

The Misunderstood Higher-Order Approach to Consciousness

Richard Brown¹

Hakwan Lau²

Joseph E. LeDoux³

]

¹ Philosophy Program LaGuardia Community College, City University of New York, Long Island City, NY.

² Department of Psychology and Brain Research Institute, University of California, Los Angeles, CA; Department of Psychology, and State Key Laboratory for Brain and Cognitive Sciences, the University of Hong Kong.

³ Center for Neural Science and Department of Psychology, New York University, New York, NY; Departments of Psychiatry and Child and Adolescent Psychiatry, New York University Langone Medical School, New York, NY; Nathan Kline Institute, Orangeburg, NY.

Abstract

Critics have often misunderstood the higher-order theory (HOT) of consciousness. Here we clarify its position on several issues, and distinguish it from other views such as the global workspace theory (GWT) and early sensory models, such as first-order local recurrency theory. The criticism that HOT over-intellectualizes conscious experience is inaccurate because in reality the theory assumes minimal cognitive functions for consciousness; in this sense it is an intermediate position between GWT and early sensory views, and plausibly accounts for shortcomings of both. Further, compared to other existing theories, HOT can more readily account for complex everyday experiences, such as of emotions and episodic memories, and make HOT potentially useful as a framework for conceptualizing pathological mental states.

Glossary

blindsight - The neurological phenomenon that after lesion to primary visual cortex patients deny having conscious experience for the affected regions in the visual field, and yet they can detect and discriminate visual targets presented there above chance levels.

consciousness – This term can be used to refer to different mechanisms and phenomena, such whether a person is awake versus in coma, or is having subjective experiences. Within this article we are primarily concerned states that have subjective phenomenal qualities, such as perceptions, memories, thoughts or emotions

first-order representations– These are generally simple mental representations, often about states of the sensory world. In the case of visual perception, for example, these involve activity in striate and extrastriate cortical areas, where neuronal firing can be directly driven by external stimuli such as lines, color, motion, and objects.

higher-order representations– These are generally *about* first-order representations, and are usually associated with areas of prefrontal cortex. However, higher-order theory typically proposes that the higher-order representation, while necessary for consciousness, is itself not one of which you are aware. To be aware of a higher-order representation requires an additional higher-order representation.

inner awareness (26 words) – this occurs when a one is aware of his or her own mental representations, a process that we propose involves higher-order representations.

introspection - In typical usage this refers to content or experience resulting from the act of intentional and effortful self-monitoring and evaluation of ongoing conscious experiences, as when one directs one's conscious thoughts to other conscious mental states. Higher-order theories do not require this type of elaborate, active, conscious introspection for simple experiences. Instead they appeal to a relatively passive (cognitively "leaner") processing involving the mere noticing of mental states and/or implicit, automatic, nonconscious processing.

phenomenal consciousness—When we say an individual is phenomenally conscious, we mean, following Thomas Nagel, that he or she is capable of having states for which there is something that it is like, from the individual's point of view, to be in such states. Term is itself neutral with respect to whether phenomenal consciousness depends on first- or higher-order representations, though our theoretical position is that the higher-order account is correct.

transitivity principle – This assumes that when you are in no way aware of being in some mental state there is nothing that it is like for you to be in that mental state. For example, if you are presented with a red stimulus but are in no way aware that you are seeing red, you are not consciously experiencing red. The converse of this is the principle with which higher-order theory is concerned, that consciously seeing red depends in some way on one being aware of seeing red.

Introduction

Consciousness [glossary], as used here, refers to subjective experience, or what is called phenomenal consciousness [1], as opposed to the condition of merely being awake and alert and behaviorally responsive to external stimuli. To be phenomenally conscious is for there to be something that it is like to be the entity in question -something that it is like for the entity itself [2].

