions: NREM sleep would operate synaptic pruning, while REM sleep would strengthen the connections via randomly generated, wake-like levels of activity. This would make sense if we think that a continuous depotentiation overnight would deplete the brain of possibly useful but freshly formed (therefore, weak) synapses. Alternating NREM synaptic pruning to REM synaptic reinforcement would avoid that loss. Additionally, reinforcement during REM sleep would be conducted without the noise disturbance of the bottom-up sensory stream but only among middle-to-high level priors, leading to a better internal coherence of the generative model. The optimisation would be thus conducted via a “two-steps mechanism” that alternates overnight. Interestingly, this idea has been already proposed in the scientific literature: Giuditta and colleagues (Giuditta et al. 1995) talked about a “sequential hypothesis” (SH) of the function of sleep. In their view, mostly related to the positive effects of sleep on memory consolidation, slow-wave (NREM) sleep would serve as a preliminary mechanism of general depotentiation and simultaneous flagging of memories, while REM sleep would help to store and potentiate those important memories. We argue that SH can be easily meshed with PP. A two-steps mechanism, as suggested above, would also be compatible with SHY, as the authors themselves admit (Tononi and Cirelli 2014, p. 27). However, it must be noted that in a more recent paper Giuditta (Giuditta 2014) remarks important differences between SH and SHY on several points: the energetic needs and the nature of the activity of the brain, the methodological approaches to support the respective claims, and the role of REM stages in memory consolidation. Establishing the exact function of each stage of sleep is important, from PP’s perspective, to understand their variation in length over the ontogeny and their possible role in the early development of the generative model (Segawa 1999; Hobson 2009). More research is needed to figure out the exact neurological mechanisms and processes: we suggest that the adoption of PP would be a profitable way to frame the future research in order to overcome the differences between competing hypotheses.

6 Conclusion

This paper aimed to provide an overview of the understanding of sleep and dreaming within the Predictive Processing framework. We described three theoretical advantages in adopting PP. First, it is an integrative, inclusive framework, insofar as it explains several cognitive phenomena under the same cognitive architecture. Second, it has a good grip on the phenomenal aspects of dreaming. Third, it provides a clear and elegant explanation of the evolutionary role of sleep. In relation to the last two aspects, we observed that PP still lacks a good answer to the hard-problem of consciousness and suggested a possible merging with Integrated Information Theory. We showed the terminological and conceptual affinities between the two theories and the solution that IIT offers in response to the hard-problem and to the question about whether dreaming has a specific evolutionary role, the strategy of starting from the phenomenal aspects of consciousness and their explanation in terms of integrated information. If our proposal proves to be theoretically robust, it might serve as a springboard for a more general theory of cognition that includes an explanation of consciousness. In the final section, we illustrated two topics for future research to focus on at the intersection of PP and dream studies. We think that they could provide mutual and helpful clarification to both fields.

References

Albantakis, L., Hintze, A., Koch, C., Adami, C. & Tononi, G. (2014). Evolution of integrated causal structures in animats exposed to environments of increasing complexity. *PLoS Comput Biol, 10* (12), e1003966.

Aserinsky, E. & Kleitman, N. (1953). Regularly occurring periods of eye motility, and concomitant phenomena, during sleep. *Science, 118* (3062), 273–274.

Casali, A. G., Gosseries, O., Rosanova, M., Boly, M., Sarasso, S., Casali, K. R., Casarotto, S., Bruno, M.-A., Laureys, S., Tononi, G. & Massimi, M. (2013). A theoretically based index of consciousness independent of sensory processing and behavior. *Science Translational Medicine, 5* (198), 198ra105. http://stm.sciencemag.org/content/5/198/198ra105.

Chalmers, D. J. (1996). *The conscious mind: In search of a fundamental theory.* New York: Oxford University Press.

Cirelli, C. & Tononi, G. (2015). Sleep and synaptic homeostasis. *Sleep, 38* (1), 161.

Clark, A. (2012). Dreaming the whole cat: Generative models, predictive processing, and the enactivist conception of perceptual experience. *Mind, 121* (483), 753-771. https://dx.doi.org/10.1093/mind/fzs106.

——— (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences, 36* (03), 181–204.

——— (2016). *Surfing uncertainty: Prediction, action, and the embodied mind.* New York: Oxford University Press.

Crick, F. & Mitchison, G. (1983). The function of dream sleep. *Nature, 304* (5922), 111–114.

Dennett, D. C. (1976). Are dreams experiences? *The Philosophical Review, 85* (2), 151–171.

———(2013). Expecting ourselves to expect: The Bayesian brain as a projector. *Behavioral and Brain Sciences, 36* (03), 209–210.

Dharani, N. E. (2005). The role of vestibular system and the cerebellum in adapting to gravitoinertial, spatial orientation and postural challenges of REM sleep. *Medical Hypotheses, 65* (1), 83–89.

Domhoff, G. W. (2001). A new neurocognitive theory of dreams. *Dreaming, 11* (1), 13–33.

