

## MINORITY REPORTS: CONSCIOUSNESS AND THE PREFRONTAL CORTEX

Matthias Michel<sup>1</sup>, Jorge Morales<sup>2</sup>

1. Sciences, Normes, et Démocratie, Sorbonne Université, CNRS
2. Department of Psychological and Brain Sciences, Johns Hopkins University

Address for correspondence: Sciences, Normes et Démocratie, Sorbonne Université, 1 Rue Victor Cousin, 75005, Paris, France.

Email: [matthias.michel.curtill@gmail.com](mailto:matthias.michel.curtill@gmail.com), [jorge.morales@jhu.edu](mailto:jorge.morales@jhu.edu)

**Abstract:** Whether the prefrontal cortex is part of the neural substrates of consciousness is currently debated. Against prefrontal theories of consciousness, many have argued that neural activity in the prefrontal cortex does not correlate with consciousness but with subjective reports. We defend prefrontal theories of consciousness against this argument. We surmise that the requirement for reports is not a satisfying explanation of the difference in neural activity between conscious and unconscious trials, and that prefrontal theories of consciousness come out of this debate unscathed.

**Keywords:** Consciousness; neural correlates of consciousness; prefrontal cortex; no-report paradigms

## Introduction

The neuroscience of consciousness aims to find a set of brain mechanisms that explains how conscious experiences come about. Accordingly, one of the main goals in the field is to discover the neural correlates of consciousness (NCCs) (Crick & Koch, 1990), namely, neural states that are minimally sufficient for mental states to be conscious (Chalmers, 2000). One way to think about NCCs is as difference-makers that transform unconscious contents into conscious contents. In this article, we won't focus on how these difference-makers work, but in how to find them in the brain. In particular, we will focus on the neural correlates of conscious visual perception.

The question of whether the prefrontal cortex (PFC) is an NCC has become the epicenter of a debate between what we will call “prefrontal theories” and “posterior theories” of consciousness. Proponents of prefrontal theories, such as the global workspace theory (Baars, 1988; Dehaene & Changeux, 2011) and higher-order theories (Lau & Rosenthal, 2011; LeDoux & Brown, 2017), defend that prefrontal cortex (PFC) is an NCC.<sup>1</sup> Roughly, this means that some neural mechanisms in the PFC are responsible for rendering unconscious contents conscious, and that neural mechanisms elsewhere in the brain normally do not have this capacity. On the other hand, proponents of posterior theories defend that anterior parts of the cortex are not NCCs, but rather, that consciousness depends mainly on the activity of posterior parts of the cortex (Lamme, 2006) or of a “posterior hot-zone,” which includes roughly the entire cortex minus the insula and the PFC (Koch et al., 2016; Tononi et al., 2016).

Proponents of posterior theories of consciousness typically adopt two lines of argument against prefrontal theories. The first argument is that lesions to the PFC do not impair consciousness (Boly et al., 2017; Kozuch, 2014; Pollen, 2011). These criticisms have been recently addressed by Odegaard et al. (2017). Hence, we will set them aside in this article. The second argument is what we call the “argument from report”, according to which, the PFC is not an NCC but, rather, a neural correlate of *re-*

---

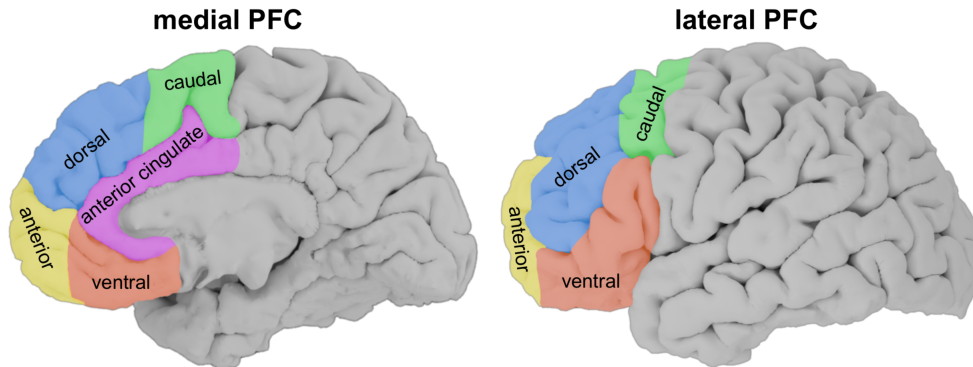
<sup>1</sup> These discussions often emphasize the existence of a main, or even unique, NCC (researchers often talk about *the* NCC). While this might turn out to be empirically correct, a problem with studying the neural basis of consciousness is that there might be more than one region or process responsible for rendering a state conscious. The brain often has redundant mechanisms that carry out the same function, and identifying or impairing an NCC might not immediately rule out the existence of secondary NCCs. We obviate this problem in what follows and talk about the PFC being *an* NCC rather than *the* NCC.

ports of consciousness (Koch et al., 2016; Storm et al. 2017; Tsuchiya et al., 2015). If this argument is sound, prefrontal theories would thus be based on a series of misunderstandings and methodological artefacts (Michel, 2017). In this paper, we defend prefrontal theories against the argument from report.

## 1. Prefrontal theories and the argument from report

### *a. The prefrontal cortex*

Having some grasp of the basic neuroanatomy of the PFC will be important in this article. In humans, the PFC constitutes a large portion of the frontal lobe that includes most of the cortical tissue anterior to the central sulcus. It can be decomposed into five main regions: the anterior prefrontal cortex (which includes the frontopolar cortex and part of the orbitofrontal cortex), the caudal prefrontal cortex (which includes the frontal eye fields on the lateral portion and the pre-supplementary motor area in the medial portion), the dorsolateral prefrontal cortex, the ventrolateral prefrontal cortex (also including part of the orbitofrontal cortex), and the medial prefrontal cortex (Figure 1).



**Figure 1. The human prefrontal cortex.** *Left:* Medial view of the right hemisphere showing the anterior cingulate cortex as well as medial portions of the anterior, dorsal, ventral, and caudal PFC. *Right:* Lateral view of the left hemisphere showing the dorsolateral, caudal, ventrolateral, and anterior prefrontal cortex. The cortical meshes used to make these images were created by Anderson Winkler (<https://brainder.org/brain-for-blender>) (distributed under a Creative Commons License CC BY-SA 3.0). The regions of interest were fake-colored by the authors following Passingham & Wise (2012).

Importantly, although the role of the PFC is not generally recognized as that of “processing sensory information”, the five main areas of the PFC are extensively connected with sensory areas. For example, the anterior PFC receives olfactory, gustatory, visceral, visual as well as somatosensory information (Carmichael & Price, 1995; Cavada, 2000), although it receives only few auditory inputs (Saleem et al., 2008). The caudal PFC is part of the so-called “dorsal visual stream”, which processes the spatial location of visual targets, and thereby receives visuo-spatial information (Milner & Goodale, 2008). Nevertheless, parts of the caudal PFC, such as the frontal eye fields, also receive information from the “ventral visual stream”, which processes colors, shapes, and textures of visual stimuli (Petrides & Pandya, 1999; Webster et al., 1994). The dorsolateral PFC receives inputs from multimodal areas, areas involved in the identification of objects, as well as somatosensory areas (Petrides & Pandya, 1999, 2002). The ventrolateral PFC receives visual as well as auditory information due to its extensive connections with the inferior and superior temporal cortices, and it also receives somatosensory information (Petrides & Pandya, 2002; Webster et al., 1994). Finally, the medial PFC also receives auditory and visual information from the superior temporal sulcus and superior temporal cortex (Kondo et al., 2003, 2005).

Hence, whereas the PFC is likely not responsible for *producing* sensory contents, all of its subcomponents receive sensory information. Consequently, sensory information is distributed throughout the PFC such that its subcomponents can achieve their specific functions (Passingham & Wise, 2012). However, our understanding of the mechanisms in the PFC through which all that sensory information could become *conscious* is still limited. As a result, it is important to note that conscious sensory representations could be *distributed* throughout the PFC, rather than represented in a single sub-part of the PFC, although this possibility remains open.<sup>2</sup> The fact that conscious information could be distributed throughout the PFC justifies talking about “the PFC”, in general, rather than only some of its specific components, as being involved (or not) in consciousness. It is important to recognize, however, that talking about the link between consciousness and the entire PFC, as opposed to some of its specific subcomponents, is more the result of our ignorance of the specific mechanisms in the PFC that could be involved in consciousness rather than a strong thesis about the nature of conscious representations in the PFC.

---

<sup>2</sup> The distributed nature of the functional organization in PFC probably applies to many cognitive functions and not just to consciousness.

