Philos Stud
DOI 10.1007/s11098-013-0123-9

Prefrontal lesion evidence against higher-order theories of consciousness

Benjamin Kozuch

Abstract According to higher-order theories of consciousness, a mental state is conscious only when represented by another mental state. Higher-order theories must predict there to be some brain areas (or networks of areas) such that, because they produce (the right kind of) higher-order states, the disabling of them brings about deficits in consciousness. It is commonly thought that the prefrontal cortex produces these kinds of higher-order states. In this paper, I first argue that this is likely correct, meaning that, if some higher-order theory is true, prefrontal lesions should produce dramatic deficits in visual consciousness. I then survey prefrontal lesion data, looking for evidence of such deficits. I argue that no such deficits are to be found, and that this presents a compelling case against higher-order theories.

Keywords Prefrontal cortex · Lesions · Higher-order theories · Consciousness

1 Introduction

The past few decades have seen the rise of a number of theories of phenomenal consciousness. Few, however, have experienced the same sustained attention as so-called higher-order theories. According to higher-order theories of consciousness, some mental state M is conscious only if it is represented by some higher-order mental state M*. For proponents of higher-order theories, common sense backs this proposition, in that we would not call a mental state conscious if the subject is unaware of it. If this is true, then a necessary condition upon a mental state being conscious is that it be something of which the subject is aware. The relevant kind of

B. Kozuch (✉)
Philosophy Department, University of Arizona, Tucson, AZ, USA
e-mail: bigben@email.arizona.edu

Published online: 21 March 2013

awareness is said to be representational, and so it will be a higher-order *representation* of some mental state M that is necessary for M being conscious.[1,2]

Exactly what kind of higher-order representation is *sufficient* for a mental state being conscious is a matter of disagreement among higher-order theorists. Some think that the right kind of higher-order state is perception-like (Locke 1690; Armstrong 1968; Lycan 1996); for others it is more thought-like (Gennaro 1996; Carruthers 2000; Rosenthal 2002; Rolls 2004).[3] Some higher-order theorists hold that the higher-order representation is numerically distinct from the state being represented (Lycan 1996; Rosenthal 2002), whereas others hold that the higher- and lower-order states are parts of the same mental state (Gennaro 1996; Carruthers 2000; Van Gulick 2004; Kriegel 2009). Some hold that the actual *occurrence* of a higher-order state is necessary for a conscious state (most higher-order theorists think this), whereas others believe there need be only a *disposition* for such a state to occur (Carruthers 2000).

While there has been much enthusiasm for the higher-order (HO) approach, HO theories have one virtue only recently recognized, namely, their amenability to empirical confirmation or disconfirmation (see, e.g., Lau and Rosenthal 2011). In this paper, lesion evidence is used for this purpose. The idea is the following: If a mental state is conscious only if it is the target of a certain kind of HO representation, then there are some area(s) in the brain such that, because they produce the right kind of HO states, they are necessary for conscious states. Let us refer to these brain areas as *integral* areas. (Alternatively, these could be *networks* of areas; see below.) Now, if we can identify those brain areas likely to produce HO representations—that is, if we can identify *potential* integral areas—then the results of lesions to these areas can be used as evidence for or against HO theories: If deficits in phenomenal consciousness follow lesions to the potential integral areas, this is evidence for HO theories. But if in no case do such deficits follow, then this is important evidence against HO theories. Whether the available lesion evidence confirms or disconfirms HO theories is the issue with which this paper is concerned.

If lesion evidence is to be used in this way, we need to know where integral areas are most likely to be. Most have thought they would be in the prefrontal cortex (Beeckmans 2007; Block 2009; Kriegel 2009; Lau 2010; Brown 2011; Lau and Rosenthal 2011; Lau and Brown forthcoming; but see Gennaro 2012). If this is correct, however, it could prove unfortunate for HO theories. If there were integral areas in the prefrontal cortex, then prefrontal lesions should—in at least some cases—bring about dramatic deficits in visual consciousness. In this paper, I argue

[1] For lucid presentations of the argument just recounted (the so-called simple argument for higher-order theory), see Lycan (2001), Rosenthal (2002). Many, perhaps all, higher-order theorists find this argument compelling (e.g., Gennaro 1996; Carruthers 2000; Kriegel 2009); but see Block (2009), which calls into question the cogency of it.

[2] Higher-order theories can be contrasted with first-order theories of consciousness (Harman 1990; Dretske 1995; Tye 1995), which hold that a first-order representational state can be conscious (whether or not it is targeted by a higher-order state), as long as the first-order state satisfies other conditions (such as being available for use in reasoning and belief-formation).

[3] This is somewhat of a simplification, as Carruthers has sometimes described his theory as a blend of higher-order perception and higher-order thought theories (see his 2004).

that the expected deficits are not to be found, and that this constitutes important evidence against HO theories of consciousness.[4]

In Sect. 2, I describe and expand upon the reasons for which the prefrontal cortex is thought to contain integral areas. I conclude that (at least most) HO theories must predict that prefrontal lesions will sometimes produce striking deficits in visual consciousness. In Sect. 3, I first survey evidence in which prefrontal lesions fail to produce the expected deficits. I then examine data sometimes thought to support there being integral areas in the prefrontal cortex, arguing that other interpretations of these data are preferable. In Sect. 4, I consider some objections to the idea that the prefrontal lesion data should act as evidence against HO theories. The conclusion of the paper will be that the effects of prefrontal lesions (or lack thereof) count against HO theories, especially those theories hypothesizing the right kind of HO state to be of a more sophisticated sort.

Before starting, some clarifications. First: HO theories are often intended as theories of *phenomenal* consciousness, understood as the "what-it's-like" aspect of mental life (Nagel 1974; Block 1995; Chalmers 1995).[5] More specifically, HO theories claim that a mental state M cannot be *phenomenally* conscious unless M is targeted by some mental state M*. It is precisely this claim that is my target. Whether an HO theory might work as an explanation of some other kind of consciousness (e.g., "higher-order consciousness," see Block 2009) is not a question with which I am concerned. Second: For sake of variety and economy, I below use the terms "phenomenal consciousness," "consciousness," and "experience" equivalently, all meant to refer to phenomenal consciousness. (Note that this does not comport with how some HO theorists use these terms.)[6] Third: In this paper, I am concerned only with whether HO theories work as an explanation of *visual* consciousness. And so I only consider data and arguments insofar as they are relevant to visual consciousness.

2 If some higher-order theory is true, then lesions to the prefrontal cortex should produce deficits in visual consciousness

In this section, I argue for the thesis entitling this section, doing so in two parts. In Sect. 2.1, I recount and expand upon reasons previously given for thinking that, if some HO theory is true, then the prefrontal cortex should contain an integral area or areas. In Sect. 2.2, I argue that a lesion to an integral area should produce striking deficits in visual experience. The conclusion of this section will be that (at least most) HO theories must predict prefrontal lesions to produce dramatic deficits in visual experience.

[4] Or at least many HO theories. See the discussions below (in Sect. 2.1, and especially Sect. 4.3) concerning how not all HO theories seem equally committed to there being integral areas in the PFC.

[5] There is, for example, something it is like to have pangs of hunger, or to see the red of a firetruck. *Experiences* like these are what phenomenal consciousness refers to.

[6] For example, some HO theorists will use the terms "experience" or even "what-it's-like" to describe states that may or may not be phenomenally conscious (e.g., Carruthers 2000). (Byrne does a good job of sorting out the terminological thicket surrounding HO theories; see his 2004.)

2.1 If some higher-order theory is true, then the prefrontal cortex should contain one or more integral areas

As discussed above, HO theories entail there to be some areas in the brain such that, because they produce (the right kind of) HO states, they are necessary for having conscious states; I have been referring to these as *integral* areas. This means that, if we know where in the brain (the right kind of) HO states might be produced (that is, where the *potential* integral areas are), then we can use results of lesions to these areas as evidence for or against HO theories.

So, where might these HO states be produced? Most who have speculated upon this take the prefrontal cortex to be the best candidate.[7] Kriegel, for example, has argued that (the right kind of) HO states should be found "in an area of the brain that is associated with quite sophisticated cognition, since the…representations are…not of simple sensory stimuli…but of psychological states" (2007, p. 902; see also 2009). The part of the brain carrying out sophisticated metacognitive activities such as these, continues Kriegel, is the prefrontal cortex. Lau has likewise thought that any areas producing HO states are in the prefrontal cortex, reasoning that "one would expect the higher-order representations to be in these brain regions, as they seem to play roles in uncertainty monitoring, cognitive control, thinking and planning, etc." (2010, p. 4).

I believe that the prevailing wisdom is correct, and that the prefrontal cortex is where HO representations are most likely produced (that is, the *right kind* of HO representations; I now stop saying this each time). In this subsection, I lay out the prima facie case for this. More specifically, I argue for this conditional: If some HO theory is true, then there should be one or more integral areas in the prefrontal cortex.

First, though, we precisify the idea of an integral area. Given the integrated nature of brain function (Bressler and Menon 2010), it is possible—perhaps likely—that it is not individual brain areas necessary for producing HO states, but rather *networks* of areas. In light of this, we can consider a brain area an integral area if it is necessary for the having of conscious states either (a) because it produces HO states, or (b) because it is an essential component in a network of areas that produces HO states.[8] (By "essential," I mean that its malfunctioning likely entails the network malfunctioning.) The motivation for defining an integral area in this way becomes clear later on. Important for time being is that this makes the goal of this subsection more modest than it otherwise seems. All I aim to provide in this section is (at least) tentative reason to believe the following: Any networks able to produce HO states will probably include, as an essential component, one or more areas in the prefrontal cortex.

In general, it is because of two considerations that researchers have taken the prefrontal cortex to contain potential integral areas. The first is that those areas most

[7] For references, see p. 3.