———(2007). Realistic simulation and bizarreness in dream content: Past findings and suggestions for future research. In D. Barrett & P. McNamara (Eds.) *The new science of dreaming: Content, recall and personality correlates* (pp. 1–28). Westport, CT: Preager Press.

———(2011a). Dreams are embodied simulations that dramatize conception and concerns: The continuity hypothesis in empirical, theoretical, and historical context. *International Journal of Dream Research, 4* (2), 50–62.

——— (2011b). The neural substrate for dreaming: Is it a subsystem of the default network? *Consciousness and Cognition, 20* (4), 1163–1174.

Dołega, K. & Dewhurst, J. (2015). Curtain call at the Cartesian theatre. *Journal of Consciousness Studies, 22* (9-10), 109–128.

Flanagan, O. (1995). Deconstructing dreams: The spandrels of sleep. *The Journal of Philosophy, 92* (1), 5–27.

——— (2000). Dreaming is not an adaptation. *Behavioral and Brain Sciences, 23* (06), 936–939.

Foulkes, D. (1985). *Dreaming: A cognitive-psychological analysis.* Mahwah, NJ: Lawrence Erlbaum Associates.

——— (1999). *Children's dreaming and the development of consciousness.* Cambridge, MA: Harvard University Press.

Fox, K. C. R., Nijeboer, S., Solomonova, E., Domhoff, G. W. & Christoff, K. (2013). Dreaming as mind wandering: Evidence from functional neuroimaging and first-person content reports. *Frontiers in Human Neuroscience, 7*, 412.

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience, 11* (2), 127–138.

Friston, K., Mattout, J. & Kilner, J. (2011). Action understanding and active inference. *Biological Cybernetics, 104* (1-2), 137–160.

Giuditta, A. (2014). Sleep memory processing: The sequential hypothesis. *Frontiers in Systems Neuroscience, 8*, 219.

Giuditta, A., Ambrosini, M. V., Montagnese, P., Mandile, P., Cotugno, M., Zucconi, G. G. & Vescia, S. (1995). The sequential hypothesis of the function of sleep. *Behavioural Brain Research, 69* (1), 157–166.

Heller, C. (2014). The ups and downs of synapses during sleep and learning. *Sleep, 37* (7), 1157.

Hinton, G. E., Dayan, P., Frey, B. J. & Neal, R. M. (1995). The "wake-sleep" algorithm for unsupervised neural networks. *Science, 268* (5214), 1158.

Hobson, J. A. (2009). REM sleep and dreaming: Towards a theory of protoconsciousness. *Nature Reviews Neuroscience, 10* (11), 803–813.

Hobson, J. A. & Friston, K. J. (2012). Waking and dreaming consciousness: Neurobiological and functional considerations. *Progress in Neurobiology, 98* (1), 82–98.

——— (2014). Consciousness, dreams, and inference: The Cartesian theatre revisited. *Journal of Consciousness Studies, 21* (1-2), 6–32.

——— (2016). A response to our theatre critics. *Journal of Consciousness Studies, 23* (3-4), 245–254.

Hobson, J. A., Pace-Schott, E. F. & Stickgold, R. (2000). Dreaming and the brain: Toward a cognitive neuroscience of conscious states. *Behavioral and Brain Sciences, 23* (06), 793–842.

Hobson, J. A., Hong, C. C.-H. & Friston, K. J. (2014). Virtual reality and consciousness inference in dreaming. *Frontiers in Psychology, 5*, 1133.

Hohwy, J. (2013). *The predictive mind.* Oxford: Oxford University Press.

Malcolm, N. (1959). *Dreaming.* London: Routledge and Kegan Paul.

Massimini, M., Boly, M., Casali, A., Rosanova, M. & Tononi, G. (2009). A perturbational approach for evaluating the brain's capacity for consciousness. *Progress in Brain Research, 177*, 201–214.

Massimini, M., Ferrarelli, F., Murphy, M. J., Huber, R., Riedner, B. A., Casarotto, S. & Tononi, G. (2010). Cortical reactivity and effective connectivity during REM sleep in humans. *Cognitive Neuroscience, 1* (3), 176–183.

Nagel, T. (1974). What is it like to be a bat? *The Philosophical Review, 83* (4), 435–450.

Nielsen, T. A. (2000). A review of mentation in REM and NREM sleep: "covert" REM sleep as a possible reconciliation of two opposing models. *Behavioral and Brain Sciences, 23* (06), 851–866.

Nieminen, J. O., Gosseries, O., Massimini, M., Saad, E., Sheldon, A. D., Boly, M., Siclari, F., Postle, B. R. & Tononi, G. (2016). Consciousness and cortical responsiveness: A within-state study during non-rapid eye movement sleep. *Scientific Reports, 6*, 30932.

Nir, Y. & Tononi, G. (2010). Dreaming and the brain: From phenomenology to neurophysiology. *Trends in Cognitive Sciences, 14* (2), 88–100.