It is worth noting that some of the evidence regarding PFC connectivity discussed above, as well as some of the evidence of the role of PFC in consciousness that we will present below, stems from animal studies—macaque monkeys specifically. Human and macaque brains in general, and PFC in particular, have numerous anatomical and functional similarities, which makes the macaque brain an adequate model for understanding human physiology and cognition (Passingham, 2009). For example, the types of neural tissue and their anatomical connectivity to other regions within PFC are largely comparable across both species (Petrides et al., 2012). The macaque PFC also has multiple functional networks similar to those in the human PFC. For example, functional magnetic resonance imaging (fMRI) has revealed similar resting-state networks in humans and macaques that functionally connect dorsolateral, ventrolateral, and ventromedial PFC with other areas of the brain (Mantini et al., 2013; Neubert et al., 2014).

Despite these similarities, we should acknowledge important differences between the human and macaque PFC. The human brain, and this is particularly true of PFC, is not just a scaled version of the macaque brain (Rilling, 2006). The macaque brain, when warped to fit a human brain, requires larger expansions in temporal, parietal, and prefrontal cortex (van Essen et al., 2007). Neuronal tissue in some regions in PFC differ across the two species. For instance, the dorsal anterior cingulate cortex (also known as BA32) probably has no homologue in the macaque brain (Öngür et al., 2003). Functional connectivity differences exist too. The lateral frontopolar cortex in humans has a connectivity profile that has no counterpart in the macaque prefrontal cortex (Neubert et al., 2014). Similarly, networks connecting dorsolateral PFC and parietal cortex, as well as anterior cingulate cortex and the insula, have been found only in humans (Mantini et al., 2013).

Notwithstanding the aforementioned anatomical and functional differences between macaque and human brains, and as we already noted, our understanding of possible PFC mechanisms supporting conscious experiences is still rudimentary, and so these differences do not affect our main argument.

*b. The argument from report*

To discover NCCs, scientists rely on a method called “contrastive analysis” (Baars, 1988), which consists in comparing neural activity from trials in which subjects consciously perceive a stimulus with trials in which they

perceive it unconsciously. As such, contrastive analysis requires experimenters to know whether participants were conscious of the displayed stimulus or not. A straightforward way of acquiring this information is by asking subjects to report whether they consciously perceived the stimulus or not. This approach is called a “report-based” experimental method (Irvine, 2013). Report-based methods encompass a wide variety of procedures, such as confidence ratings (Cheesman & Merikle, 1986), reports on the visibility of the stimulus (Sergent & Dehaene, 2004), reports using the Perceptual Awareness Scale (Ramsøy & Overgaard, 2004), and post-decision wagering (Persaud et al., 2007).

The specific differences between these methods will not be central here. What matters is what is common to all report-based methods, namely, the fact that subjects are required to make reports from which experimenters can infer subjects’ conscious perceptual states. Most experiments in consciousness science use report-based methods, and it would negate most scientific work in this field if this strategy turned out to be deeply flawed. Several authors think that this is the case (Boly et al., 2017; Koch et al., 2016; Pitts et al., 2014), and that consciousness science should turn instead to experimental paradigms that do not require subjects to produce reports (Tsuchiya et al., 2015). We now explain why we find these worries to be unwarranted.

The general structure of experiments using the contrastive analysis method is as follows: experimenters present visual stimuli that, through a variety of methods, are rendered invisible for the subject on some trials. Subjects’ task is to detect the stimulus and report seeing or not seeing it (typically by pressing a button), or to make a visual discrimination and subsequently rate their level of confidence in that visual discrimination. These tasks are performed while subjects’ brain activity is measured, typically by using fMRI or electroencephalography (EEG). Experimenters can then identify the brain activity that best correlates with trials in which subjects reported consciously seeing a stimulus and the brain activity that best correlates with trials in which they did not. A comparison of these two kinds of activity should reveal the NCC.<sup>3</sup>

Over and over again, researchers have found that conscious perception elicits increased activity in the PFC while unconscious perception

---

<sup>3</sup> Besides conscious versus unconscious seeing, experimenters need to control for differences in *perceptual processing* such as perceptual signal strength, performance, and attention, so that consciousness is the *only* difference between the two contrasted conditions (Morales et al., 2015).

does not. The PFC has thus been considered an NCC. Many studies support this hypothesis, using a wide variety of experimental paradigms to render stimuli invisible, such as stimulus masking (e.g., Dehaene et al., 2001; Del Cul et al., 2007; Fisch et al., 2009; Lau & Passingham, 2006), the attentional blink (Sergent et al. 2005; Williams et al., 2008), or binocular rivalry (Lumer & Rees, 1999; Sterzer et al., 2007; Weilhhammer et al., 2013). These results are also replicated across several sensory modalities (Sadaghiani et al., 2009; De Lafuente & Romo, 2005).

Nevertheless, and despite the abundant evidence, proponents of alternative theories, such as the integrated information theory (Tononi et al., 2016) or the local recurrency theory (Lamme, 2006), do not consider that the PFC is a central mechanism for conscious perception. Rather, they argue that the NCCs are “in the back” of the cerebral cortex, and not “in the front” (Boly et al., 2017). But, how do they account for the large body of research that finds PFC activity only when stimuli are consciously perceived? What is supposed to explain this difference in brain activity between seen and unseen trials, if not consciousness? One of the main arguments against PFC theories of consciousness is that activity in the PFC does not correlate with consciousness itself, but with cognitive processes following conscious perception (Aru et al., 2012; de Graaf et al., 2012). On this view, asking subjects to report their awareness of the stimulus introduces a confound: purported NCCs are “overestimated, in the sense that neural mechanisms not directly generating the experience, but only necessary to report it, are included” (Tsuchiya et al. 2015, p. 758).

As a result, in its simplest form, the so-called “argument from report” states that subjects’ explicit reports of awareness (and unawareness) should be eliminated from contrastive analyses that search for the NCCs. On this view, the contrast should be between conditions in which subjects consciously see a stimulus and conditions in which they do not without requiring that subjects report back in any way whether they were aware of the stimulus or not. Naturally, this restriction requires that the experimenter knows how to classify each trial (aware vs unaware), which requires establishing a previous link between certain types of stimulation and awareness (perhaps via reliable psychophysical manipulations or autonomic biophysical markers that correlate with reports of awareness). By eliminating reports, the argument continues, the difference in neural activity between aware and unaware conditions should reveal the NCCs *and only* the NCCs.

Indeed, producing a report requires participants not only to engage motor-planning neural circuits, but also to maintain and access infor-

mation about whether they saw the stimulus or not. Moreover, because stimuli in this kind of experiments are often at threshold, or otherwise hard to perceive, subjects need to evaluate the nature and content of their experiences. To the extent that formulating a report involves the subject's ability to evaluate her own perceptual states, the requirement for reports also engages the subject's metacognitive capacities. The problem is that metacognition, that is, the ability to monitor and control one's perceptual and cognitive states, is itself a complex cognitive capacity largely supported by neuroanatomy and neural activity in the PFC, particularly in the anterior, medial, and dorsolateral PFC (Fleming et al. 2010; McCurdy et al 2013; Morales et al 2018; Fleck et al., 2006; Hebart et al 2016); but this activity might not itself be an NCC. Scientists would thus be mistaking the neural correlates of consciousness with the neural activity involved in *reports of* consciousness (Michel, 2017).

A seminal experiment by Dehaene and colleagues (2001) could provide a good case study of how this argument is supposed to work. In this experiment, visually presented words were rendered invisible by using a method called visual masking, in which a target stimulus becomes invisible in some trials when a second stimulus (a "mask") is presented quickly (<100ms) at the same location. In this experiment, subjects had to report the word whenever they thought a word had appeared. Dehaene et al. observed that trials in which words were detected elicited an increase in PFC activity, especially in the ventrolateral PFC, whereas trials in which words were not detected did not. The problem with this experiment is that consciousness of the word was not the only changing variable between seen and unseen trials. When subjects did not see the stimulus, they did not have to report anything, but when they saw it, they had to report the identity of the stimulus. Hence, the report itself introduced a confound and could explain the increase in PFC activity on seen trials.

Nonetheless, whereas a simple version of the argument from report could successfully be applied in this particular example, in most experiments the increase in PFC activity on conscious trials is unlikely to be due to reports. In the vast majority of subsequent studies, subjects have been required to make a report *both* on conscious *and* unconscious trials (for reviews, see Dehaene et al. 2014; Koch et al. 2016; Odegaard et al. 2017). If the requirement to make a report were the sole driver of PFC activity increases during these experiments, one should expect to see it every time subjects produce a report, namely, during conscious as well as unconscious trials. This would effectively cancel out any detectable increase in PFC neural activity when comparing these trials. This is not the case. What is



repeatedly observed is an increase in PFC activity during conscious trials only. Consequently, increases in PFC activity are unlikely to be due to reports themselves. Hence, proponents of the argument from report have to explain why there is no increase in PFC activity on trials in which subjects *report* not seeing the stimulus. To do so, they have to provide a slightly more complex version of the argument.

Proponents of the argument from report could avoid this problem by claiming that different cognitive capacities are involved during trials in which subjects report seeing the stimulus compared to trials in which they report that they do not. As we will explain in more detail below, those cognitive capacities are thought to be, essentially, post-perceptual demands related to decision-making and to maintaining and accessing stimulus-specific information in working memory. On this more complex version of the argument, the increase in PFC activity during conscious trials is explained by the engagement of cognitive capacities to report seeing the stimulus. The observed decrease in PFC activity during unconscious trials is explained by the fact that those cognitive capacities are not engaged when subjects report that they do not see the stimulus.