[8] This is *still* a bit imprecise: If there were more than one network individually sufficient for the production of HO states, then a brain area might fail to be necessary for the having of conscious states, even though it plays an essential role in the production of any number of token conscious states. But the above formulation is sufficient for present purposes.

likely to produce HO states are those involved in monitoring and supervision of activity in other parts of the brain; the second is that these functions are carried out in the prefrontal cortex. Now we examine these considerations more closely.

The prefrontal cortex (PFC) consists of the foremost parts of the brain, those anterior to the primary motor cortex. The morphology of the human brain is not very different from other primates until one considers the PFC (Brodmann 1909; Deacon 1990). Here, ancestors of *Homo sapiens* underwent a massive expansion between 2.5 million and 100,000 years ago (Wills 1993). This expansion of the PFC is commonly thought to be what made possible many distinctively human capacities, such as reasoning, long-term planning, and advanced metacognition (Fuster 2002).

The PFC facilitates these and other sophisticated cognitive activities by supervising and coordinating activity in other brain areas (Alvarez and Emory 2006), this being enabled by its prolific feedforward and feedback connections with each other cortical (and often subcortical) part of the brain (Nauta 1972). Accordingly, the PFC carries out tasks falling under the rubric of *executive function* (Miyake et al. 2000), including working memory, action inhibition, the building of associations between stimulus and reward, error detection, conflict resolution, and (perhaps) theory of mind (Rolls 2000; Botvinik et al. 2001; Pardo et al. 1990; Frith and Frith 2003).[9]

Plausibly, many facets of executive function involve HO representation of visual states. For example, working memory consists of a limited-capacity buffer, one allowing information to be "held in mind" so that high-level tasks—ones linking long-term memory, perception, and action—can be performed (Baddeley 2003). Part of what (visual) working memory must do, presumably, is select certain first-order visual states for retention in the buffer. This plausibly involves HO representation of first-order visual states. If so, those prefrontal areas involved in working memory (namely, the lateral PFC; Owen 1997) can be considered potential integral areas.[10] Likewise, other aspects of executive control appear to involve HO representation of first-order visual states, including (but not necessarily limited to)

[9] There is some question as to whether the PFC plays an important role in theory of mind. (For arguments and data in favor of this idea, see Frith and Frith (2003), Gallagher and Frith (2003); for arguments and data against it, see Bird et al. (2004), Saxe et al. (2006).) This could be an issue of consequence: At least one HO theorist (Carruthers 2000) has thought that the theory of mind mechanism might produce (the right kind of) HO states (but see Kriegel 2007). If this were right, *and* it turned out that the theory of mind mechanism was located entirely outside the PFC, then prefrontal lesion evidence would be less relevant to evaluating HO theories than it is taken to be in this paper.

This, however, looks like an outlying possibility, as it rests on at least two contentious theses. The first is that the theory of mind mechanism does not depend essentially on structures in the PFC. As just seen, whether or not this is the case is a matter of ongoing debate. The second is that the theory of mind mechanism not only represents others' mental states, but also one's own mental states (more specifically, sensory states) (Carruthers 2010). The former, but not the latter, is a function traditionally ascribed to the theory of mind mechanism. Because of the contentiousness of these theses, I leave exploration of this possibility for future research.

[10] It should be mentioned that Baddeley's theory of working memory is in some danger of being supplanted by a newer theory (Zimmer 2008), one giving less of a central role to the PFC than Baddeley's. This is not of consequence to present purposes, since Zimmer's theory still gives the PFC an essential role; meaning that, if the network of brain areas involved in short-term memory constitute an integral network, we should expect deficits in visual consciousness to result when the PFC is damaged.

tasks such as building stimulus-reward relationships, conflict resolution, and direction of attention.

However, not all aspects of executive function involve HO representation of *visual* states. Inhibitory functions in the ventromedial PFC, for example, probably involve the monitoring of incipient actions, rather than visual states. So the ventromedial PFC probably should not be considered a potential integral area (at least not in virtue of this function). For similar reasons, Kriegel has claimed that those brain areas involved in theory of mind might not qualify as potential integral areas (2007).

However, it is best to put aside the issue of precisely *which* areas in the PFC qualify as potential integral areas. While we know much about the PFC, we still lack fine-grained knowledge of its functional architecture. Yet such knowledge is needed if we want to decisively rule out particular parts of the PFC from being integral areas. Accordingly, we settle for a more general conclusion; namely, that if some HO theory is true, it is likely that there is one or more integral areas *somewhere* in the PFC. Supporting this is the idea that the PFC is involved in the kinds of monitoring and supervisory functions with which the production of HO states would be associated.

While we are limiting ourselves to this more general conclusion, it is worth noting that momentum has built for the idea that the *dorsolateral* PFC (dlPFC) is a particularly good candidate for a potential integral area. In his (2007), Beeckmans identifies the conceptual short-term memory system (Potter 1999) as that which is most likely to produce HO states,[11,12] a system likely involving the dlPFC (Owen 1997). Kriegel (2007) and Lau (2008) have also thought the dlPFC a good candidate, on basis of neuroimaging studies in which dlPFC activity correlates with conscious perception of a stimulus (e.g., Sahraie et al. 1997; Lau and Passingham 2006; Maniscalco et al. 2009). Because this hypothesis is popular, we pay particular attention to the dlPFC when looking at the prefrontal lesion evidence below.

As noted, my goal in this subsection is only to support the idea that if some HO theory is true, then there should be one or more integral areas in the PFC. Admittedly, I have not presented a decisive case. Doing so involves showing it very unlikely that (the right kind of) HO states could be formed without help from the PFC; more precisely, it would involve showing that, according to our best neuroscience, it is very unlikely that there are any brain areas (or networks of areas), that are both (a) able to produce the right kind of HO states, and (b) are not (or do not include) areas in the PFC. Accomplishing this requires more space than available here. Nonetheless, it seems that *at least many* HO theories are committed to there being one or more integral areas in the PFC. I give detailed explanation as

[11] The primary function of the conceptual short term memory store is to create 'gists' of visual scenes: fleeting, conceptual representations of the objects composing one's current milieu.

[12] Beeckmans makes this identification even in the case of higher-order perception theory (e.g., Lycan 1996). One might wonder if it is apt to identify the production of higher-order perceptions with a *conceptual* short-term memory system, since some commentators (e.g., Carruthers 2007) have described higher-order perceptions as being *nonconceptual*. Beeckmans' reasons for making this identification are complex, and appeal to the specifics of Lycan's theory. Interested readers should refer to Beeckmans (2007, p. 103).

to why later (in Sect. 4.3). Broadly speaking, the reason comes from many popular HO theories (e.g., HO thought theory) hypothesizing the right kind of HO state to be of a sophisticated sort. When it comes to such HO theories, it is especially likely that the production of an HO state will involve at least *some* area in the PFC, given the association of the PFC with sophisticated kinds of cognition.

2.2 The lesioning of an integral area would cause significant deficits in visual experience

If it is right that (many) HO theories must predict there to be one or more integral areas in the PFC, we may examine results of lesions to the PFC to confirm or disconfirm HO theories. To do this, however, we first need an idea as to what effects losing an integral area would have.

It is somewhat straightforward what would happen if *each* of a subject's integral areas (or networks) were disabled. Since the subject could no longer produce those HO states necessary for visual consciousness, we may reasonably predict this results in something phenomenologically similar to blindness. But what should we expect in cases where some, but not all, integral areas are disabled?[13]

While perhaps not as dramatic, one suspects that such a subject would still experience striking deficits in visual experience: Before being lesioned, the integral area would have made possible many visually conscious states, ones that presumably played a part in composing the subject's ongoing visual experience. Relative to what the subject's visual experience had been like prior to the lesion, her ongoing visual experience would now have lacunae. Within her visual experience, there would no longer be visually conscious states where, prior to the lesion, she would have expected to find them. This would constitute some remarkable change in her ongoing visual experience, one likely manifest to the subject.[14,15]

One might object that the loss of an integral area could be masked by there being more than one integral area, or by another brain area taking over the lesioned area's duties (Kriegel 2007; Lau and Rosenthal 2011). I assess these objections in due course (in Sect. 4). For time being, I proceed on the assumption that the loss of an integral area would produce striking deficits in visual experience.

[13] Or in a parallel case, wherein a lesion to an integral area causes it to only *partially* lose its ability to produce HO states.

[14] This does raise a question: While it seems true that, if an integral area is damaged, this would result in a dearth of visually conscious states, and it also seems true that this loss would constitute a remarkable change in one's ongoing visual experience, one might ask: What exactly would this be like, to experience a deficit of visually conscious states? In *what* remarkable way would this lack of visually conscious states be a departure from what was previously typical of the subject's ongoing visual experience? For our purposes, we need not answer this question: It is safe to assume that, whatever this would be like, it would constitute *some* remarkable change in one's visual experience.

[15] I say that such a change would *likely* be manifest to the subject, since significant changes in visual consciousness can, in some instances, go unnoticed for a period of time. Subjects suffering from hemianopia (blindness for half the visual field) sometimes will not recognize their visual deficit until they have undergone testing. However, this is not so important for our purposes: When it comes to subjects in the prefrontal lesion data we look at below, it will look very likely that, if the subjects had any striking deficits in visual consciousness, such deficits would be discovered in the course of examination.

3 Lesions to the prefrontal cortex do not produce the expected deficits in visual consciousness

In the last section, we saw that the PFC is a good candidate for containing potential integral areas, and that damage to an integral area would bring about striking deficits in visual consciousness. If all this is right, then if some HO theory is true, prefrontal lesions should—at least in some cases—bring about dramatic deficits in visual consciousness. In this section, I argue that prefrontal lesions fail to produce the relevant kinds of deficit.