Phenomenally conscious experiences are the stuff of novels, poems, and songs, the essence of being a human. It's hard to imagine what it would be like to not be sentient in the way we are. Unsurprisingly, then, the science of consciousness is currently a vibrant and thriving area of research. However, there is no generally accepted theory of the phenomena being studied, and the phenomena themselves often do not include the many of the kinds of complex experiences that we normally have in the course of day to day life, such as of our emotions and memories.

We will argue that the foundation for a viable theory of such experiences exists but has not been given the credit it deserves. We are referring to various ideas known collectively as higher-order theory (HOT). We believe that one of the main reasons HOT has been sidelined is that it has been misunderstood, and our main goal in this paper is identify and correct some of the most common misconceptions. But first, a brief description of HOT is in order.

What Is HOT?

In general, there are two features that make a theory of consciousness a higher-order theory. The first is the commitment to the claim that a mere (first-order) representation is not sufficient for conscious experiences to arise - some higher-order mechanisms are also needed. For example, having a first-order perceptual state, a state in which the brain represents something which is not a mental state, for example, something in its environment, is often crucial for the organism to respond meaningfully to external stimuli. However, a fundamental tenet of HOT is that first-order states occur unconsciously, and are not sufficient for phenomenally conscious experiences to occur [3].

This leads to the second requirement of HOT, the claim that if an organism is in a mental state, such as a perceptual state, but is in no way aware of itself as being in that state, then the organism is not in a phenomenally conscious state. This is a logical consequence of what is called the Transitivity Principle [3], [glossary]. The basic idea, according to HOT, is that conscious experiences entail some kind of minimal inner awareness [glossary] of one's ongoing mental functioning.

Misconception 1: HOT Is a Single Theory

HOT is not a single entity, as it comes in many varieties (see [4-6]), and the different versions disagree on what may be the exact mechanisms for the inner awareness being postulated. This diversity of HOT has not always been acknowledged by critics; few arguments apply to the entire family of theories. The traditional view, sometimes attributed to John Locke and Immanuel Kant, refers to the mechanism of inner awareness as an Inner Sense, akin to perception [7, 8]. However, Inner Sense HOTs, also referred to as Higher-Order Perception theories [9], have recently fallen out of favor because of a failure to find a neural implementation of an inner sense [10]. Similarly, Dispositional HOT, which posits that the mere availability of first-order content to higher-order mechanisms accounts for consciousness [11], has been abandoned [12]. Currently active versions of HOT include the Phenomenal Self theory [13], the radical plasticity hypothesis [14], and several versions of Higher-Order Thought Theory (HOTT) [3, 6, 15-19]. HOTT postulates that the higher-order state is thought-like, and is the subject of much current discussion and debate [20-29].

Misconception 2: HOT suggests that sophisticated thoughts are necessary for conscious experiences.

One common objection to HOT is that it makes consciousness overly sophisticated [30, 31]. This criticism mainly applies to HOTT, and even in that case the criticism is misplaced. For example, critics sometimes see the invocation of higher-order thoughts in HOTT as entailing complex forms of cognition, including introspection and self-awareness, that may not be present in mammals besides humans, and in humans are more complex than would likely be required for phenomenal perceptual experience.

Confusion results in part from the ways "thought," "introspection" and "self," as used in HOT, are interpreted by critics [27]. Traditionally, introspection refers to an active process in which one becomes conscious of their inner states [32, 33]. Similarly, "self" is often used to refer to conscious awareness of one's self. The fact that higher-order philosophers tend to use propositional statements involving personal pronouns ("I see an apple") to describe the higher-order thought has led to the idea that HOT implies a conscious self that has introspective knowledge of its experiences. However, HOTT proponents typically are in fact calling upon a cognitively lean conception of both thought and self.