In sum, the argument from report is successful only if changes in PFC activity between aware and unaware trials are best explained by the engagement or disengagement of a *set of report-related cognitive capacities* than by the presence or absence of *consciousness*. As a result, proponents of the argument from report make two crucial predictions.

First, a specific set of report-related cognitive capacities should be engaged during conscious trials *only*, and not during unconscious trials. However, if it turns out that essentially the *same* cognitive capacities are used by subjects *both* when they report seeing a stimulus, and when they report not seeing a stimulus, then the engagement of those capacities is ill-suited to explain the difference in PFC activity between conscious and unconscious trials. In Section 2, we argue that this prediction is not verified. On our view, there is little reason to believe that “seen” and “unseen” reports involve fundamentally different cognitive capacities. Consequently, the difference in PFC activity between conscious and unconscious trials cannot be explained by the fact that subjects use different report-related cognitive capacities during seen and unseen trials.

Second, proponents of the argument from report also predict that there should be no, or negligible, differences in PFC activity between unconscious and conscious trials when subjects do not have to provide reports. This is for a simple reason: if the engagement of report-related cog-

nitive capacities explains the increase in PFC activity during conscious trials only, suppressing the need for a report should also cancel the observed difference in PFC activity between conscious and unconscious trials. In Section 3, we argue that this prediction is inconsistent with the available data: there are differences in PFC activity between conscious and unconscious trials even when subjects do not have to provide reports.

## 2. PFC activity with reports

Proponents of the argument from report, such as Pitts et al. (2014), recognize that “regardless of whether the stimulus was consciously perceived, the subject always has to report something even if the report is ‘I saw nothing’” (p. 2). Pitts and colleagues provide a *prima facie* plausible argument for explaining the observed difference in PFC activity. They propose that a crucial difference between “seen” reports and “unseen” reports is that there is nothing for the subject to report in unconscious trials. Thus, subjects do not have to *remember* the identity of the stimulus by encoding and maintaining a perceptual representation of the stimulus in working memory on unconscious trials:

On aware trials, subjective reports rely on the maintenance of perceptual information in working memory and access of this information by higher-level cognitive systems that enable decision-making and response planning/execution. On unaware trials, there is no conscious perceptual information to maintain or access even though a decision must be made and a negative response must be planned and executed. Thus, in addition to differences in conscious perception, these two types of trials differ in terms of post-perceptual processing such as maintenance in working memory and access of perceptual information for decision-making. (Pitts et al., 2014, p. 2)

As already suggested above, it is important to remark that the simple name “argument from report” might be misleading when we consider more complex versions of this argument, such as Pitts et al.’s. Indeed, proponents of this view often insist that the critical contrast between conscious and unconscious trials should be restricted to passive viewing conditions where, besides eliminating subjects’ explicit reports, post-perceptual demands related to decision-making and to maintaining and accessing stimulus-specific information in working memory are also absent. Thus, the “argument from report” might just as well be relabeled the “argument from report, working memory, and decision-making” as

this would capture better its specificity. For simplicity, and to keep the terminology in line with the literature, we use the shorter label in what follows.

On this version of the argument from report, PFC activity on conscious trials is best explained by the maintenance of stimulus-specific perceptual information in order to report the identity or location of the stimulus. For example, in the aforementioned experiment by Dehaene et al. (2001), subjects not only had to remember whether they saw the stimulus or not, but also had to remember *the identity* of the stimulus. In this case, there was a crucial difference between seen and unseen trials: in seen trials, subjects had to maintain a perceptual representation of a word in working memory, whereas in unseen trials they did not have to do so. This difference between seen and unseen reports could explain that an increase in PFC activity was found only on seen trials. Hence, on this version of the argument from report, the maintenance of stimulus-specific perceptual information in working memory to generate a report is the best explanation of the difference in brain activity between seen and unseen trials. And, indeed, the ability to maintain behaviorally relevant sensory information for a brief amount of time has long been associated with activation of PFC neurons (e.g., Funahashi et al., 1989; Miller & Cohen, 2001). However, there are three problems with this version of the argument from report.

*a. Unconscious information about the stimulus is maintained in the brain*

The first problem is that the claim that “on unaware trials, there is no conscious perceptual information to maintain or access” is contentious. Indeed, there is evidence that information about identity or location of stimuli can be maintained in working memory for several seconds during *unseen* trials (e.g., King et al., 2016; Soto et al., 2011; Trübtschek et al., 2017). This makes the maintenance of stimulus-specific perceptual information in working memory an unlikely candidate to explain the difference in brain activity between seen and unseen trials, because visual information could very well be maintained in working memory during unconscious trials too.

For example, in a study by Salti et al. (2015), participants had to detect a stimulus and indicate its location. Backward masking was used to render the stimulus invisible on roughly sixty percent of the trials, but participants still had to indicate the location of the stimulus on *every* trial. Participants were able to detect the location of the stimulus well above

chance even when the stimulus was categorized as unseen. Salti et al. trained pattern classifiers on EEG and magnetoencephalography (MEG) data to assess where in the brain information about the stimulus was represented in both seen-correct and unseen-correct trials. They found that the location of the stimulus could be decoded up to 800 milliseconds, independently of whether the stimulus was seen or not. Importantly, locations of *unseen* stimuli could be decoded from signals in the dorsolateral PFC (superior frontal regions). Since maintenance of information in working memory has been associated with PFC activity, the fact that the dorsolateral PFC encoded stimulus location during unseen trials could indicate that the maintenance of unseen locations involved working memory. Hence, this could suggest that information about the stimulus was maintained in working memory in both seen and unseen conditions. Despite this, activity patterns, particularly in the dorsolateral PFC, were classified more effectively in seen-correct compared to unseen-correct trials. Here, the maintenance of visual information in working memory is unlikely to explain the difference in brain activity patterns because information about the stimulus was maintained in both seen and unseen conditions. Moreover, this information was used to correctly detect the location of the target, which means that it must have been accessed by cognitive systems in both seen and unseen conditions to generate the correct response. Hence, even when performance is the same (i.e., subjects correctly report the stimulus location) and involves the same cognitive capacities, namely, maintenance of visual information in working memory, differences in PFC activity patterns between seen and unseen conditions are still observed.

*b. Maintaining stimulus-specific information in working memory is not required*

The second problem with this version of the argument from report is that it relies on several contentious assumptions about the types of information that need to be maintained in working memory across different kinds of tasks.

To understand these assumptions, it is important to introduce a set of distinctions, beginning with the difference between *identification* tasks and *detection* tasks. During identification tasks, subjects have to provide a report about some features that are specific to a target stimulus, such as its identity, location, or other specific perceptual features such as its color or orientation. On the other hand, during detection tasks, subjects have to

provide a report about the presence or absence of a stimulus, with no need to consider its specific perceptual features.

Beside this task difference, we should also distinguish between several types of information that could be maintained in working memory when subjects perform these tasks. First, subjects could maintain *stimulus-specific information*, namely, *perceptual information*, such as perceptual features of the stimulus, its identity or its location. Second, subjects could maintain a *perceptual decision* in addition to, or instead of, maintaining stimulus-specific information. For example, they could maintain in working memory the decision that the stimulus was present or absent, the decision that one saw the stimulus or not, or the decision that a stimulus was a square, or that it was located on the right side of the screen. Note that when subjects maintain a perceptual decision, they do not need to encode perceptual information about the stimulus in a perceptual format. For instance, they can maintain *that* they saw or did not see a square without thereby perceptually encoding any specific shape, size, color, identity, or location. With these distinctions at hand, we can now analyze the argument from report further.

The first assumption made by proponents of the report argument is that subjects store *stimulus-specific perceptual information* in working memory to perform *identification tasks*. If that is true, as Pitts et al. argue, the demands imposed on working memory during seen trials are higher than during unseen trials, because there is (on their view) no stimulus-specific perceptual information to maintain when the stimulus is not consciously seen. However, contrary to Pitts et al.'s assumption, subjects do not need to encode stimulus-specific information in working memory to perform these tasks. Instead, they could directly maintain a perceptual decision about their awareness or unawareness of the stimulus.<sup>4</sup> If that's the case, the same type of information is maintained in both conscious and unconscious trials during identification tasks. That is, even when subjects do not see the stimulus, they still have to maintain the perceptual decision *that they did not see the stimulus*. In that case, working memory demands are matched between unconscious and conscious trials since, in both cases, subjects maintain the same type of information in working memory.