It might appear that I am swimming upstream. There has, after all, been a recent movement among some neuroscientists towards thinking that the PFC is important, perhaps necessary,[16] for visual consciousness (Sahraie et al. 1997; Sergent and Dehaene 2004; Lau and Passingham 2006; Lau 2008, 2010; Maniscalco et al. 2009; Dehaene and Changeux 2011). If the PFC were necessary for visual consciousness, this is of course favorable to HO theories, as it constitutes evidence for there being integral areas in the PFC. In large part, this trend is based on recent neuroimaging studies showing correlations between PFC activity and consciousness (e.g., Dehaene et al. 2001; Lau and Passingham 2006; Del Cul et al. 2009). More recently, these and other data were marshaled into an empirically based argument for HO theories (Lau and Rosenthal 2011).

However, not everyone is convinced that the PFC is important for consciousness (Pollen 1995, 2008; Prinz 2000; Zeki 2003; Block 2005; Macknik 2006; Gennaro 2012). In discussing the studies just mentioned, Pollen recently argued that "although prefrontal activity can secondarily modify activity in the posterior brain regions, longstanding neurological evidence suggests that prefrontal activity is not essential for [visual consciousness]" (2008, p. 1992). To support this, Pollen cites a collection of prefrontal lesion studies, claiming that they show prefrontal lesions do not bring about disturbances in visual consciousness. If right, this of course would be important evidence against HO theories.

Whether or not prefrontal lesions bring about deficits in visual consciousness, and what this means for HO theories, is the topic of this section. In Sect. 3.1, I survey a number of PFC lesion studies in which visual consciousness remains unaffected, arguing that these studies constitute evidence against the PFC containing integral areas. In Sect. 3.2, I consider neuroscientific evidence cited in favor of the PFC being necessary for visual consciousness, arguing that these data fail to support HO theories.

3.1 Evidence showing that lesions to the prefrontal cortex do not produce the expected deficits in visual consciousness

Often the PFC is divided into three major divisions: the orbital, lateral, and medial PFC. Though not in the prefrontal lobe, the anterior cingulate cortex (ACC) is also

[16] It is not always clear whether these researchers think the PFC to be *necessary* for visual consciousness. When expressing what they think is the relationship between visual consciousness and the PFC, they often put it in terms of the PFC being "critical," or "essential" for visual consciousness.

sometimes considered part of the PFC (Andrewes 2001, Chap. 3). I will follow this convention, since the ACC plays a role in executive function and is therefore a candidate for producing HO states. Moreover, two HO theorists (Kriegel 2007; Gennaro 2012) have recently argued that the ACC is a potential integral area. Sometimes the ACC is considered a separate part of the PFC, other times it is considered part of the medial PFC. I will consider it separate from the medial PFC, since unique kinds of deficit—ones worth looking at individually—can result from ACC lesions.

Lesions to the orbital, lateral, or medial PFC (what I will refer to as the *PFC proper*) produce so-called executive dysfunction (see, e.g., Alvarez and Emory 2006). Depending on the precise lesion location, subjects with damage to one of these areas have problems inhibiting inappropriate actions, switching efficiently from task to task, or retaining items in short-term memory. However, lesions to these areas appear not to produce notable deficits in visual consciousness: Tests of the perceptual abilities of subjects with lesions to the PFC proper reveal no such deficits; as well, PFC patients never report their visual experience to have changed in some remarkable way.[17]

Consider a study carried out by Heath et al. (1949). Prior to this experiment, one researcher had reported a monkey to have visual field defects following removal of prefrontal areas (Kennard 1939; see also Ferrier and Turner 1898), but later studies contradicted these results (Mettler 1944; Lashley and Clark 1946). To resolve these conflicting data, Heath and colleagues looked at 23 psychiatric patients who had various parts of their PFC bilaterally ablated, examining them for visual deficits pre- and post-operatively. The lesions included various areas within the medial, orbital, and lateral PFC. In three subjects, specifically the dorsolateral PFC (dlPFC) was ablated. (Recall that some HO theorists consider the dlPFC an especially good candidate for an integral area.)

The researchers used a battery of tests to examine the patients for visual deficits. Visual acuity was tested using Snellen charts (those posters imprinted with gradually smaller letters, familiar to us from doctors' offices and driving exams). A "tangent screen" was used to check for visual field defects: In a candle-lit room, small, white objects were presented against a black screen in various parts of the subjects' visual field in order to detect blind spots. Other methods were employed to search for visual deficits.

None of the 23 subjects—including those with bilateral dlPFC lesions—showed visual field defects, a loss of visual acuity, or abnormalities of color perception. Nor did they report that their visual experience had changed since the lesions. Heath et al. (1949, p. 147) conclude: "Bilateral, simultaneous ablation of various Brodmann frontal areas, and combinations of these, does not produce, in the human being, any impairment in the visual field nor in acuity of peripheral vision…[and] color vision is not adversely affected".

A survey of other studies involving lesions to the PFC proper reveals that, while such lesions have dramatic effects, they do not produce deficits in visual

[17] I am indebted to Pollen (2008) for calling my attention to some of the PFC lesion studies discussed below.

consciousness. For instance, subjects with damage to the orbital PFC, while often testing of normal intelligence or higher, have problems with impulsive and short-sighted behavior (Damasio 1994). However, these subjects show no abnormalities of vision when tested. Nor do they offer reports indicating their visual experience has changed in some remarkable way. The same is true of subjects with medial PFC damage, though they might have problems with memory, planning, or action initiation (Bird et al. 2004).

So far we have seen that isolated lesions to the orbital, medial, or lateral PFC do not produce deficits in visual experience. This is also true when all three areas are unilaterally ablated. Penfield and Evans (1935) describe two patients, each of whom had their right or left PFC removed to alleviate intractable epilepsy. While these patients post-operatively showed problems with action-initiation, and with planning and carrying out multi-stage tasks (such as preparing a meal), perceptual tests uncovered no visual deficits. And though they complained of disorganized thinking, they mentioned no way in which their visual experience had changed.[18]

There is even one recorded case of *bilateral* excision of prefrontal areas (Brickner 1936). (The removed portion included all of the PFC anterior to language and motor areas.) After the operation, the patient was childlike in intelligence and behavior, and habitually engaged in self-aggrandizement (people who knew him said that this latter behavior was not new, but rather greatly exaggerated since the operation). However, perceptual tests presented no evidence of visual deficits (ibid., p. 13), and the subject himself never mentioned any remarkable change in his visual experience: The book written about his case includes numerous quotes from the patient, but nowhere to be found are instances of him discussing some difference in how he visually experienced the world.

It looks, then, as if lateral, medial, and orbital PFC lesions do not produce deficits in visual experience. Now we turn to one last division of the PFC, the ACC. Whether or not ACC lesions result in deficits in visual consciousness is not straightforward. Subjects with ACC damage might suffer from lethargy and a diminished interest in things formerly enjoyed (Devinsky et al. 1995), or be insensitive to painful stimuli. None of this, of course, gives reason to think that they are experiencing deficits in visual consciousness. However, in instances where damage is particularly severe,[19] ACC lesions can bring about *akinetic mutism* (Cairns et al. 1941). Subjects with akinetic mutism sit or stand passively, appearing emotionally flat. Though such a subject appears awake, and tracks objects with his eyes, he will not initiate actions, and remains unresponsive. Some have wished to infer from this lack of responsiveness and initiative that akinetic mutes are not phenomenally conscious (Damasio 1994). This is unjustified. That those with akinetic mutism present with a lack of responsiveness and initiative does not shed light on whether they are phenomenally conscious or not: They might have these symptoms and have a visual phenomenology, or they might have these symptoms

[18] It is worth noting that one of the patients was Penfield's sister, and was therefore someone with whom Penfield had frequent and intimate interactions. This makes it seem especially likely that any deficits would have been discovered.

[19] And especially when the supplementary motor area is also damaged (see below).

and not have a visual phenomenology. Usually we settle issues of what a subject's experience is like by getting a report from the subject, but those suffering from akinetic mutism (by definition) make no reports.

However, there is indirect evidence indicating that subjects with akinetic mutism lack deficits in consciousness. Some who recover from akinetic mutism lack memories of their time afflicted (Damasio 1999, p. 262), but those remembering describe themselves as having been aware of their surroundings, but entirely lacking in motivation. One patient said she could follow conversations, but would not participate because she had "nothing to say," and that her mind had been "empty" (Damasio and Van Hoesen 1983, p. 98). Such reports suggest that akinetic mutes have a paralyzing lack of motivation, not deficits in visual experience. This appears all the more plausible given the role the ACC is thought to play in motivation (Devinsky et al. 1995), and because akinetic mutism is most common in cases where the supplementary motor area (an area neighboring the ACC) is also damaged. The supplementary motor area, like the ACC, is thought to be involved in the initiation of motor behavior (ibid.).[20] Subjects with akinetic mutism, it seems, are best not interpreted as having deficits in visual consciousness.

To conclude, damage to any of the four major divisions of the PFC does not appear to bring about deficits in visual consciousness; at least not the kind we would expect if the PFC contained integral areas. This suggests rather forcefully that there are no integral areas in the PFC. However, before drawing this conclusion, I will consider (in the next subsection), data some researchers have claimed to support HO theories.

Before moving on, there is one issue to address, having to do with whether subjects in the above studies were ever *explicitly* asked if they had deficits in visual consciousness. If not, one might wonder if the subjects had deficits, and the researchers merely failed to discover them. But consider that, in the above studies, subjects were being examined specifically to see what kinds of deficit had been created by a lesion. If this is the point of an examination, it seems likely that any deficits in visual consciousness would be detected, even if the experimenters never specifically asked about them.[21] This seems especially true in regards to the kind of dramatic deficits the loss of an integral area would likely produce. And so the question as to whether subjects were ever explicitly asked about deficits in visual consciousness appears not so important.