For example, the thoughts involved in higher-order representations are *not* usually themselves viewed as consciously experienced [3, 5, 34, 35] (but see [18,19, 36, 37]). A further step, involving a more elaborate, active, form of

introspection is typically required to be conscious of the higher-order state. To clarify this, we build on Tulving's conception of noetic vs. auto-noetic consciousness [38], which distinguishes awareness of facts (an apple is present) from temporally-grounded reflective self awareness (I know that I, in particular, am seeing an apple at this particular moment in time). Shea and colleagues recent proposal that explicit, supra-personal metacognition, which allows one to say that "I am sure" or "I am doubtful," may be unique to human consciousness [45], and might be especially relevant to auto-noetic states. It should be noted that both the leaner/passive and thicker/active versions of self and introspection relevant to HOT are distinct from more basic biological mechanisms referred to with terms such as core self [39-41] or the machinery of self [42].

With the leaner sense of introspection and self, higher-order theorists are free to speculate that non-human animals, infants, and even non-biological agents, could have the necessary kind of thoughts to have simple conscious experiences, such as conscious perceptions. Viewed this way, it is an open question as to whether they do or not. However, we are not currently in a position to answer this question empirically [37]. Further most HOTs are agnostic regarding the status of consciousness in animals, infants, and machines [Box 1]. Rather than this agnosticism being a short-coming of the theory, we see it as prudent reliance on the norms of good science. Below, we focus on HOTT (unless indicated) since most research, theorizing, and debate of late has been about this variant of HOT.

Misconception 3: HOT says consciousness is the same as metacognition or confidence

The putative higher-order mechanisms posited by HOT to be necessary for conscious experiences to occur are often confused with experimental tasks that have a somewhat higher-order 'flavor', e.g. tasks involving metacognition. In psychological experiments on metacognition, subjects are often required to give confidence ratings after a simple perceptual or memory task. While these tasks are useful experimental devices for probing mechanisms relevant to conscious awareness, most proponents of HOT do not treat metacognition and consciousness as conceptually equivalent [43, 44]. To the extent some metacognitive process is involved in consciousness, we treat this process as mostly implicit, in the sense that it is engaged at a sub-personal (and nonconscious) level. Shea and colleagues recent proposal that explicit, supra-personal metacognition may be important in some conditions [45], perhaps those involving auto-noetic, as opposed to noetic, states of consciousness (see above).

Why then do proponents of HOT employ explicit metacognitive tasks in experiments? In part this was due to an empirical observation: when neural activity in prefrontal cortex, which is known to be important for conscious perception [46], is manipulated with transcranial magnetic stimulation (TMS), explicit metacognition is also affected [47-49]. This suggests the possibility that the brain mechanisms responsible for higher-order conscious experience and explicit metacognition may be evolutionarily 'recycled' or shared. Therefore, explicit metacognition tasks can be used as a convenient tool for studying

mechanisms relevant to consciousness, even though these tasks do not represent the *sine qua non* of consciousness.

Relatedly, some claim that metacognitive tasks can be performed with nonconscious (subliminal) stimulus presentations [Box 2]. However, HOT itself is agnostic as to whether this form of implicit metacognition can occur, despite the fact that HOT assumes that nonconscious cognition (but not necessarily nonconscious metacognition), precedes higher-order awareness.

Misconception 4: HOT is a variant of Global Workspace Theory

Like HOT, variants of Global Workspace Theory (GWT) [50-53] also suggest that first-order representations alone are insufficient for consciousness; additional downstream processes are needed. The two theories thus have this in common, but also differ in important ways.

For instance, in GWT the additional mechanisms required beyond first-order representations are functionally defined as serving the purpose of global broadcast, allowing conscious information to be shared among different (first-order) 'modules' for different modalities (e.g. perception, memory, language, motor control). As a result of such broadcasting, conscious signals will be strengthened and stabilized. In other words, consciousness is associated with important functional advantages, especially in behavioral and cognitive control.

HOTs, by contrast, do not typically propose that the relevant higher-order mechanisms have functional advantages [54]; but see [55]. This does not mean that higher-order representations are epiphenomena; it simply means that the purpose of the higher-order state may be *primarily* to give rise to conscious experience. Thus, the mere existence of the higher-order representations does not itself a guarantee behavioral output, much less superior behavioral performance, in a task.