Nevertheless, even if working memory demands were different between seen and unseen trials in *identification tasks*, this is certainly not the case for *detection tasks*. In experiments using *detection* tasks, subjects are not required to remember and report properties that are *specific* to the

---

<sup>4</sup> We thank an anonymous reviewer for suggesting this argument.

stimulus, such as *what* the stimulus is, or *where* it is, but simply have to report seeing something or not. As such, it seems plausible that during detection tasks subjects do not maintain *perceptual information* in working memory, since they do not need to use stimulus-specific perceptual information to perform the task, but, instead, maintain the *perceptual decision* of having seen something or not (or the perceptual decision that the stimulus was present or not). Importantly, maintaining a perceptual decision in this way does not require one to maintain any perceptual information about the specific properties of the stimulus. Consequently, whereas in identification tasks subjects have to maintain stimulus-specific information on seen trials only, in detection tasks, the same type of information is maintained in both seen and unseen trials. As such, post-perceptual processing is matched between the seen and unseen conditions, provided that the working memory demands to maintain a perceptual decision of having seen something (or that the stimulus was present) are not much more important or involve fundamentally different processes than those required to maintain a perceptual decision of having seen nothing (or that the stimulus was absent).

Let us use a recent experiment by van Vugt et al. (2018) as a case study to show why the argument from report fails to account for the difference in brain activity between seen and unseen conditions in detection tasks. van Vugt et al. trained monkeys to detect very low-contrast stimuli displayed for 50ms. 450ms after the presentation of the stimulus, monkeys had to make a saccade towards the location of the stimulus when they detected it, or towards a rejection dot when they did not. Crucially, the *location* of the stimulus remained the same across all trials, and monkeys did not have to report on the *identity* of the stimulus. So, they did not have to actively maintain stimulus identity or location in working memory, but only had to maintain, during 450ms, whether they saw something or not (that is, whether they had to saccade to one location or another). Moreover, monkeys were extensively trained to perform this task, allowing them to perform it quite automatically, thereby reducing any residual load on working memory. In this experiment, dorsolateral PFC activity increased during trials in which monkeys were aware of stimuli and did not increase during unconscious trials (i.e., when they failed to saccade to the stimulus's location when a stimulus was presented). Proponents of the argument from report would have to explain why very similar cognitive processes such as, on one hand, maintaining and accessing seeing something, and on the other hand, maintaining and accessing seeing nothing, would recruit *fundamentally different brain areas*,

i.e., the PFC for “seen” reports, and posterior areas for “unseen” reports. It seems implausible that such a small difference in reports could explain the difference in brain activity between seen and unseen trials in those simple detection tasks. In contrast, consciousness of the stimulus could be a good explanation of this difference in brain activity.

Two important studies allow us to push this argument further. First, it seems that “seen” and “unseen” reports rely on similar mechanisms in the brain. Merten and Nieder (2012) recorded activity from randomly selected neurons in the lateral PFC in monkeys during a detection task. A stimulus with varying intensities was first presented for 100ms, and monkeys had to maintain their perceptual decision during a delay period before making a report. Importantly, during this period, monkeys did not know how they would have to report seeing or not seeing the stimulus, thus preventing motor planning. Merten and Nieder observed that, *during the delay period*, groups of neurons in the lateral PFC specifically encoded the “seen” decision, while other groups of lateral PFC neurons encoded the “not seen” decision, confirming that reports of seeing and not seeing stimuli rely on similar mechanisms in the PFC. However, *during the presentation of the stimulus*, only “seen” neurons discharged when the stimulus was presented, while “unseen” neurons maintained baseline activity whether or not the stimulus was presented. This result is at odds with the idea that neural activity in the PFC represents only *the maintenance* of “seeing the stimulus” in order to make a report. Indeed, if that were the case, we should expect to observe “seen” and “unseen” neurons active only during the delay period. Instead, “seen” neurons discharged both during stimulus presentation and during the delay period (when the stimulus was seen), whereas “unseen” neurons discharged only during the delay period when the stimulus was unseen.<sup>5</sup> Following the authors, this could indicate

two discrete processing steps involved in abstract decisions in detection tasks, implemented by “yes” cells in the stimulus phase and by

---

<sup>5</sup> Moreover, in this experiment, specific groups of neurons in the lateral PFC modulated their activity depending on the intensity of seen stimuli during the stimulus phase, despite the fact that monkeys did not have to report on the intensity of stimuli. Qualitative aspects of stimuli thus seem to be represented in the PFC despite the fact that these aspects are not task-relevant and do not have to be maintained for the report. This result is in line with a study by Mante et al. (2013) showing that stimuli that are irrelevant for a perceptual task (i.e., stimuli that need not be maintained for a report) nonetheless lead to neural activity in the PFC.

“yes” and “no” cells during the delay phase. During the stimulus phase, the responses might reflect the subjective experience of the stimulus, based on the accumulation of sensory information. The emergence of “no” neurons in the delay phase likely constitutes a second active decision-processing step transforming the subjective experience to abstract categories in rule-based detection tasks. (p. 6292)

Second, a study by de Lafuente and Romo (2005) further supports the view that activity in the PFC does not primarily reflect the report decision. Neurophysiological recordings in macaque monkeys during a tactile detection task demonstrate a strong correlation of trial-by-trial subjective reports (button pressings) with late frontal activity, specifically activity in the pre-supplementary motor area (itself densely connected with dorsolateral and dorsomedial prefrontal cortex (Luppino et al., 1993)), whereas activity in posterior areas (somatosensory cortex) is correlated with physical properties of the stimulus and not with perceptual decisions. Monkeys were not required to remember or maintain any perceptual stimulus-specific information, but only to detect whether the stimulus was present or not. Here again, proponents of the argument from report could argue that late frontal activity reflects the report decision, and not consciousness of the perceptual content. However, de Lafuente and Romo designed a control task in which the correct response button was illuminated at the beginning of every trial, which means that monkeys already knew which button they would have to press *before* the stimulus was presented. This implies that, in this control task, *the decision to report was already made before stimulus presentation*. If neural activity in the PFC was only linked to the decision to *report* seeing the stimulus or not, then, those neurons should discharge *before* the stimulus was presented, namely, at the moment when the decision to report was actually made. On the contrary, de Lafuente and Romo found that neural activity was nearly identical in the standard detection task and in the control task (see also de Lafuente & Romo (2006)). In accordance with Merten and Nieder’s results, this strongly suggests that recorded neurons in the PFC discharging during stimulus presentation represent the perceptual decision that a stimulus was seen or unseen, and not only the decision to *report* seeing a stimulus or not.



*c. Subjective reports are often about performance or visibility, not about stimuli themselves*

So far, we have argued that the argument from report is unlikely to work, first, because information about stimuli could be maintained in both seen and unseen trials; second, because stimulus-specific information does not need to be maintained during detection tasks—and possibly not even during identification tasks. But even when one sets aside detection tasks, subjects are not usually required to maintain any stimulus-specific information for subjective reports. They only have to maintain information about their *performance* on a discrimination task or *on the visibility* of the stimulus,<sup>6</sup> regardless of whether the stimulus was seen or unseen. For example, in a seminal study by Lau and Passingham (2006), participants had to make a forced-choice judgment as to whether a masked stimulus was a square or a diamond, and then report the degree of visibility of the stimulus. Lau and Passingham found that mid-dorsolateral PFC activity increased in seen trials and not in unseen trials, even when performance on the discrimination task was matched, thus making sure that activity in this region of the PFC did not reflect performance on the discrimination task. In this experiment, although the first-order discrimination involved maintaining stimulus-specific information (in *both* seen and unseen conditions), the subjective report itself did not involve maintaining any stimulus-specific information. Here again, the only information that needed to be maintained for the subjective report throughout the trial was one’s perceptual decision of having seen the stimulus or not. The only way in which the argument from report would work in this and similar cases is by presupposing that maintaining a perceptual decision of having seen a stimulus involves brain mechanisms that are fundamentally different from

---

<sup>6</sup> The property *being visible* is not a stimulus-specific property, because it could be a property of *any* stimulus. To this extent, maintaining a visibility rating is not different from maintaining a decision, say, about one’s performance (e.g., a confidence rating). Nevertheless, one could still worry that visibility ratings entail different demands in the “seen” and “unseen” cases *while* making the perceptual decision, even if the decision *per se* can be maintained in a similar fashion without any further stimulus-specific information. However, if there is an introspective difference between seeing something and not seeing anything *while* having (or not having) the experience, it cannot be said, like Pitts et al. do, that the difference is “post-perceptual”. While making the decision whether the stimulus was visible or not, the introspective effort should be matched between unaware and aware cases, and the differences should be limited precisely to the variable of interest: consciousness. This, however, does not entail that the post-perceptual requirement of reporting stimulus visibility *after* stimulus presentation is distinct for “aware” and “unaware” cases.

those that are used to maintain a perceptual decision of *not* having seen a stimulus. But we have no reason for thinking that this is the case.

### 3. PFC activity without reports

The results reported above could appear to be in stark contrast with results from so-called “no-report experimental paradigms” (Tsuchiya et al., 2015; Koch et al., 2016). But two facts are not emphasized nearly enough in this debate (although see Lew & Lau, 2017). First, proponents of the argument from report generally ignore, or repress maybe, that in the majority of experiments that do not require reports, researchers still observe neuronal activity in the PFC exclusively on conscious trials. Second, a few null findings from fMRI studies are often overemphasized without proper acknowledgement of the limited sensitivity of this imaging technology. Here we address studies that use binocular rivalry, masking, and dreams to probe neural activity associated with consciousness without report.