Noreika, V., Valli, K., Markkula, J., Seppälä, K. & Revonsuo, A. (2010). Dream bizarreness and waking thought in schizophrenia. *Psychiatry Research, 178* (3), 562–564.

Oizumi, M., Albantakis, L. & Tononi, G. (2014). From the phenomenology to the mechanisms of consciousness: Integrated information theory 3.0. *PLoS Computational Biology, 10* (5), e1003588.

Perogamvros, L. & Schwartz, S. (2012). The roles of the reward system in sleep and dreaming. *Neuroscience & Biobehavioral Reviews, 36* (8), 1934–1951.

——— (2013). Sleep and emotional functions. In P. Meerlo, R. M. Benca & T. Abel (Eds.) *Sleep, neuronal plasticity and brain function* (pp. 411–431). Springer.

Pliushch, I. (2017). The overtone model of self-deception. In T. Metzinger & W. Wiese (Eds.) *Philosophy and predictive processing.* Frankfurt am Main: MIND Group.

Revonsuo, A. (2000). The reinterpretation of dreams: An evolutionary hypothesis of the function of dreaming. *Behavioral and Brain Sciences, 23* (06), 877–901.

Revonsuo, A. & Tarkko, K. (2002). Binding in dreams–The bizarreness of dream images and the unity of consciousness. *Journal of Consciousness Studies, 9* (7), 3–24.

Revonsuo, A., Tuominen, J. & Valli, K. (2015). The avatars in the machine. In T. K. Metzinger & J. M. Windt (Eds.)

Open MIND. Frankfurt am Main: MIND Group. https://dx.doi.org/10.15502/9783958570375.

Sauvageau, A., Nielsen, T. A. & Montplaisir, J. (1998). Effects of somatosensory stimulation on dream content in gymnasts and control participants: Evidence of vestibulomotor adaptation in REM sleep. *Dreaming, 8* (2), 125.

Scarone, S., Manzone, M. L., Gambini, O., Kantzas, I., Limosani, I., D'Agostino, A. & Hobson, J. A. (2008). The dream as a model for psychosis: An experimental approach using bizarreness as a cognitive marker. *Schizophrenia Bulletin, 34* (3), 515–522.

Segawa, M. (1999). Ontogenesis of REM sleep. In B. N. Mallick & S. Inoué (Eds.) *Rapid eye movement sleep* (pp. 39–50). New York: Marcell Dekker, Inc.

Seth, A. K. (2013). Interoceptive inference, emotion, and the embodied self. *Trends in Cognitive Sciences, 17* (11), 565–573.

Solms, M. (2000). Dreaming and REM sleep are controlled by different brain mechanisms. *Behavioral and Brain Sciences, 23* (6), 843–850.

Stickgold, R. & Walker, M. P. (2013). Sleep-dependent memory triage: Evolving generalization through selective processing. *Nature Neuroscience, 16* (2), 139–145.

Stickgold, R., Hobson, J. A., Fosse, R. & Fosse, M. (2001). Sleep, learning, and dreams: Off-line memory reprocessing. *Science, 294* (5544), 1052–1057.

Tononi, G. (2008). Consciousness as integrated information: A provisional manifesto. *The Biological Bulletin, 215* (3), 216–242.

——— (2012). Integrated information theory of consciousness: An updated account. *Arch Ital Biol, 150* (2-3), 56–90.

——— (2015). Integrated information theory. *Scholarpedia, 10* (1), 4164.

Tononi, G. & Cirelli, C. (2014). Sleep and the price of plasticity: From synaptic and cellular homeostasis to memory consolidation and integration. *Neuron, 81* (1), 12–34.

Tononi, G. & Koch, C. (2015). Consciousness: Here, there and everywhere? *Phil. Trans. R. Soc. B, 370* (1668), 20140167.

Tononi, G., Boly, M., Massimini, M. & Koch, C. (2016). Integrated information theory: From consciousness to its physical substrate. *Nature Reviews Neuroscience*.

Valli, K. & Revonsuo, A. (2009). The threat simulation theory in light of recent empirical evidence: A review. *The American Journal of Psychology*, 17–38.

Von Hippel, W. & Trivers, R. (2011). The evolution and psychology of self-deception. *Behavioral and Brain Sciences, 34* (01), 1–16.

Windt, J. M. (2015). *Dreaming: A conceptual framework for philosophy of mind and empirical research.* Cambridge, MA: MIT Press.

Windt, J. M. & Noreika, V. (2011). How to integrate dreaming into a general theory of consciousness—A critical review of existing positions and suggestions for future research. *Consciousness and Cognition, 20* (4), 1091–1107.

Windt, J. M., Harkness, D. & Lenggenhager, B. (2014). Tickle me, I think I might be dreaming! Sensory attenuation, self-other distinction, and predictive processing in lucid dreams. *Frontiers in Human Neuroscience, 8*, 1–11.