This last point may be particularly relevant for understanding powerful forms of unconscious processing such as blindsight, in which patients deny having subjective experiences following damage to the visual cortex, even though they can guess the identity of visual stimuli well above chance level [56]. On GWT, this lack of conscious experience must mean that the relevant signal is not globally available to all major modules of the brain; somehow, some local pathway must have made possible the successful guessing and stimulus identification. According to HOT, on the other hand, a mere difference in the higher-order state determines whether a nonconscious perceptual process is consciously experienced or not, independent of the nature of the first-order process itself, or even in the absence of a first order state in some views [17, 19, 85]. Therefore, a percept can reach a relatively stable global state and remain nonconscious, which accounts for potentially powerful forms of unconscious perception

As such, HOT can be considered as a middle position between GWT and local theories, such as first-order theories [57-60]. GWT posits that consciousness is tightly associated with high-level global cognition and attention. First-order local theories suggest there are no such required links at all between consciousness and high-level cognition and attention. HOT, by contrast, posits

that not all high level global cognitive and attentional processes benefit functionally from consciousness, even though all conscious experiences involve mechanisms in regions traditionally thought to be important for high level cognition and attention. Just because the same regions are involved does not mean that all of the same cognitive functions are engaged in different conditions. Consciousness must, in the end, be measured—neural correlates of consciousness, though suggestive, do not demonstrate a causal relation between phenomenal experience and brain activity and do not allow the reverse engineering of consciousness simply based on areas of activation.

One recent review global broadcast and higher-order monitoring were cast as two orthogonal dimensions of processing [61]. The authors agreed that both kinds of processes, and especially their conjunction, are typically involved in everyday conscious experiences; if one were to speculate about the way in which conscious artificial intelligence could be built, both dimensions would be relevant. However, according to HOT, for subjective experience to occur, what is absolutely essential is higher-order awareness. Global broadcasting tends to go hand-in-hand with higher-order awareness in typical cases, but not always. For example, in peripheral vision one has the sense of seeing the details, but the details are typically not globally accessible as such. HOT accounts for this by proposing that global broadcasts do not necessarily happen for all subjective experiences.

The contrary may also occur: some information may be globally broadcasted but the phenomenology of normal conscious perception may be missing. For example, when holding visual information in working memory, the content is globally accessible, but one does not typically confuse this with normal conscious seeing. Imagine one is maintaining images of a few abstract shapes in order to respond a few seconds later about what the shapes are. During the delay, active vivid imagery may not be necessary. In fact, vividness of imagery varies between individuals [62], but even those who don't experience anything like normal seeing during the delay tend to have no problem holding the information. Given that the information is globally accessible in both cases (seeing vs holding information online in memory), what accounts for the difference in phenomenology? It is plausible that the difference is in higher-order processes.

Misconception 5: HOT is about reports & access, not experiences per se

Some proponents of GWT posit that global broadcast is a mechanism for conscious access, implying that access rather than phenomenology *per se* is all one can experimentally study [63]; see [53] for a possible exception. However, to the extent that one can distinguish between access and subjective phenomenology at all, proponents of HOT typically argue that subjective experiences are not simply about access to information.

Related to the case of peripheral vision mentioned above, arguments have been made against GWT based on the basis of experiments in which one reports phenomenally experiencing a large array of stimuli but is unable to report the details - because the mechanism that allows access is occupied with doing other

tasks or overflowed with other content [64]. These cases have been taken to support first-order views [65]. But HOT readily provides an account of these findings too [Box 3], precisely because HOT is about experience not reports per se. That is, participants' impression of rich and unreportable experience may be reflected by higher-order processes.