3.2 Evidence thought to support the PFC containing integral areas

As mentioned above, some neuroscientists hypothesize PFC activity to be essential for visual consciousness. This has been based, in large part, upon experiments revealing correlations between PFC activity and consciousness. More recently, Lau

[20] Akinetic mutism resulting from damage just to the ACC is rare and usually short-lasting. It can also be brought about by damage to subcortical structures connected to the ACC (Mega and Cohenour 1997).

[21] In what particular manner should we expect the deficits to be detected? I do not precisely know, but one could guess that a subject with a disabled integral area would not perform normally on (at least some) tests of their visual abilities. Failing that, we could probably still expect the subject to volunteer information indicating that things "seemed" visually different to her.

and Rosenthal have claimed it is a "myth" that prefrontal lesions do not affect visual consciousness (2011; see also Dehaene and Changeux 2011), citing newer studies in which subjects with prefrontal lesions seem to experience deficits in visual consciousness. I argue in this subsection, however, that we can explain these data without hypothesizing the PFC to contain integral areas. Moreover, since prefrontal lesions fail to produce the kinds of deficit we would expect were it to contain integral areas—even in those studies cited in favor of HO theories—we should prefer those explanations that do not.

3.2.1 Lesion evidence thought to support the PFC containing integral areas

We start with a representative example of the lesion studies offered in support of HO theories, one involving visual masking. In visual masking, a figure whose interior border overlaps the outer border of a target stimulus is shown to the subject shortly (tens of milliseconds) after the target stimulus. Done at the right interval, the target stimulus becomes invisible to the subject (it is "masked"). Del Cul and colleagues found that prefrontal patients[22] would fail to see stimuli normal subjects reported as visible. More specifically, the interval between the target and masking stimulus needed to be an average of about 20 ms longer before the prefrontal patients reported having seen the target.

In this study, there is an apparent difference between what is experienced by subjects with and without prefrontal lesions. In certain conditions, normal subjects see both the target and mask, whereas those with prefrontal lesions see only the mask. This might look like evidence for integral areas being in the PFC. However, these are not the kinds of deficit one would associate with the damaging of an integral area. Remember that, in Sect. 2.2, we assumed that a lesioned integral area would produce remarkable changes in one's ongoing visual experience (an assumption defended further below). But all that we see in this experiment are minor deficits: Compared to normal subjects, prefrontal patients find an already hard-to-see stimuli, a little harder to see. What we lack evidence for are the dramatic changes in visual experience that the loss of an integral area would cause.[23]

The same can be said of most other lesion studies cited in favor of HO theory: Barceló et al. (2000) found subjects with unilateral PFC lesions to be slightly worse than normals when attempting a "difficult bi-field visual-discrimination task" (p. 399), that of detecting a target stimulus heavily crowded by distractors. In another experiment (Turatto et al. 2004), subjects given repetitive bursts of transcranial magnetic stimulation (a technique used to induce temporary "lesions")[24] to the dlPFC were less likely than those given sham stimulation to correctly identify whether two visual scenes were different (i.e., they exhibited

[22] The location of the lesions differed from subject to subject.

[23] As discussed in fn. 14, something I am not speculating on is exactly *what* kinds of changes we would expect. Perhaps the subjects would find even non-masked stimuli hard to see. Perhaps they would balk at even attempting the task, because it would phenomenologically seem to them as if they were blind, or mostly blind. I am not sure of any of this. What I am certain of is that we would expect deficits of a magnitude greater than those seen in this experiment.

[24] A magnetic field is used to temporarily take "off-line" one or more brain areas.

"change blindness"). In each study, what we find are subjects with prefrontal lesions performing at a slightly lower level than controls at a task that even the controls find difficult. What we do not find—to repeat—are the remarkable disruptions in visual experience we would expect, in the case that an integral area had been damaged. (Parallel observations could be made about [Rounis et al. 2010], another lesion study cited by HO theorists).

Another line of lesion evidence cited in favor of HO theories comes from *hemispatial neglect* (Driver and Vuilleumier 2001). Subjects with this disorder will often fail to notice objects appearing in half their visual field (usually the left). Such a subject might only eat food on half of his plate, or shave just half his face. Neglect usually follows lesions to the inferior parietal cortex, but also more rarely results from lesions to prefrontal areas (Nakamura and Mishkin 1986; Maeshima et al. 1995; Husain and Kennard 1996). If one considers neglect a disorder of consciousness (Driver and Vuilleumier 2001; Koch and Tsuchiya 2007; Vosgerau and Newen 2008), then this looks like evidence for there being integral areas in the PFC.

However, not all believe neglect is a deficit of visual consciousness, instead thinking it is merely a disorder of *attention* (see Lamme 2006; Block 2007; and especially Jacob and De Vignemont 2010). This is plausible given the lesioned area's participation in an attentional network (Mesulam 1999; Bartolomeo et al. 2007). In addition, sometimes prefrontal neglect (as opposed to neglect caused by parietal lesions) is thought to consist just of a deficit in the planning and execution of motor actions towards objects in the neglected hemifield (Liu et al. 1992; Mesulam 1999; Verdon et al. 2010). Consider also that, if neglect involved the lesioning of an integral area, it is hard to explain why it occurs in but one hemifield at a time (i.e., there is no "bispatial neglect"): If neglect involved damage to an integral area, we would expect that, in cases of bilateral damage to the area in question, a loss of visual consciousness for the entire visual field would result; but such results are never found.[25]

Nonetheless, of the lesion evidence offered on behalf of HO theories, hemispatial neglect is the most promising. Whether it might truly be considered evidence for HO theories is an issue deserving further treatment. It is clear, however, that there are obstacles to overcome before it could serve as such. Given this, and because the prefrontal lesion evidence reviewed in Sect. 3.1 seems to not include instances in which an integral area has been damaged, the prefrontal lesion evidence overall counts against HO theories.

There is, however, one question needing addressed. What explains the *minor* deficits in visual consciousness some prefrontal patients apparently have, if not the damaging of an integral area? This question is answered next, in the course of examining other relevant evidence.

[25] There is another lesion case some HO theorists (Lau and Rosenthal) have thought to support HO theories: At the 14th annual meeting of the Association for the Scientific Study of Consciousness (June 2010, Toronto), neuroscientist Robert Knight briefly discussed a subject with large bilateral prefrontal lesions who was only minimally responsive to stimuli. However, since this is an unpublished study, many relevant details are unavailable, making it difficult to say anything with confidence about what it means for HO theories. Nonetheless, my guess would be that such a study fails to support HO theories for reasons similar to why the case of akinetic mutism does (see Sect. 3.1).

3.2.2 *Prefrontal activity and visual consciousness*

Another line of evidence taken to support HO theories consists of studies showing correlations between visual consciousness and PFC activity. Consider, as an example, another study involving visual masking (Lau and Passingham 2006). Often when a stimulus is masked, the subject nonetheless unconsciously perceives it. In the present study, such unconscious perception was demonstrated by subjects performing above-chance when asked to identify the target stimulus in a forced-choice question. Lau and Passingham used fMRI to compare the neural activation in those cases where subjects reported seeing the target stimulus with those cases in which they reported not seeing it. On average, the dlPFC showed significantly more activation when subjects reported seeing the target stimulus. Intriguingly, Lau and Passingham also identified two masking intervals where performance was *matched* between those who consciously perceived the target and those who unconsciously perceived it. Again, the dlPFC showed increased activation when the stimulus was consciously perceived.

Other studies have the same flavor. Dehaene et al. (2001) found that masked and unmasked words brought about roughly the same amount of activation in visual areas, while unmasked words also caused greater PFC activity. In another experiment, Sahraie et al. (1997) scanned the brain of blindsighted subject GY while presenting motion stimuli to either his blind or non-blind hemifield.[26] Increased dlPFC activity was found in those cases where GY reported being conscious of motion. Other data show similar associations between visual consciousness and the PFC (Lumer et al. 1998; Marois et al. 2004; Maniscalco et al. 2009).[27]

There are three available interpretations of these correlational data, along with the prefrontal lesion data offered in support of HO theories (those just examined above). According to one interpretation, these data show that PFC activity is *necessary* for visual consciousness: PFC activity correlates with the stimulus being consciously perceived because PFC activity is *directly responsible* for its being consciously perceived (Sahraie et al. 1997; Sergent and Dehaene 2004; Maniscalco et al. 2009; Lau 2010). This interpretation is of course friendly to HO theories (Kriegel 2007; Lau and Rosenthal 2011), as the HO theorist can argue that, in those cases where the stimulus is consciously perceived, the increased PFC activity reflects the production of an extra HO state; i.e., the HO state responsible for the stimulus being perceived consciously.

In another interpretation, the above data show, not that the PFC is necessary for visual consciousness, but rather only that it plays an indirect role in determining what visual states become conscious (Pollen 2008). According to this interpretation, activity in non-PFC areas (probably extrastriate sensory areas) is sufficient for constituting visually conscious states (Pollen 1995; Zeki 2003; Block 2005; Lamme 2006),[28] while the PFC can only alter what visual states become conscious, doing so

[26] GY is probably the blindsighted subject who has been most thoroughly studied; see Weiskrantz (1997).

[27] For reviews of data suggesting a close relationship between the PFC and visual consciousness, see Rees et al. (2002), Dehaene and Changeux (2011).

through use of feedback connections to modify activity in sensory areas (perhaps by biasing neural competition; Mack and Rock 1998). In the above studies, then, the increased prefrontal activity comes from the PFC influencing activity in the visual areas, thereby causing the stimulus to be perceived consciously.