#### *a. Binocular rivalry*

Binocular rivalry is a visual phenomenon during which an image presented to one eye becomes unconscious because it competes with a rival and incompatible image presented to the other eye, causing conscious perception to alternate between the two images every few seconds. Frässle et al. (2014) compared neural activity elicited by binocular rivalry switches in report and no-report conditions. In this experiment, subjects experienced binocular rivalry under two conditions. In the report condition, subjects had to report the switch from one image to the other. In the no-report condition, the switch was inferred by observing reflexes such as pupil dilation and ocular micro-saccades, which have been observed to occur when perceptual experience alternates during binocular rivalry (Einhäuser et al., 2008; Naber et al., 2011). They concluded that dorsolateral prefrontal cortex was activated during the perceptual switch in the report condition, but not in the no-report condition.

Several caveats need to be addressed when considering binocular rivalry experiments. First, it should be remarked that whether the PFC is involved in the switch between two images during binocular rivalry is unimportant for the debate, because the perceptual switch in itself may have nothing to do with consciousness. In fact, Zou et al. (2016) showed that binocular rivalry occurs between two invisible dichoptic gratings, thus demonstrating that the perceptual switch in binocular rivalry is inde-

pendent of the mechanisms of consciousness (see also Giles et al., 2016). On the other hand, what is important is whether the PFC is involved in both consciousness of the switch, and in consciousness of contents once the switch has occurred. It seems, indeed, that activity in the PFC is involved in consciousness of the perceptual switch. In an ingenious experiment, Brascamp et al. (2015) induced perceptual alternations presenting different yet indistinguishable displays of randomly moving dots to each eye, such that alternations were unnoticed by participants. They found that, when the switch between the two images goes unnoticed, activity in the PFC does not increase, whereas consciousness of the switch involves PFC activity, especially in the frontal eye fields, ventro- and dorsolateral prefrontal cortex. They conclude that fronto-parietal regions do not have a “driving role... in perceptual switches”, but that they could have “a central role in visual awareness” (p. 1667). All things being equal, without any change in conscious experience,<sup>7</sup> one shouldn’t expect any change in PFC activity. That’s precisely what Brascamp et al. have found.

Putting aside these methodological considerations, contrary to Frässle and colleagues’ explicit interpretation, some prefrontal areas in the superior frontal gyrus and the inferior frontal gyrus remained clearly active even during the no-report condition (these activations are visible in their Fig. 3; see Zaretskaya and Narinyan, 2014). This is not surprising, because multiple studies using binocular rivalry have consistently found that conscious perception correlates with PFC activity whereas unconscious perception does not (see Brascamp et al., (2018) for a review). More important for the current discussion, Lumer & Rees (1999) designed a binocular rivalry experiment in which subjects saw the stimuli freely and were not instructed to make any motor or verbal report to indicate perceptual transitions. During subsequent debriefing, subjects confirmed that they experienced visual alternations. Lumer & Rees found that activity in the lateral PFC, and particularly in the ventrolateral PFC, was temporally correlated with activity in early extrastriate visual cortex, which had been previously associated with consciously experiencing alternating images in binocular rivalry paradigms. In contrast, activity in the PFC was not correlated with activity in striate cortex (V1), which is typically not considered relevant for consciously perceiving alternating images. Importantly, the PFC was distinctively coupled with visual areas despite the

---

<sup>7</sup> One could argue that *there was* a change in phenomenal consciousness, but that this change was not accessed (Block, 1995). At present, however, we cannot determine whether this is true or not. For all we know, the data suggests that subjects’ perceptual experiences did not change when the switch occurred, and thus, we shouldn’t expect a change in the NCC.

fact that no report was required. Moreover, when compared with a condition in which no conscious alternation of images was experienced (because the images monocularly presented were made compatible and thereby rendered stable), the PFC was more coupled with extrastriate visual areas, highlighting its role in experiencing conscious perceptual shifts independently from report (for similar results, see also Wilcke et al., 2009). Using binocular flash suppression, a method similar to binocular rivalry in which an image presented in one eye is suppressed by another image flashed in the other eye, Panagiotaropoulos et al. (2012) arrived at the same conclusion: by using single unit neuron recordings in monkeys, they confirmed that lateral PFC activity increased for consciously perceived stimuli even when no report was required.

*b. Masking*

In a study by Tse et al. (2005), subjects were presented with simple target stimuli that were preceded and followed by masks, causing a perceptual illusion called the “standing wave of invisibility”. When the masks were presented in close temporal proximity to the target, it was rendered invisible, but it remained visible when the masks were presented with a short delay. Importantly, subjects were not asked to report the visibility of the targets at any time while inside the scanner. Rather, the paradigm was validated using indirect visibility ratings previously collected from five subjects. By using univariate fMRI analyses, Tse and colleagues found that only regions in visual cortex were more active when the visible (delay-masked) and invisible (immediately-masked) conditions were compared.

This result invites us to reflect on the problems surrounding sensitivity of fMRI measures and the study of the PFC. We ought to be cautious about null findings like the one reported by Tse and colleagues. Failing to detect a statistically significant difference in neural activity between conscious and unconscious conditions does not immediately warrant the inference that a brain region is not involved in sustaining consciousness. To put this simply, absence of evidence is not immediate evidence of absence.<sup>8</sup> This is especially true when using low sensitivity neuroimaging

---

<sup>8</sup> To state the statistically obvious: it is a fallacy to consider that a failure to reject the null hypothesis means that the null hypothesis is true. Those who interpret Tse et al. as indicating that PFC is not involved in consciousness commit this fallacy. Roughly, statistical significance as assessed by p-values protects against Type I errors (rejecting the null hypothesis when it is in fact true), but it does not

techniques such as fMRI, which poses special difficulties in detecting and decoding meaningful neural activity in PFC related to consciousness.

Univariate analyses of fMRI data where the overall level of activity is compared across conditions are most useful for detecting large effects. However, the difference in neural activity between conscious and unconscious conditions is often subtle compared to the kind of activity found in sensory and motor cortex. This subtlety is evidenced by the fact that neural activity related to consciousness that is missed in fMRI univariate analyses is often detectable with more sensitive methods. Indeed, the lack of sensitivity of fMRI could explain the discrepancy in the results between no-report studies using fMRI to measure brain activity, and no-report studies that use other procedures (Morales and Lau, Forthcoming). For example, the limited spatio-temporal sensitivity of fMRI measures would explain why lateral PFC activity related to consciousness is detected in passive viewing conditions when using single-unit recordings in macaques (Panagiotaropoulos et al., 2012).

This interpretation of the data is also supported by several human studies that have found PFC activity during conscious trials without reports when more sensitive measurement techniques were employed. For example, electrocorticography (ECoG) allows placing electrodes directly on the surface of the brain to record cortical activity, and is, therefore, more sensitive than EEG or fMRI. Using ECoG, Noy et al. (2015) observed activity in fronto-parietal regions associated with conscious experiences even when no report was required from subjects. Notably, even with a highly sensitive method such as ECoG, the signal of the fronto-parietal electrodes was less than one third of that recorded from visual areas, contributing to the hypothesis that the NCCs in PFC are comparatively subtle. Also using ECoG, Vidal et al. (2015) measured brain activity in subjects that experienced a form of perceptual suppression induced by a sudden contrast decrease of a stimulus (a dark gray fuzzy circle), causing it to vanish from consciousness. They first conducted a behavioral task to confirm that a sudden 50% contrast decrease caused a complete disappearance of the stimulus in 95% of the cases. Next, Vidal et al. compared the brain activity involved in either the disappearance of a stimulus from con-

---

tell us anything about the probability of the null hypothesis being true. Researchers are typically worried about Type I errors, but largely let Type II errors (accepting the null hypothesis when it is in fact false) go uncontrolled (Lieberman & Cunningham, 2009). Here, we suggest that a Type II error was committed. Unfortunately, this type of errors does not come as a great surprise in a field as underpowered as neuroscience (Cremers et al., 2017; Szucs & Ioannidis, 2017). Null results should be interpreted very carefully.

consciousness (perceptual suppression) or its re-appearance in consciousness during perceptual suppression release, with identical stimuli in both conditions, and without reports. Activity in PFC, and more specifically in the inferior frontal gyrus, as well as in the anterior insular cortex, was one of the main differences between invisible and visible conditions, even though no report was required in this experiment.

Another important consideration when interpreting fMRI univariate results is that neural activity in PFC is usually not linearly correlated with experimental manipulations. While neural activity in sensory cortex is typically linearly correlated to stimulus and behavioral profiles, frontal neurons have a mixed-selectivity profile with high dimensionality and contextual encoding (Fusi et al., 2016; Rigotti et al. 2013; Mante et al. 2013). This further suggests that fMRI univariate analyses are in general suboptimal for detecting specialized neural activity in PFC. It is of utmost importance, then, to rely on multivariate analyses through which complex patterns of activity in PFC can be detected.