Another line of criticism for HOT concerning report is that much of the empirical evidence supporting HOT comes from tasks requiring subjects to report about the stimuli, and the act of reporting could be a confound [66]. They thus suggest the use of so-called No Report paradigms in which subjects are not asked to make reports about the relevant stimuli [67]. Advocacy of such paradigms, in our opinion, has unfortunately generated undue excitement. It is, of course, important to control for experimental confounds. Indeed, task demand (i.e. the need for subjects to report about the relevant stimuli) has long been controlled for and addressed in the literature [68-70]. It is true that in some studies when subjects were required to direct their attention away, basic neuroimaging analysis failed to find a difference in higher-order activity tracking conscious perception [66]. But it is a mistake to generalize from these null findings, because studies using more sensitive methods showed clear involvement of higher-order mechanisms, even when the subjects were not required to report about the stimuli [71-73].

Misconception 6: HOT is implausible because conscious experiences are more tightly associated with first-order (sensory) activity

One reason that critics mistakenly think that HOT is about report and access is that, pre-theoretically, conscious experiences seem to be associated with first-order perception [1, 74]. After all, activity in the visual cortex is typically understood to reflect the content in the conscious experiences concerned, even in some versions of HOT [18]. For example, if V4 processes information about color, and MT processes information about motion, then one might expect that these are the areas that subserve conscious experience of these properties [57]. Therefore, anything beyond first-order content areas may just reflect post-perceptual downstream processing that allows cognitive access and reporting. One reason that this may be an appealing idea is that, if correct, it would greatly simplify the task of understanding the neural mechanisms of consciousness since sensory areas are better understood at this time than higher-cognitive circuits [75].

However, while it is true that the various visual areas seem to code relevant content, it is unclear that their link to consciousness is straightforward. It is known that subliminal or unconscious stimulation also engages similar sensory circuits as consciously experienced stimuli [76]. Also, ongoing spontaneous activity in the very same neurons giving rise to normal perception is typical unconscious [77]. One possibility is that such activity has to reach a certain threshold [59], or that it has to achieve a right sort of dynamic profile [58, 77], such as involving feedback to V1. But patients with damage to V1 can still have visual experiences [18, 78]. So at least from the outset, an alternative seems just

as plausible: that (extrastriate) visual activity has to be re-represented by higher-order processes to become conscious, as suggested by HOT.

Misconception 7: HOT suggests that consciousness is *in* the Prefrontal Cortex

HOT is not in about prefrontal cortex (PFC) *per se*. It is a theory about fundamental cognitive requirements of consciousness. But combined with advances in cognitive neuroscience implicating PFC in higher cognition [82-84], the hypothesis is that the prefrontal cortex plausibly plays a key role in consciousness because it supports essential cognitive mechanisms that underlie higher-order representations.

Related to Misconception 1 above, different HOTs do not agree on the degree to which the PFC contributes to consciousness. Lau and Rosenthal proposed that in HOTT a first-order visual state being conscious consists in prefrontal cortex activity of the right sort [35]. A related view proposed by Lau is that higher-order prefrontal representations index the relevant first-order content in sensory areas to allow that content to enter consciousness [18]. In this sense, consciousness is “jointly determined” by concurrent activity in the higher-order areas and areas that maintain the lower-order representations. Alternatively, in Brown’s HOROR theory, the higher-order state in prefrontal cortex itself is phenomenally conscious [17, 19, 85]. Which of these, if any, might be correct is an open question.

Despite these differences between theories, virtually no current proponents of HOT assume that one is ‘more conscious’ whenever there is more PFC activity. Unlike early sensory neural representations, PFC neurons do not straightforwardly signal constant stimulus features - neuronal coding may be complex and involve local and distributed ensembles within and between areas of PFC [75]. As such, observations that activity in a given PFC region, as measured with neuroimaging, may decrease while one is dreaming [86], or in a psychedelic state [87], do not pose a challenge to HOT -- such a simple relationship at this coarse-grained level is not generally presumed to be the case by HOT theorists.