In a third interpretation, the above data only show the PFC necessary for the *report* of visually conscious states (Prinz 2000; Zeki 2003; Block 2005), it again being activity in sensory areas alone that is sufficient for visually conscious states. The proponent of this interpretation can argue that, in those instances where a conscious representation of a stimulus is especially fleeting (as when a mask is employed), the PFC must make more of an "effort" to bring the content of that representation into a report; this being what explains the extra PFC activity in cases where subjects report being conscious of the stimulus.

We have three prima facie plausible interpretations of the data showing connections between the PFC and visual consciousness. One hypothesizes the PFC to be necessary for consciousness (and therefore possibly contain integral areas), two do not. Which should we prefer? Well, the PFC lesion data examined in Sects. 3.1 and 3.2.1 showed that lesions to the PFC do not bring about the kinds of deficit in visual experience we would expect, were the PFC necessary for visual consciousness. This seems to show that the PFC *cannot* be necessary for visual consciousness, meaning we should prefer either of the other two interpretations.

I will put this another way, appealing to the asymmetry between confirmation and disconfirmation (Popper 1959). In the case of data showing correlations between PFC activity and visual experience, what we have are *confirmations* of the hypothesis that the PFC is necessary for visual experience, since if the PFC were necessary, we should see such correlations (Lau 2008). But we should also expect such correlations if the PFC is merely able to determine which visual states become conscious, or what contents show up in reports. So the data confirm these hypotheses as well.

However, in the case of the PFC lesion data surveyed in the last two subsections, what we have is a disconfirmation, and therefore *falsification*, of the hypothesis that the PFC is necessary for visual experience: If the PFC was necessary for visual experience, then lesions to the PFC should, in at least some cases, bring about striking changes in one's visual experience. But prefrontal lesions bring about no such changes. Given this, the PFC lesion data *cannot* be reconciled with any interpretation in which the PFC is necessary for visual experience. And so we should prefer either of the interpretations that hold the PFC to be not necessary for visual experience.

Given that the PFC contains integral areas only if it were necessary for visual experience, the data showing connections between visual experience and PFC activity do not help the HO theorist. This means that, if the PFC lesion data are to not count strongly against HO theories, there should be some flaw in the reasoning I advanced so far. In the next and final section, we consider whether this is the case.

[28] Given certain background systems are functioning normally, such as areas in the upper brain stem and thalamus; see Block (2009) for discussion.

4 Objections

Above (Sect. 2.2) I assumed that an integral area being disabled would bring about striking deficits in visual experience. One might doubt this for two reasons. The first concerns the possibility of there being more than one integral area: If, though an integral area had been disabled, there was another functioning integral area, then the person would continue to have visually conscious states. Perhaps this could mask the loss of an integral area. The second reason concerns the ability of brain areas to take on the function of lesioned areas (a form of "neuroplasticity"): Maybe some subjects in the above lesion studies actually had an integral area damaged, but did not experience deficits in visual experience because another brain area had taken on the job of producing (the right kind of) HO states.

Additionally, one might want to re-open the issue of whether there might be integral areas outside the PFC, on the basis of the lesion data discussed above. The idea here is that the ability of subjects to perform normally on visual examinations is evidence for them still having the ability to monitor their mental states, and therefore produce HO representations. Perhaps whatever brain area (or network of areas) accounting for this ability is entirely[29] outside the PFC. Were this true, we would need to see the results of lesions to this area (or areas) before thinking we had a strong case against HO theories.

Below, I discuss these three objections.

4.1 An extra integral area would not prevent deficits in visual experience

I discussed above how the dlPFC is considered by many HO theorists as a prime candidate for an integral area, and we also saw how even bilateral lesions to the dlPFC fail to cause the expected conscious deficits. However, Kriegel has argued that, if the dlPFC was lesioned, but no deficits resulted, this would not necessarily show that it is not an integral area, because it might be that "the dlPFC is not the only seat of higher-order representations, and when it is incapacitated, another brain area can still produce such representations" (2007, p. 910). Were this a possibility, then perhaps the prefrontal lesion data examined above do not disconfirm HO theories. One could argue that, though the prefrontal lesion data contain instances in which an integral area has been disabled, the subjects do not experience deficits in visual consciousness because of another functioning integral area or areas.[30]

To help assess this objection, I introduce the idea of "redundancy": If, at time *t*, there are two HO representations H and H*, which both have some lower-order state L as their content, let us say that H and H* *redundantly* represent L.[31] Now, the

[29] This is to say that the network of areas does not contain, as an essential component, any area in the PFC (see Sect. 2.1).

[30] Naturally, this objection might not get off the ground in the case of the patient who had his entire PFC removed (Sect. 3.1), but I put this aside in considering the objection.

[31] This idea of redundancy would need modification were we to consider an (at least superficially) different version of HO theory, in which it is but one HO state providing for all of a subject's conscious states. (Instead of a large number of HO states, there would be just one HO state with complex contents.) In this case, the redundancy of representation all happens within the same HO state: Different components of the HO state would have the same lower-order state as their content.

question we are considering is, when some integral area is lesioned, could the presence of an extra integral area (or areas) prevent the appearance of some remarkable change in visual consciousness? It seems that it could do this only if there was significant redundancy of representation. More specifically, it could prevent this only if, prior to the lesion, a significant number of the lower-order states represented by the now-lesioned area had been redundantly represented by the extra integral area.[32]

To see why, let us consider what happens if there is not significant redundancy: Since an integral area was disabled, there would be a reduction in the number of HO states routinely produced. And if, prior to the lesioning, most of the lower-order states that the integral area had represented were not redundantly represented, this would lead to a deficit of visually conscious states. The subject's ongoing visual experience would be henceforth bereft of many of the conscious states the integral area had made possible. Relative to what the subject's typical visual experience had been like prior to the lesion, there would now be "gaps."

Given this, if an extra integral area is to forestall the appearance of manifest deficits in one's visual experience, there needs to have been redundancy of representation. How much redundancy? It is reasonable to think that close to all of those conscious states that had been represented by the now-lesioned area would need to have been redundantly represented by the extra integral area.

However, we are hard-pressed to find evolutionary justification for such widespread redundancy. Consider the richness of our visual phenomenology (see, e.g., Carruthers 2000, p. 221). It is plausible to think that constructing such a visual experience requires many, many visually conscious states; meaning that, if HO theory is true, building a visual experience such as ours requires many, many HO states.[33] Carruthers (2000, pp. 213–222; 2007) has argued that there is no evolutionary justification for so many cognitive resources being dedicated to the production of HO states.[34] Notice now that whatever merit this argument possesses is amplified in the present case, for it seems even harder to find justification for so many cognitive resources being dedicated—not just to representing lower-order states—but also *redundantly* representing them.[35]

In this subsection, we have considered whether an extra integral area (or areas) would preclude the appearance of deficits in consciousness. But, as we have seen, if

[32] It is also possible that the extra integral area never redundantly represented the lower-order states, but rather came on-line only after the first integral area was damaged (it being a "back up" system of sorts). Instances in which one brain area takes on the function of another are examples of "neuroplasticity" (Grafman 2000). I postpone consideration of neuroplasticity until the next subsection.

[33] Or just one HO state with *very* complex contents (see fn. 31).

[34] This is, in fact, Carruthers' primary motivation for adopting a dispositionalist form of HO theory, according to which a mental state is conscious even if there is merely a disposition for an HO state to be formed about that mental state.

[35] It has been argued that our visual phenomenology is not as rich as it seems (Dennett 1991; O'Regan 1992; Simons 2000; but see Block 2001), and some HO theorists have availed themselves of these arguments in hopes of minimizing the amount of HO states required for the typical visual experience (Weisberg 1999; Gennaro 2004; but see Carruthers 2000, pp. 299–301). Perhaps a similar strategy could be successfully adopted in response to the problem of redundant representation that I have raised here, but this remains to be seen.

there was not significant redundancy in representation of lower-order states, we can count on the lesioning of an integral area to produce noticeable lacunae in the subject's visual experience. And it is evolutionarily implausible that cognitive resources would be dedicated to such widespread redundancy.

4.2 Neuroplasticity would not prevent deficits in visual consciousness

Neuroplasticity is the general ability of the brain to reorganize itself in response to experience or damage (Grafman 2000). Because of neuroplasticity, sometimes a brain area neighboring a lesioned area (or in the same location in the other hemisphere) will take on functions lost in the lesion. Neuroplasticity could give reason to think that, even if the PFC contains integral areas, we should not necessarily expect prefrontal lesions to produce deficits in visual experience. Kriegel seems to have this in mind when discussing how a cingulotomy (ablation of part of the ACC) does not affect consciousness. Kriegel argues that the ACC could nonetheless be an integral area, since it is possible that, "although the ACC performs monitoring functions in healthy subjects, the functions are recovered by another brain area if the ACC is incapacitated" (2007, p. 910). Likewise, Lau and Rosenthal (2011) cite evidence that, in cases of unilateral damage, "the undamaged hemisphere can dynamically 'take over' and contribute to functions normally carried out by the damaged side" (p. 369). Were these possibilities, then perhaps there are integral areas in the PFC after all, lesion evidence notwithstanding.

When considering this idea, we should first note that recovery of function would need to be complete (or near complete) if it were to prevent a subject with a lesioned integral area from having a notably different visual experience, for reasons similar to those discussed in the last subsection: If the recovery was not complete (or near complete), then HO states would not be produced in sufficient quantities, causing the appearance of "gaps" in the subject's ongoing visual experience. However, while the effects of neuroplasticity are sometimes impressive, it is unlikely that it could have prevented dramatic deficits in the visual experience of subjects in the studies discussed above, had they a damaged integral area.