To emphasize, without considering the logic behind the absence of evidence using insensitive measures like fMRI, without considering the mixed-selectivity profile of neurons in PFC, and without subjecting neural data to more sensitive statistical analyses such as multivariate pattern classification, a null result like Tse et al.’s (or finding decreased activity like the one found by Frässle et al.) should be taken with a grain of salt.

### *c. Dreams*

Binocular rivalry and masking paradigms are not the only way of testing the NCCs without requiring report. The study of dreams, a natural no-task and no-report paradigm, has also been thought to provide support for the idea that PFC does not contribute to consciousness (Koch et al., 2016). For example, Boly et al. (2017) claim that “within-state studies contrasting dreaming versus non-dreaming during NREM sleep and REM sleep, have pointed to a ‘posterior hot zone’ in parieto-occipital areas, possibly extending to mid-cingulate regions, as a reliable NCC” (p. 9606), and argue that “during REM sleep, a “no-task” state, specific dream contents, such as faces, places, movement, and speech, can be predicted from posterior, but not anterior, cortex (Siclari et al., 2017)” (p. 9607).

First, proponents of prefrontal theories of consciousness do not deny that the contents of consciousness can be predicted from patterns of neural activity in posterior areas. However, the fact that dream contents

can be decoded from these areas falls short from showing that these regions are responsible for consciousness of these contents. Second, as recently reported by Fazekas & Nemeth (2018), a wide variety of studies, including Siclari et al. (2017), have found significant increases in activity in PFC during dreams, particularly in medial PFC (Braun et al., 1997; Maquet et al., 1996, 2000), but also in dorsolateral PFC, using various measurement methods such as positron emission tomography (PET) (Hong et al., 1995; Nofzinger et al., 1997), fMRI (Hong et al., 2009), MEG (Ioannides et al., 2007), near-infrared spectroscopy (Kubota et al., 2011) and intracranial recordings (Vijayan et al., 2017). Hence, the view that PFC is hypoactive during dreams does not seem to be warranted by the data.

## Conclusion

We showed, first, that in multiple report-based experiments comparing unconscious and conscious conditions, the difference in brain activity in the PFC is unlikely to be explained solely by its role in generating reports. Second, we argued that no-report experiments either elicited PFC activity correlated with consciousness, as in cases of binocular rivalry and dreams, or produced null results that shouldn't be taken as indicating that the PFC is not involved in consciousness of perceptual contents. Indeed, these null results can be explained away by appealing to the lack of sensitivity of fMRI measures. We compared them to empirical results based on more sensitive measurement procedures and reached the conclusion that these null results were likely to be false negatives.

One clear lesson from the literature we discussed is that the PFC is probably involved both in generating reports of conscious experiences and in consciousness of contents. Although we rejected the merits of the argument from report based on available empirical evidence, we do not think that proponents of prefrontal theories of consciousness should necessarily shy away from no-report paradigms. As argued above, common neuroimaging techniques are not sufficiently sensitive to detect subtle differences in neural activity, but NCCs might indeed be subtle. Moreover, unlike sensory cortices, neural activity in prefrontal neural populations is highly multidimensional (Mante et al. 2013; Parthasarathy et al., 2017), affecting the chances of success of simple univariate analyses contrasting the overall level of neural activity between conscious and unconscious conditions. This makes multivariate pattern decoding approaches more appropriate for establishing a functional role of subtle neural activity (and, even then,

decoding information from PFC is not without challenges (Bhandari et al. 2018)). Taken together, these facts about the NCC, the PFC, and our current statistical and neuroimaging methods suggest that reducing the potential confounds created by reports can be beneficial. By using experimental paradigms that use task-irrelevant or non-reported perceptual experiences to identify the specific mechanisms involved in supporting consciousness, proponents of prefrontal theories can focus on finding the NCC rather than on disentangling reports from conscious awareness.

The potential benefit of using no-report paradigms should not obscure the fact that we already have solid evidence that PFC is involved in consciousness *despite* the use of paradigms that require subjective reports. Disentangling reports from consciousness might be beneficial for refining our understanding of the PFC's role. But it is by no means necessary, and perhaps it is not even a completely feasible task. Experimenters might have devised ways of inferring whether subjects are undergoing conscious experiences or not without a concomitant explicit report from them. However, despite their name, no-report paradigms rely on a pre-established association between subjects' conscious perception and their explicit reports. For all the ingenuity and potential advantages of no-report paradigms, the neuroscience of consciousness is ultimately about subjective reports.

**Acknowledgments:** We thank Hakwan Lau, Brian Odegaard, the editor Steven Fleming, as well as two anonymous reviewers for their helpful comments.

## References

- Aru, J., Bachmann, T., Singer, W., & Melloni, L. (2012). Distilling the neural correlates of consciousness. *Neuroscience and Biobehavioral Reviews*, 36(2), 737–746.
- Baars, B. J. (1988). *A Cognitive Theory of Consciousness*. New York: Cambridge University Press.
- Bhandari, A., Gagne, C., & Badre, D. (2018). Just above Chance: Is It Harder to Decode Information from Human Prefrontal Cortex Blood Oxygenation Level-dependent Signals? *Journal of Cognitive Neuroscience*, 30(10), 1473-1498.
- Block, N. (1995). On a confusion about a function of consciousness. *Behavioral and Brain Sciences*, 18(2), 227–247.
- Boly, M., Massimini, M., Tsuchiya, N., Postle, B. R., Koch, C., & Tononi, G. (2017). Are the Neural Correlates of Consciousness in the Front or in the Back



- of the Cerebral Cortex? Clinical and Neuroimaging Evidence. *The Journal of Neuroscience*, 37(40), 9603–9613.
- Brascamp, J., Blake, R., & Knapen, T. (2015). Negligible fronto-parietal BOLD activity accompanying unreportable switches in bistable perception. *Nature Neuroscience*, 18(11), 1672–1678.
- Brascamp, J., Sterzer, P., Blake, R., & Knapen, T. (2018). Multistable Perception and the Role of Frontoparietal Cortex in Perceptual Inference. *Annual Review of Psychology*, 69(1), 77-103.
- Braun, A.R., Balkin, T.J., Wesenten, N.J., Carson, R.E., Varga, M., Baldwin, P., ..., Herscovitch, P. (1997). Regional cerebral blood flow throughout the sleep-wake cycle. An H2(15)O PET study. *Brain*, 120(7), 1173–1197.
- Cavada, C. (2000). The Anatomical Connections of the Macaque Monkey Orbitofrontal Cortex. A Review. *Cerebral Cortex*, 10(3), 220–242.
- Carmichael, S. T., & Price, J. L. (1995). Sensory and premotor connections of the orbital and medial prefrontal cortex of macaque monkeys. *Journal of Comparative Neurology*, 363(4), 642-664.
- Chalmers, D. (2000). What is a neural correlate of consciousness? In Metzinger, T. (Ed.). *Neural Correlates of Consciousness: Empirical and Conceptual Issues*, Cambridge, MA: MIT Press.
- Cheesman, J., & Merikle, P. M. (1986). Distinguishing conscious from unconscious perceptual processes. *Canadian Journal of Psychology*, 40(4), 343–367.
- Cremers, H. R., Wager, T. D., & Yarkoni, T. (2017). The relation between statistical power and inference in fMRI. *PLoS ONE*, 12(11).
- Crick, F., & Koch, C. (1990). Towards a neurobiological theory of consciousness. *Seminars in the Neurosciences*, 2, 263–275.
- De Graaf, T. A., Hsieh, P. J., & Sack, A. T. (2012). The “correlates” in neural correlates of consciousness. *Neuroscience and Biobehavioral Reviews*, 36(1), 191–197.
- De Lafuente, V., & Romo, R. (2005). Neuronal correlates of subjective sensory experience. *Nature Neuroscience*, 8(12), 1698–1703.
- De Lafuente, V., & Romo, R. (2006). Neural correlate of subjective sensory experience gradually builds up across cortical areas. *Proceedings of the National Academy of Sciences*, 103(39), 14266–14271.
- Dehaene, S., & Changeux, J. P. (2011). Experimental and Theoretical Approaches to Conscious Processing. *Neuron*, 70(2), 200–227.
- Dehaene, S., Charles, L., King, J.-R., & Marti, S. (2014). Toward a computational theory of conscious processing. *Current Opinion in Neurobiology*, 25, 76–84.
- Dehaene, S., Naccache, L., Cohen, L., Bihan, D. L., Mangin, J. F., Poline, J. B., & Rivière, D. (2001). Cerebral mechanisms of word masking and unconscious repetition priming. *Nature Neuroscience*, 4(7), 752–758.
- Del Cul, A., Baillet, S., & Dehaene, S. (2007). Brain dynamics underlying the nonlinear threshold for access to consciousness. *PLoS Biology*, 5(10), 2408–2423.
- Einhäuser, W., Stout, J., Koch, C., & Carter, O. (2008). Pupil dilation reflects perceptual selection and predicts subsequent stability in perceptual rivalry. *Proceedings of the National Academy of Sciences*, 105(5), 1704–1709.
- Fazekas, P., & Nemeth, G. (2018). Dream experiences and the neural correlates of perceptual consciousness and cognitive access. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 373(1755).
- Fisch, L., Privman, E., Ramot, M., Harel, M., Nir, Y., Kipervasser, S., ... Malach, R. (2009). Neural “Ignition”: Enhanced Activation Linked to Perceptual Awareness.