In younger patients, the powers of neuroplasticity are legendary. Vargha-Khadem et al. (1997) report on a child of nine who had his entire left cortical hemisphere removed. Though the left hemisphere is normally where language abilities are predominantly localized, the boy nonetheless learned to use and understand language within a few years. However, in adult subjects—like those in the PFC lesion studies—the effects of neuroplasticity are comparatively muted, it being rare that a function lost due to a lesion is completely recovered (Frost et al. 2003). And in many cases of apparent recovery of function, it is not due to neuroplasticity, but rather what is known as "compensatory masquerade": The subject has learned to approach tasks in a novel way, one not relying on functions lost in the lesion (Grafman and Litvan 1999; Grafman 2000).

Overall, it looks unlikely that neuroplasticity is powerful enough to have prevented the appearance of deficits in visual experience in the case of subjects in the PFC lesion studies examined above. This also brings to light what is missing in the evidence for "dynamic neuroplasticity" to which Lau and Rosenthal appeal

(2011). In Voytek et al. (2010), it was found that, in subjects with unilateral lesions, increased contralesional prefrontal activity often accompanied correct responses, suggesting undamaged areas were contributing to tasks previously handled by the now-lesioned area. However, this study did not show the undamaged PFC to enable a *normal* level of performance. But this is what the study needed to show, if it were to suggest that neuroplasticity could prevent deficits in visual experience after an integral area is damaged.

There is also reason to think that, in the case of subjects in the PFC lesion studies discussed above, neuroplasticity did not prevent any such deficits: We can guess that, if one of these subjects had a dramatic recovery in her ability to produce (the right kind of) HO states, she would *probably* also have had a dramatic recovery of whatever other functions were lost because of the lesion. However, in the PFC studies discussed above, we lack evidence for anything like this: The subjects still suffer, in no insignificant way, from deficits in executive function.

4.3 Residual abilities to produce higher-order representations

The final objection comes from the observation that patients in the PFC lesion studies did not present with deficits during examinations of their visual abilities. These examinations typically consist of asking subjects to produce reports about what they are seeing (e.g., what letters they see on a Snellen chart). One might think that performing such a task involves producing an HO representation of a mental state; namely, of the mental state having the stimulus as its content (see, e.g., Rosenthal 1997, pp. 746–747). (According to this line of thought, reporting that there is, for example, a "Z" on the chart involves forming an HO representation of the visual state that has the "Z" as its content.) If this is true—the objection continues—and it is also true that the prefrontal patients could perform tasks like these, then each of these subjects must still have had some brain area (or areas) able to produce HO states. And if these subjects have a brain area (or areas; I will stop saying this each time) able to produce HO representations, then perhaps the prefrontal lesion evidence does not yet count against HO theories. Before it did, we would need reason for thinking the area producing the HO states is not an integral area.

The first question to ask about this objection concerns its key assumption, which is that reporting on stimuli requires forming HO states. The idea here seems to be that, for one to report on a stimulus, one needs not only a representation of the stimulus, but also a representation of the representation of the stimulus. But why think this? To me, this appears to mix up what is required for reporting on a stimulus with what is required for reporting that one is perceiving a stimulus. And so it is not clear that, just because subjects can report on visual stimuli, they also must have some brain area able to produce HO representations.

However, let us assume that reporting on a stimulus requires HO representations, so we may see what follows. If true, then each subject in the prefrontal lesion studies possessed some area still able to produce HO states. We will refer to this as the *spared* area. The question arises now as to whether we should think that, in the case of the PFC-lesioned subjects, the spared area is in the PFC. This is important,

since a spared area in the PFC would not help the HO theorist: The prefrontal lesion evidence examined above provided reason to think that there are no integral areas in the PFC. So the objection depends on whether it is plausible that the spared area is outside the PFC.

Before considering this issue, it should be noted that nothing prevents us from holding that it is some area in the PFC responsible for subjects' residual ability to produce HO states. There are two ways that this might be the case. First, the spared area could be some prefrontal area other than the lesioned area. Given the PFC's association with monitoring activities, it is of course plausible that the spared area would be in the PFC. Second, it could be that the spared area is the lesioned area itself. This could happen if the damaged area still had a *partial* ability to produce HO states. Such a subject, while unable to produce the same *volume* of HO states as before, could still generate reports about her mental states.[36] In sum, we need not appeal to areas outside the PFC to explain subjects' ability to report on their perceptions.

On the other hand, if the possibility of a spared area outside of the PFC is to help HO theories, it should be plausible that there is a potential integral area (or network of areas) entirely outside the PFC. This returns us to the topic of Sect. 1, which concerned where in the brain we should expect (the right kind of) HO states to be produced. As seen above, many commentators have thought that any such areas will be in the PFC, since this is where monitoring duties are performed. However, given subjects' retained ability to report on their mental states, one might wish to re-visit the issue of whether (the right kind of) HO states could be produced without the PFC. After all, what I presented in Sect. 1 looked more like reasons for thinking that the PFC is involved in producing (the right kind of) HO representations, rather than reasons for thinking that areas outside the PFC are not able to (i.e., not able to do so without help from the PFC). And one might observe that, for a decisive case against HO theories, the latter task needs accomplished.

Indeed, one HO theorist (Gennaro 2012)—mindful of the prefrontal lesion data examined above—has offered a tentative case for there being integral areas outside the PFC, suggesting that the PFC is required only for producing HO states that have *other* HO states as their target. There is not space here to give his multiple proposals fair treatment. However, it is worth noting that they come in two categories, neither of which are obviously helpful to the HO theorist. The first includes scientific theories according to which the ACC is necessary for having conscious states (Damasio 1999; Newen and Vogeley 2003). But, as seen above, the ACC is probably not an integral area. The second category includes theories according to which it is recurrent neural activity within sensory areas that is necessary and sufficient for visually conscious states (Edelman and Tonini 2000; Bullier 2001; Pascual-Leone and Walsh 2001). But these sorts of views are often thought to *conflict* with HO theories, on the supposition that such activity would *not* involve the production of HO states (Block 2007; Lau and Brown forthcoming). Thus, there are two issues to be worked out before we can see how this latter proposal bears on HO theories. The first is whether such recurrent neural

[36] Note that, in such a scenario, if the lesioned area was an integral area, we would expect the subject both to experience deficits in visual consciousness (she would have lost the ability to produce HO states in the same quantity), and have the ability to report that she was experiencing such deficits; thereby allowing her condition to be discovered.

activity actually is necessary and/or sufficient for visual experience. The second is whether such activity plausibly involves the instantiation of (the right kind of) HO states. I leave this for future research.[37]

I close with one general observation. This is that it seems like those HO theories building a lot of complexity into (the right kind of) HO state will be those theories most likely to be vulnerable to the argument presented in this paper. Consider, for example, HO *thought* theory (Gennaro 1996; Carruthers 2000; Rosenthal 2002). According to how this theory is commonly explicated, a mental state is conscious (if and) only if it is targeted by a conceptual and belief-like mental state. Furthermore, the to-be-conscious state must be represented as a mental state of one's self.[38] Having an HO thought, then, requires significant cognitive sophistication[39]: One must represent mental states; represent oneself; and do all of this conceptually. It also involves building *complex* representations: One must represent mental states as *mental states of oneself*. Given the apparent cognitive sophistication involved in an HO thought, and given that the PFC is where sophisticated cognitive functions are thought to be carried out (Fuster 2002), it seems especially likely that, if HO thought theory is true, there are integral areas in the PFC. More specifically, it seems especially likely that the production of (the right kind of) HO states essentially involves one or more areas in the PFC.

More generally, the lesson seems to be this: The more sophistication an HO theorist builds into (the right kind of) HO state, the more problematic it is that prefrontal lesions do not produce deficits in visual experience. On the other hand, an HO theorist that makes HO states not so cognitively demanding—e.g., by not requiring the mental states to be represented as states of oneself—has a better chance of finding some area

[37] Flohr offers a neurologically based version of HO theory, one appearing not to necessarily implicate the PFC. The theory is pitched at the neuronal level, hypothesizing the *NMDA synapse* to be what ultimately enables building "higher-order, self-reflexive representational structures" (1999, p. 255). The NMDA synapse is qualified to do this, argues Flohr, because of its ability to produce both transient and permanent changes in synaptic weights. Given that the theory is (for the most part) based at the neuronal level, it appears not wed to it being any particular brain areas that produce the HO states, and therefore presents one promising avenue for the HO theorist to explore.

[38] Here, in their own words, is how these philosophers describe an HO thought: According to Rosenthal, "a mental state is conscious just in case it is accompanied by a non-inferential…assertoric thought to the effect that one is in that very state" (2002, p. 410). Carruthers holds that conscious states must be disposed to bring about an "activated belief (possibly a non-conscious one) that I have M, and to cause it non-inferentially" (2000, p. 227). Gennaro says that an HO thought is "the thought 'that *I* am in M'…What exactly is its content? Of course it involves reference to a subject ('I') and to a mental state ('M')" (1993, p. 58).

[39] Some HO thought theorists have attempted to minimize this conclusion (see Gennaro 1993, 1996; Rosenthal 2002), this being an attempt to respond to philosophers who have argued that, if HO thought theory was true, then infants and animals would not be conscious, since these beings lack the cognitive sophistication required for HO thoughts (Carruthers 1989, 1992, 1999; Dretske 1995; Tye 1995). Generally speaking, the HO thought theorists in question have argued that the concepts used in an HO thought can be rather crude, including the concept of oneself, and so producing HO thoughts is not as cognitively demanding as we might first think.

I think the arguments these HO theorists have offered do not make it look as if there could be areas outside the PFC that are able to produce HO thoughts. While not having space to go into this, I will nonetheless note that nothing these HO thought theorists have said so far chips away at the most cognitively-demanding aspect of an HO thought, which is its complexity: An HO thought is a representation in which a mental state is represented *as being a state of oneself*.