- ness in Human Ventral Stream Visual Cortex. *Neuron*, 64(4), 562–574.
- Fleck, M. S., Daselaar, S. M., Dobbins, I. G., & Cabeza, R. (2006). Role of Prefrontal and Anterior Cingulate Regions in Decision-Making Processes Shared by Memory and Nonmemory Tasks. *Cerebral Cortex*, 16(11), 1623–1630.
- Fleming, S. M., Weil, R. S., Nagy, Z., Dolan, R. J., & Rees, G. (2010). Relating Introspective Accuracy to Individual Differences in Brain Structure. *Science*, 329(5998), 1541–1543.
- Frässle, S., Sommer, J., Jansen, A., Naber, M., & Einhauser, W. (2014). Binocular Rivalry: Frontal Activity Relates to Introspection and Action but Not to Perception. *Journal of Neuroscience*, 34(5), 1738–1747.
- Funahashi, S., Bruce, C.J., and Goldman-Rakic, P.S. (1989). Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *Journal of Neurophysiology*, 61, 331–349.
- Fusi, S., Miller, E. K., & Rigotti, M. (2016). Why neurons mix: high dimensionality for higher cognition. *Current Opinion in Neurobiology*, 37, 66–74.
- Giles, N., Lau, H., & Odegaard, B. (2016). What Type of Awareness Does Binocular Rivalry Assess? *Trends in Cognitive Sciences*, 20(10), 719–720.
- Hebart, M. N., Schriever, Y., Donner, T. H., & Haynes, J. D. (2014). The Relationship between Perceptual Decision Variables and Confidence in the Human Brain. *Cerebral Cortex*, 26(1), 118–130.
- Hong, C. C.-H., Harris, J. C., Pearlson, G. D., Kim, J.-S., Calhoun, V. D., Fallon, J. H. ... Pekar, J. J. (2009). fMRI evidence for multisensory recruitment associated with rapid eye movements during sleep. *Human Brain Mapping*, 30(5), 1705–1722.
- Hong, C. C., Gillin, J. C., Dow, B. M., Wu, J., & Buchsbaum, M. S. (1995). Localized and lateralized cerebral glucose metabolism associated with eye movements during REM sleep and wakefulness: a positron emission tomography (PET) study. *Sleep*, 18(7), 570–580.
- Ioannides, A. A., Kostopoulos, G. K., Liu, L., & Fenwick, P. B. C. (2009). MEG identifies dorsal medial brain activations during sleep. *NeuroImage*, 44(2), 455–468.
- Irvine, E. (2013). Measures of Consciousness. *Philosophy Compass*, 8(3), 285–297.
- King, J.-R., Pescetelli, N., & Dehaene, S. (2016). Brain Mechanisms Underlying the Brief Maintenance of Seen and Unseen Sensory Information. *Neuron*, 92(5), 1122–1134.
- Koch, C., Massimini, M., Boly, M., & Tononi, G. (2016). Neural correlates of consciousness: progress and problems. *Nature Reviews Neuroscience*, 17(5), 307–321.
- Kondo, H., Saleem, K. S., & Price, J. L. (2003). Differential connections of the temporal pole with the orbital and medial prefrontal networks in macaque monkeys. *Journal of Comparative Neurology*, 465(4), 499–523.
- Kondo, H., Saleem, K. S., & Price, J. L. (2005). Differential connections of the perirhinal and parahippocampal cortex with the orbital and medial prefrontal networks in macaque monkeys. *Journal of Comparative Neurology*, 493(4), 479–509.
- Kozuch, B. (2014). Prefrontal lesion evidence against higher-order theories of consciousness. *Philosophical Studies*, 167(3), 721–746.
- Kubota, Y., Takasu, N. N., Horita, S., Kondo, M., Shimizu, M., Okada, T. ... Toichi, M. (2011). Dorsolateral prefrontal cortical oxygenation during REM sleep in humans. *Brain Research*, 1389, 83–92.

- Lamme, V. A. F. (2006). Towards a true neural stance on consciousness. *Trends in Cognitive Sciences*, 10(11), 494–501.
- Lau, H. C., & Passingham, R. E. (2006). Relative blindsight in normal observers and the neural correlate of visual consciousness. *Proceedings of the National Academy of Sciences*, 103(49), 18763–18768.
- Lau, H., & Rosenthal, D. (2011). Empirical support for higher-order theories of conscious awareness. *Trends in Cognitive Sciences*, 15(8), 365–373.
- LeDoux, J. E., & Brown, R. (2017). A higher-order theory of emotional consciousness. *Proceedings of the National Academy of Sciences*, 114(10), E2016–E2025.
- Lew, S., & Lau, H. (2017). Crucial Role of the Prefrontal Cortex in Conscious Perception. In *Executive Functions in Health and Disease* (pp. 129–141). Academic Press.
- Lieberman, M. D., & Cunningham, W. A. (2009). Type I and Type II error concerns in fMRI research: Re-balancing the scale. *Social Cognitive and Affective Neuroscience*, 4(4), 423–428.
- Luppino, G., Matelli, M., Camarda, R., & Rizzolatti, G. (1993). Corticocortical Connections of Area-F3 (SMA-proper) and Area F6 (pre-SMA) in the Macaque Monkey. *The Journal of Comparative Neurology*, 338(1), 114–140.
- Lumer, E. D., & Rees, G. (1999). Covariation of activity in visual and prefrontal cortex associated with subjective visual perception. *Proceedings of the National Academy of Sciences*, 96(4), 1669–1673.
- Mante, V., Sussillo, D., Shenoy, K. V., & Newsome, W. T. (2013). Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature*, 503(7474), 78–84.
- Mantini, D., Corbetta, M., Romani, G. L., Orban, G. A., & Vanduffel, W. (2013). Evolutionarily Novel Functional Networks in the Human Brain? *The Journal of Neuroscience*, 33(8), 3259–3275.
- Maquet, P. (2000). Functional neuroimaging of normal human sleep by positron emission tomography. *Journal of Sleep Research*, 9(3), 207–231.
- Maquet, P., Peters, J. M., Aerts, J., Delfiore, G., Degueldre, C., Luxen, A., & Franck, G. (1996). Functional neuroanatomy of human rapid-eye-movement sleep and dreaming. *Nature*, 383(6596), 163–166.
- McCurdy, L. Y., Maniscalco, B., Metcalfe, J., Liu, K. Y., de Lange, F. P., & Lau, H. (2013). Anatomical coupling between distinct metacognitive systems for memory and visual perception. *The Journal of Neuroscience*, 33(5), 1897–1906.
- Merten, K., & Nieder, A. (2012). Active encoding of decisions about stimulus absence in primate prefrontal cortex neurons. *Proceedings of the National Academy of Sciences*, 109(16), 6289–6294.
- Michel, M. (2017). Methodological artefacts in consciousness science. *Journal of Consciousness Studies*, 24(11–12), 94–117.
- Miller, E.K., & Cohen, J.D. (2001) An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, 24, 167–202.
- Milner, A. D., & Goodale, M. A. (2008). Two visual systems re-viewed. *Neuropsychologia*, 46(3), 774–85.
- Morales, J. & Lau, H. (forthcoming). The neural correlates of consciousness. In Uriah Kriegel (ed.), *The Oxford Handbook of the Philosophy of Consciousness*. Oxford University Press.
- Morales, J., Chiang, J., & Lau, H. (2015). Controlling for Performance Capacity Confounds in Neuroimaging Studies of Conscious Awareness. *Neuroscience of Consciousness*, 1(1).