(or network of areas) outside the PFC that might produce (the right kind of) HO states. If successful in doing so, she would have grounds for claiming that the prefrontal lesion data do not count strongly against her version of HO theory.

However, even with HO theories not building too much sophistication into (the right kind of) HO states, it is no foregone conclusion that some area can be found outside of the PFC that qualifies as a potential integral area. Indeed, we saw above that it looks as if the best candidates for producing (the right kind of) HO states are areas in the PFC. And so finding a suitable brain area outside of the PFC might be difficult for HO theorists of any stripe.

5 Conclusion

The starting point for this paper was the observation that HO theories are amenable to empirical confirmation or disconfirmation by lesion data: If some HO theory is true, then it should be the case that, when those brain areas that produce (the right kind of) HO states are lesioned, this brings about remarkable changes to one's ongoing visual experience. Many researchers have thought that the production of HO states will require the PFC. But it appears that areas in the PFC can be lesioned without producing the deficits in experience we would expect, were it the case that an integral area was damaged. One might wish to explain these data in a way friendlier to HO theories, claiming that neuroplasticity, or the presence of more than one integral area, is the reason that these subjects fail to have deficits in visual experience. But I have argued that these sorts of explanation either suggest massive redundancy of representation, or overestimate the power of neuroplasticity. One might also suggest that subjects in the prefrontal lesion studies had retained the ability to produce HO representations, claiming this as a reason to think that there is some brain area that is still a candidate for being an integral area. But I have argued that even if we grant that these subjects still possessed some area able to produce HO representations, this area would likely be in the PFC, and hence not an integral area. I also considered evidence showing correlations between PFC activity and visual consciousness. But these data are compatible with explanations not hypothesizing the PFC to be necessary for visual experience, and I argued that these other explanations should be preferred, given the PFC lesion data. Overall, the prefrontal lesion data appear to constitute noteworthy evidence against HO theories, especially those HO theories appearing to require a cognitively sophisticated sort of HO state.

Acknowledgments I would like to thank Uriah Kriegel and Shaun Nichols for comments on earlier drafts of this paper, and Rocco Gennaro, Hakwan Lau, Brian Fiala, and numerous others for illuminating discussions on this topic.

References

Alvarez, J., & Emory, E. (2006). Executive function and the frontal lobes: A meta-analytic review. *Neuropsychology Review, 16*(1), 17–42.

Andrewes, D. (2001). *Neuropsychology: From theory to practice*. New York: Psychology.

Armstrong, D. (1968). *A materialist theory of the mind*. London: Routledge.

Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience, 4*, 829–839.
Barceló, F., et al. (2000). Prefrontal modulation of visual processing in humans. *Nature Neuroscience, 4*, 399–403.
Bartolomeo, P., Thiebaut de Schotten, M., & Doricchi, F. (2007). Left unilateral neglect as a disconnection syndrome. *Cerebral Cortex, 17*(11), 2479–2490.
Beeckmans, J. (2007). Can higher-order representation theories pass scientific muster? *Journal of Consciousness Studies, 14*(9–10), 90–111.
Bird, C., et al. (2004). The impact of extensive medial frontal lobe damage on "Theory of mind" and cognition. *Brain, 127*, 914–928.
Block, N. (1995). On a confusion about a function of consciousness. *Behavioral and Brain Sciences, 18*, 227–287.
Block, N. (2001). Paradox and cross-purposes in recent work on consciousness. *Cognition, 79*, 197–219.
Block, N. (2005). Two neural correlates of consciousness. *Trends in Cognitive Sciences., 9*(2), 46–52.
Block, N. (2007). Consciousness, accessibility, and the mesh between psychology and neuroscience. *Behavioral and Brain Sciences, 30*, 481–548.
Block, N. (2009). Comparing the major theories of consciousness. In M. Gazzaniga (Ed.), *The cognitive neurosciences IV*. Cambridge: MIT Press.
Botvinik, M., et al. (2001). Conflict monitoring and cognitive control. *Psychological Review, 108*, 624–652.
Bressler, S., & Menon, V. (2010). Large-scale brain networks in cognition: Emerging methods and principles. *Trends in Cognitive Sciences, 14*, 277–290.
Brickner, R. (1936). *The intellectual function of the frontal lobes*. New York: Macmillan.
Brodmann, K. (1909). *Vergleichende Lokalisationslehre der Grosshimrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues*. Leipzig: Barth.
Brown, R. (2011). The myth of phenomenological overflow. *Consciousness and Cognition*. doi: 10.1016/j.concog.2011.06.005.
Bullier, J. (2001). Feedback connections and conscious vision. *Trends in Cognitive Sciences, 9*, 369–370.
Byrne, A. (2004). What phenomenal consciousness is like. In R. Gennaro (Ed.), *Higher-order theories of consciousness: An anthology*. Philadelphia: John Benjamins.
Cairns, H., et al. (1941). Akinetic mutism with an epidermoid cyst of the 3rd ventricle. *Brain, 64*, 273–290.
Carruthers, P. (1989). Brute experience. *Journal of Philosophy, 86*, 258–269.
Carruthers, P. (1992). *The animals issue*. Cambridge: Cambridge University Press.
Carruthers, P. (1999). Sympathy and subjectivity. *Australasian Journal of Philosophy, 77*, 465–482.
Carruthers, P. (2000). *Phenomenal consciousness: A naturalistic theory*. Cambridge: Cambridge University Press.
Carruthers, P. (2004). HOP over FOR, HOT theory. In R. Gennaro (Ed.), *Higher-order theories of consciousness: An anthology*. Philadelphia: John Benjamins.
Carruthers, P. (2007). Higher-order theories of consciousness. In *Stanford encyclopedia of philosophy*. http://plato.stanford.edu/entries/consciousness-higher/. Accessed 18 Mar 2013.
Carruthers, P. (2010). Introspection: Divided and partly eliminated. *Philosophy and Phenomenological Research, 80*, 76–111.
Chalmers, D. (1995). Facing up to the problem of consciousness. *Journal of Consciousness Studies, 2*(3), 200–219.
Damasio, A. (1994). *Descartes' error: Emotion, reason and the human brain*. New York: Putnam.
Damasio, A. (1999). *The feeling of what happens*. New York: Harcourt Inc.
Damasio, A., & Van Hoesen, G. (1983). Focal lesions of the limbic frontal lobe. In K. Heilman & P. Satz (Eds.), *Neuropsychology of human emotion*. New York: Guilford Press.
Deacon, T. (1990). Rethinking mammalian brain evolution. *American Zoology, 30*, 629–705.
Dehaene, S., & Changeux, J. (2011). Experimental and theoretical approaches to conscious processing. *Neuron, 70*, 200–227.
Dehaene, S., et al. (2001). Cerebral mechanisms of word masking and unconscious repetition priming. *Nature Neuroscience, 4*(7), 752–758.
Del Cul, A., et al. (2009). A causal role of prefrontal cortex in the threshold for access to consciousness. *Brain, 132*(9), 2531–2540.
Dennett, D. (1991). *Consciousness explained*. Boston: Little Brown.

Devinsky, O., Morrell, M., & Vogt, B. (1995). Contributions of anterior cingulate cortex to behaviour. *Brain, 118*, 279–306.

Dretske, F. (1995). *Naturalizing the mind*. Cambridge: MIT Press.

Driver, J., & Vuilleumier, P. (2001). Perceptual awareness and its loss in unilateral neglect and extinction. *Cognition, 79*, 39–88.

Edelman, G., & Tononi, G. (2000). Reentry and the dynamic core: Neural correlates of conscious experience. In T. Metzinger (Ed.), *Neural correlates of consciousness: Empirical and conceptual questions*. Cambridge: MIT Press.

Ferrier, D., & Turner, W. (1898). An experimental research upon cerebro-cortical afferent and efferent tracts. *Philosophical Transactions of the Royal Society of London B, 190*, 1–44.

Flohr, H. (1999). NMDA receptor-mediated computational processes and phenomenal consciousness. In Metzinger, T. (Ed.), *Neural correlates of consciousness: Empirical and conceptual questions*. Boston: MIT press.

Frith, U., & Frith, C. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society of London B, 358*, 459–473.

Frost, S., et al. (2003). Reorganization of remote cortical regions after ischemic brain injury: A potential substrate for stroke recovery. *Journal of Neurophysiology, 89*, 3205–3214.

Fuster, J. (2002). Frontal lobe and cognitive development. *Journal of Neurocytology, 31*, 373–385.

Gallagher, H., & Frith, C. (2003). Functional imaging of "theory of mind". *Trends in Cognitive Science, 7*, 77–83.

Gennaro, R. (1993). Brute experience and the higher-order thought theory of consciousness. *Philosophical Papers, 22*(1), 51–69.

Gennaro, R. (1996). *Consciousness and self-consciousness*. Amsterdam: John Benjamins.

Gennaro, R. (2004). In R. Gennaro (Ed.), *Higher-order theories of consciousness: An anthology*. Philadelphia: John Benjamins.

Gennaro, R. (2012). *The consciousness paradox*. Cambridge: MIT Press.

Grafman, J. (2000). Conceptualizing functional neuroplasticity. *Journal of Communication Disorders, 33*(4), 345–356.

Grafman, J., & Litvan, I. (1999). Evidence for four forms of neuroplasticity. In J. Grafman & Y. Christen (Eds.), *Neuronal plasticity: Building a bridge from the laboratory to the clinic*. New York: Springer.

Harman, G. (1990). The intrinsic quality of experience. *Philosophical Perspectives, 4*, 31–52.

Heath, R., et al. (1949). Visual apparatus: Visual fields and acuity, color vision, autokinesis. In F. Mettler (Ed.), *Selective partial ablation of the frontal cortex: A correlative study of its effects on human psychotic subjects*. New York: P. B. Hoeber.