- Morales, J., Lau, H., & Fleming, S. M. (2018). Domain-General and Domain-Specific Patterns of Activity Supporting Metacognition in Human Prefrontal Cortex. *The Journal of Neuroscience*, *38*(14), 3534–3546.
- Naber, M., Frässle, S., & Einhäuser, W. (2011). Perceptual rivalry: Reflexes reveal the gradual nature of visual awareness. *PLoS ONE*, *6*(6).
- Neubert, F.-X., Mars, R. B., Thomas, A. G., Sallet, J., & Rushworth, M. F. S. (2014). Comparison of human ventral frontal cortex areas for cognitive control and language with areas in monkey frontal cortex. *Neuron*, *81*(3), 700–713.
- Nofzinger, E. A., Mintun, M. A., Wiseman, M., Kupfer, D. J., & Moore, R. Y. (1997). Forebrain activation in REM sleep: An FDG PET study. *Brain Research*, *770*(1–2), 192–201.
- Noy, N., Bickel, S., Zion-Golombic, E., Harel, M., Golan, T., Davidesco, I. ... Malach, R. (2015). Ignition’s glow: Ultra-fast spread of global cortical activity accompanying local “ignitions” in visual cortex during conscious visual perception. *Consciousness and Cognition*, *35*, 206–224.
- Öngür, D., Ferry, A. T., & Price, J. L. (2003). Architectonic subdivision of the human orbital and medial prefrontal cortex. *The Journal of Comparative Neurology*, *460*(3), 425–449.
- Odegaard, B., Knight, R. T., & Lau, H. (2017). Should a few null findings falsify prefrontal theories of conscious perception? *The Journal of Neuroscience*, *37*(40), 9593–9602.
- Panagiotaropoulos, T. I., Deco, G., Kapoor, V., & Logothetis, N. K. (2012). Neuronal Discharges and Gamma Oscillations Explicitly Reflect Visual Consciousness in the Lateral Prefrontal Cortex. *Neuron*, *74*(5), 924–935.
- Parthasarathy, A., Herikstad, R., Bong, J. H., Medina, F. S., Libedinsky, C., & Yen, S. C. (2017). Mixed selectivity morphs population codes in prefrontal cortex. *Nature Neuroscience*, *20*(12), 1770–1779.
- Passingham, R. (2009). How good is the macaque monkey model of the human brain? *Current Opinion in Neurobiology*, *19*(1), 6–11.
- Passingham, R. E., & Wise, S. P. (2012). *The Neurobiology of the Prefrontal Cortex. Anatomy, Evolution, and the Origin of Insight*. Oxford: Oxford University Press.
- Persaud, N., McLeod, P., & Cowey, A. (2007). Post-decision wagering objectively measures awareness. *Nature Neuroscience*, *10*(2), 257–261.
- Petrides, M., & Pandya, D. N. (1999). Dorsolateral prefrontal cortex: Comparative cytoarchitectonic analysis in the human and the macaque brain and corticocortical connection patterns. *European Journal of Neuroscience*, *11*(3), 1011–36.
- Petrides, M., & Pandya, D. N. (2002). Comparative cytoarchitectonic analysis of the human and the macaque ventrolateral prefrontal cortex and corticocortical connection patterns in the monkey. *European Journal of Neuroscience*, *16*(2), 291–310.
- Petrides, M., Tomaiuolo, F., Yeterian, E. H., & Pandya, D. N. (2012). The prefrontal cortex: Comparative architectonic organization in the human and the macaque monkey brains. *Cortex*, *48*(1), 46–57.
- Pitts, M. A., Metzler, S., & Hillyard, S. A. (2014). Isolating neural correlates of conscious perception from neural correlates of reporting one’s perception. *Frontiers in Psychology*, *5*(SEP), 1–16.
- Pollen, D. A. (2011). On the emergence of primary visual perception. *Cerebral Cortex*, *21*(9), 1941–1953.
- Ramsøy, T. Z., & Overgaard, M. (2004). Introspection and subliminal perception.

- Phenomenology and the Cognitive Sciences*, 3(1), 1–23.
- Rilling, J. K. (2006). Human and nonhuman primate brains: Are they allometrically scaled versions of the same design? *Evolutionary Anthropology*, 15(2), 65–77.
- Rigotti, M., Barak, O., Warden, M. R., Wang, X.-J., Daw, N. D., Miller, E. K., & Fusi, S. (2013). The importance of mixed selectivity in complex cognitive tasks. *Nature*, 497(7451), 585–590.
- Sadaghiani, S., Hesselmann, G., & Kleinschmidt, A. (2009). Distributed and antagonistic contributions of ongoing activity fluctuations to auditory stimulus detection. *The Journal of Neuroscience*, 29(42), 13410–13417.
- Saleem, K. S., Kondo, H., & Price, J. L. (2008). Complementary circuits connecting the orbital and medial prefrontal networks with the temporal, insular, and opercular cortex in the macaque monkey. *Journal of Comparative Neurology*, 506(4), 659–93.
- Salti, M., Monto, S., Charles, L., King, J.R., Parkkonen, L., & Dehaene, S. (2015). Distinct cortical codes and temporal dynamics for conscious and unconscious percepts. *eLife*, 4, e05652.
- Sergent, C., & Dehaene, S. (2004). Is consciousness a gradual phenomenon? Evidence for an all-or-none bifurcation during the attentional blink. *Psychological Science*, 15(11), 720–728.
- Sergent, C., Baillet, S., & Dehaene, S. (2005). Timing of the brain events underlying access to consciousness during the attentional blink. *Nature Neuroscience*, 8(10), 1391–1400.
- Siclari, F., Baird, B., Perogamvros, L., Bernardi, G., LaRocque, J. J., Riedner, B., ... Tononi, G. (2017). The neural correlates of dreaming. *Nature Neuroscience*, 20(6), 872–878.
- Soto, D., Mäntylä, T., Silvanto, J., 2011. Working memory without consciousness. *Current Biology*, 21:R912–R913.
- Sterzer, P., & Kleinschmidt, A. (2007). A neural basis for inference in perceptual ambiguity. *Proceedings of the National Academy of Sciences*, 104(1), 323–328.
- Storm, J. F., Boly, M., Casali, A. G., Massimini, M., Olcese, U., Pennartz, C. M. A., & Wilke, M. (2017). Consciousness Regained: Disentangling Mechanisms, Brain Systems, and Behavioral Responses. *The Journal of Neuroscience*, 37(45), 10882–10893.
- Szucs, D., & Ioannidis, J. P. A. (2017). Empirical assessment of published effect sizes and power in the recent cognitive neuroscience and psychology literature. *PLoS Biology*, 15(3).
- Tononi, G., Boly, M., Massimini, M., & Koch, C. (2016). Integrated information theory: from consciousness to its physical substrate. *Nature Reviews. Neuroscience*, 17(7), 450–461.
- Trübtschek, D., Marti, S., Ojeda, A., King, J. R., Mi, Y., Tsodyks, M., & Dehaene, S. (2017). A theory of working memory without consciousness or sustained activity. *eLife*, 6.
- Tse, P. U., Martinez-Conde, S., Schlegel, A. A., & Macknik, S. L. (2005). Visibility, visual awareness, and visual masking of simple unattended targets are confined to areas in the occipital cortex beyond human V1/V2. *Proceedings of the National Academy of Sciences*, 102(47), 17178–17183.
- Tsuchiya, N., Wilke, M., Frässle, S., & Lamme, V. A. F. (2015). No-Report Paradigms: Extracting the True Neural Correlates of Consciousness. *Trends in Cognitive Sciences*, 19(12), 757–770.

- Van Essen, D. C., & Dierker, D. L. (2007). Surface-Based and Probabilistic Atlases of Primate Cerebral Cortex. *Neuron*, *56*(2), 209–225.
- van Vugt, B., Dagnino, B., Vartak, D., Safaai, H., Panzeri, S., Dehaene, S., & Roelfsema, P. R. (2018). The threshold for conscious report: Signal loss and response bias in visual and frontal cortex. *Science*, *7186*(March), 537–542.
- Vidal, J. R., Perrone-Bertolotti, M., Kahane, P., & Lachaux, J. P. (2015). Intracranial spectral amplitude dynamics of perceptual suppression in fronto-insular, occipito-temporal, and primary visual cortex. *Frontiers in Psychology*, *6*(JAN), 1–14.
- Vijayan, S., Lepage, K. Q., Kopell, N. J., & Cash, S. S. (2017). Frontal beta-theta network during REM sleep. *eLife*, *6*.
- Webster, M. J., Bachevalier, J., & Ungerleider, L. G. (1994). Connections of inferior temporal areas TEO and TE with parietal and frontal cortex in macaque monkeys. *Cerebral Cortex*, *4*(5), 470-83.
- Weilnhammer, V. A., Ludwig, K., Hesselmann, G., & Sterzer, P. (2013). Frontoparietal Cortex Mediates Perceptual Transitions in Bistable Perception. *Journal of Neuroscience*, *33*(40), 16009–16015.
- Wilcke, J. C., O’Shea, R. P., & Watts, R. (2009). Frontoparietal activity and its structural connectivity in binocular rivalry. *Brain Research*, *1305*, 96–107.
- Williams, M. A., Visser, T. A. W., Cunnington, R., & Mattingley, J. B. (2008). Attenuation of neural responses in primary visual cortex during the attentional blink. *The Journal of Neuroscience*, *28*(39), 9890–9894.
- Zaretskaya, N., & Narinyan, M. (2014). Introspection, attention or awareness? The role of the frontal lobe in binocular rivalry. *Frontiers in Human Neuroscience*, *8*(527), 1–2.
- Zou, J., He, S., & Zhang, P. (2016). Binocular rivalry from invisible patterns. *Proceedings of the National Academy of Sciences*, *113*(30), 8408–8413.