Husain, M., & Kennard, C. (1996). Visual neglect associated with frontal lobe infarction. *Journal of Neurology, 243*, 652–657.

Jacob, P., & De Vignemont, F. (2010). Spatial coordinates and phenomenology in the two-visual systems model. In N. Gangopadhyay (Ed.), *Perception, action, and consciousness: Sensorimotor dynamics and two-visual systems*. Oxford: Oxford University Press.

Kennard, M. (1939). Alterations in response to visual stimuli following lesions of frontal lobe in monkeys. *Archives of Neurology and Psychiatry, 41*(6), 1153–1165.

Koch, C., & Tsuchiya, N. (2007). Attention and consciousness: Two different processes. *Trends in Cognitive Science, 11*(1), 16–22.

Kriegel, U. (2007). A cross-order integration hypothesis for the neural correlate of consciousness. *Consciousness and Cognition, 16*, 897–912.

Kriegel, U. (2009). *Subjective consciousness*. Oxford: Oxford University Press.

Lamme, V. (2006). Separate definitions of visual consciousness and visual attention: A case for phenomenal awareness. *Neural Networks, 17*, 861–872.

Lashley, K., & Clark, G. (1946). The cytoarchitecture of the cerebral cortex of ateles: A critical examination of architectonic studies. *Journal of Comparative Neurology, 85*(2), 223–305.

Lau, H. (2008). A higher order Bayesian decision theory of consciousness. *Progress in Brain Research, 168*, 35–48.

Lau, H. (2010). Theoretical motivations for investigating the neural correlates of consciousness. *Wiley Interdisciplinary Reviews: Cognitive Science, 2*(1), 1–7.

Lau, H., & Brown, R. (forthcoming). The emperor's new phenomenology? The empirical case for conscious experiences without first-order representations. In A. Pautz & D. Stoljar (Eds.), Festschrift for Ned Block. Cambridge: MIT.

Lau, H., & Passingham, R. (2006). Relative blindsight in normal observers and the neural correlate of visual consciousness. *Proceedings of the National Academy of Science USA, 103*, 18763–18768.
Lau, H., & Rosenthal, D. (2011). Empirical support for higher-order theories of conscious awareness. *Trends in Cognitive Sciences, 15*(8), 365–373.
Liu, G., et al. (1992). Dissociated perceptual-sensory and exploratory-motor neglect. *Journal of Neurology, Neurosurgery and Psychiatry, 55*, 701–706.
Locke, J. (1690). *An essay concerning human understanding*. (A number of different editions of this book are available.)
Lumer, E., Friston, K., & Rees, G. (1998). Neural correlates of perceptual rivalry in the human brain. *Science, 280*, 1930–1934.
Lycan, W. (1996). *Consciousness and experience*. Cambridge: MIT Press.
Lycan, W. (2001). A simple argument for a higher-order representation theory of consciousness. *Analysis, 61*, 3–4.
Mack, A., & Rock, I. (1998). *Inattentional blindness*. Cambridge: MIT Press.
Macknik, S. (2006). Visual masking approaches to visual awareness. *Progress in Brain Research, 155*, 177–215.
Maeshima, S., et al. (1995). Unilateral spatial neglect due to a haemorrhagic contusion in the right frontal lobe. *Journal of Neurology, 242*, 613–617.
Maniscalco, B., et al. (2009). Theta-burst transcranial magnetic stimulation to the prefrontal cortex impairs metacognitive visual awareness. *Journal of Vision, 9*(8), 764.
Marois, R., Yi, D., & Chun, M. (2004). The neural fate of consciously perceived and missed events in the attentional blink. *Neuron, 41*, 465–472.
Mega, M. S., & Cohenour, R. C. (1997). Akinetic mutism: Disconnection of frontal–subcortical circuits. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology, 10*, 254–259.
Mesulam, M. (1999). Spatial attention and neglect: parietal, frontal and cingulate contributions to the mental representation and attentional targeting of salient extrapersonal events. *Philosophical Transactions of the Royal Society London B, 354*, 1325–1346.
Mettler, F. (1944). Physiologic effects of bilateral simultaneous frontal lesions in the primate. *The Journal of Comparative Neurology, 81*(2), 105–136.
Miyake, A., et al. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41*, 49–100.
Nagel, T. (1974). What is it like to be a bat? *Philosophical Review, 83*, 435–456.
Nakamura, R. K., & Mishkin, M. (1986). Chronic "blindness" following lesions of nonvisual cortex in the monkey. *Experimental Brain Research, 63*, 173–184.
Nauta, W. (1972). Neural associations of the frontal cortex. *Acta Neurobiologiae Experimentalis (Wars), 32*(2), 125–140.
Newen, A., & Vogeley, K. (2003). Self-representation: Searching for a neural signature of self-consciousness. *Consciousness and Cognition, 12*, 529–543.
O'Regan, J. (1992). Solving the 'real' mysteries of visual perception: the world as an outside memory. *Canadian Journal of Psychology, 46*, 461–488.
Owen, A. (1997). The functional organization of working memory processes within human lateral frontal cortex: The contribution of functional neuroimaging. *European Journal of Neurosciences, 9*, 1329–1339.
Pardo, J., et al. (1990). The anterior cingulate cortex mediates processing selection in the Stroop attentional conflict paradigm. *Proceeding of the National Academy of Sciences USA, 87*, 256–259.
Pascual-Leone, A., & Walsh, V. (2001). Fast backprojections from the motion to the primary visual area necessary for visual awareness. *Science, 292*, 510–512.
Penfield, W., & Evans, J. (1935). The frontal lobe in man: A clinical study of maximal removals. *Brain, 58*, 115–133.
Pollen, D. (1995). Cortical areas in visual awareness. *Nature, 377*(6547), 293–295.
Pollen, D. (2008). Fundamental requirements for primary visual perception. *Cerebral Cortex, 18*, 1991–1998.
Popper, K. (1959). *The logic of scientific discovery*. New York: Basic Books.
Potter, M. (1999). Understanding sentences and scenes: The role of conceptual short-term memory. In V. Coltheart (Ed.), *Fleeting memories*. Cambridge: MIT Press.
Prinz, J. (2000). A neurofunctional theory of visual consciousness. *Consciousness and Cognition, 9*, 243–259.
Rees, G., Kreiman, G., & Koch, C. (2002). Neural correlates of consciousness in humans. *Nature Reviews Neuroscience, 3*, 261–270.

Rolls, E. (2000). The orbitofrontal cortex and reward. *Cerebral Cortex, 10*, 284–294.
Rolls, E. (2004). A higher-order syntactic thought (HOST) theory of consciousness. In R. Gennaro (Ed.), *Higher-order theories of consciousness: An anthology*. Philadelphia: John Benjamins.
Rosenthal, D. (1997). A theory of consciousness. In N. Block, O. Flanagan, & G. Güzeldere (Eds.), *The nature of consciousness*. Cambridge, MA: MIT Press.
Rosenthal, D. (2002). Explaining consciousness. In D. Chalmers (Ed.), *Philosophy of mind: Contemporary and classical readings*. Oxford: Oxford University Press.
Rounis, E., Maniscalco, B., Rothwell, J. C., Passingham, R. E., & Lau, H. (2010). Theta-burst transcranial magnetic stimulation to the prefrontal cortex impairs metacognitive visual awareness. *Cognitive Neuroscience, 1*(3), 165–175.
Sahraie, A., et al. (1997). Pattern of neuronal activity associated with conscious and unconscious processing of visual signals. *Proceedings of the National Academy of Science USA, 94*, 9406–9411.
Saxe, R., Schulz, L., & Jiang, Y. (2006). Reading minds versus following rules: Dissociating theory of mind and executive control in the brain. *Social Neuroscience, 3–4*, 284–298.
Sergent, C., & Dehaene, S. (2004). Neural processes underlying conscious perception: Experimental findings and a global neuronal workspace framework. *Journal of Physiology: Paris, 98*, 374–384.
Simons, D. (2000). Attentional capture and inattentional blindness. *Trends in Cognitive Science, 4*(4), 147–155.
Turatto, M., et al. (2004). The role of the right dorsolateral prefrontal cortex in visual change awareness. *NeuroReport, 15*, 2549–2552.
Tye, M. (1995). *Ten problems of consciousness*. Cambridge: MIT Press.
Van Gulick, R. (2004). Higher-order global states (HOGS): An alternative higher-order model of consciousness. In R. Gennaro (Ed.), *Higher-order theories of consciousness: An anthology*. Philadelphia: John Benjamins.
Vargha-Khadem, F., et al. (1997). Onset of speech after left hemispherectomy in a nine-year old boy. *Brain, 120*, 159–162.
Verdon, V., et al. (2010). Neuroanatomy of hemispatial neglect and its functional components: A study using voxel-based lesion-symptom mapping. *Brain, 133*, 880–894.
Vosgerau, G., & Newen, A. (2008). Orthogonality of phenomenality and content. *American Philosophical Quarterly, 45*, 329–348.
Voytek, B., et al. (2010). Dynamic neuroplasticity after human prefrontal cortex damage. *Neuron, 68*, 401–408.
Weisberg, J. (1999). Active, thin and hot! An actualist response to Carruthers' dispositionalist HOT view. *Psyche, 5*(6). http://www.theassc.org/files/assc/2431.pdf. Accessed 18 Mar 2013.
Weiskrantz, L. (1997). *Consciousness lost and found*. Oxford: Oxford University Press.
Wills, C. (1993). *The runaway brain: The evolution of human uniqueness*. New York: Basic Books.
Zimmer, H. D. (2008). Visual and spatial working memory: From boxes to networks. *Neuroscience & Biobehavioral Reviews, 32*(8), 1373–1395.
Zeki, S. (2003). The disunity of consciousness. *Trends in Cognitive Sciences., 7*(5), 214–218.