In evaluating current data, it is important to note that ‘prefrontal cortex’ is a generic conception that subsumes a number of different sub-regions and functional specializations [88-96]. Unfortunately, this diversity has not always been kept in mind when discussing the involvement of prefrontal in HOT or other theories about consciousness. Some have even argued against constitutive involvement of prefrontal cortex in *any* form of consciousness, using findings from cases with partial lesions of some areas [97]. However, the anatomical details reported were at times grossly inaccurate (see [71]). In Box 4 we suggest how the heterogeneity of PFC, and the rich diversity of its inputs, can be leveraged to achieve a broader, more nuanced view of higher-order representations in consciousness.

Coda: Implications for Everyday Experiences

The above shows that HOT is not just a philosophical conjecture. It is a reasonable account that is empirically testable. And compared to existing theories, it has several notable advantages in accounting for the conscious experiences that permeate our everyday lives.

According to local recurrency [57-58], consciousness is entirely detached from attention and cognitive functions. We question the empirical plausibility of this idea, but even if it turns out to be true, it is unclear if this particular notion of consciousness - that of pure phenomenology disconnected *entirely* from cognitive mechanisms - is, on its own, useful for understanding complex, everyday experiences. One example is our experiences of emotions, such as fear, anger, or empathy, or when we are moved by music or art, or simply enjoy a tasty meal. Another involves semantic and conceptual memory contributions to conscious perceptions (as when we use prior knowledge in perception) or recollect personal, episodes of the days of our lives.

Phenomenology is obviously important, and HOT provides a conception that connects it with specific self-monitoring mechanisms, such as those enabled by the human prefrontal cortex. Although GWS also calls upon prefrontal mechanisms, in this theory 'consciousness' essentially amounts to the effectiveness or depth of information processing. This is an important aspect of cognition in general that applies independent of consciousness. However, we question whether it is needed, in addition to the well-studied mechanisms of attention, perception, and wakefulness, to describe the available data about consciousness itself.

Finally, we believe that the notion of consciousness characterized by HOT allows for a nuanced intermediate position that may have particularly meaningful utility in helping us to understand richness, not only of routine everyday experiences, but also of the experiences of those who suffer from mental disorders. For decades, the subjective experiences of patients have been marginalized in establishment of the evaluation of the effectiveness of behavioral, cognitive, and pharmaceutical treatments [98]. In medical models, for example, the goal has been to find biological correlates that, if changed, especially by drugs, will solve the problem [37, 79]. But so unsuccessful have such efforts been to develop novel agents through studies of animal that the pharmaceutical industry has significantly scaled back research in this area [99].

Such considerations underlie recent debates about whether targeting the amygdala circuitry will help to alleviate symptoms of fear and anxiety in disorders such as phobias [79-81]. Of course, in a sense, the amygdala is important and relevant for fear. But ample evidence also suggests it is unlikely to be the full story [19, 37, 98]. It is known that threats can elicit amygdala activity and trigger physiological responses nonconsciously [100, 101]. Further, direct electrical stimulation of the amygdala reliably elicits physiological responses, but only rarely subjective experiences [102]; even when subjective experiences are elicited, it is unclear that these arise from the amygdala itself, as opposed to resulting from activity spreading to downstream higher-order processes [37]. Also, patients with amygdala lesions continue to consciously report emotional experiences [103], including fear [104]. These may be some of the reasons why

anti-anxiety medications targeting the amygdala in animals have not been as successful as once hoped in changing human subjective experiences.

HOT provides a principled account to address what may be the missing: to effectively treat problems related to fear and anxiety, the higher-order mechanisms of consciousness involving various areas of PFC, and the lower-order mechanisms involving subcortical defensive circuitry, may both have to be treated. But to do so effectively, each may have to be targeted separately, otherwise either can reinstate the other.

The science of consciousness may thus offer insights into clinical problems. But it is also the case that evaluation the utility of consciousness theories in the clinic may offer an underused source of insights about the nature of consciousness at both the psychological and neural levels.

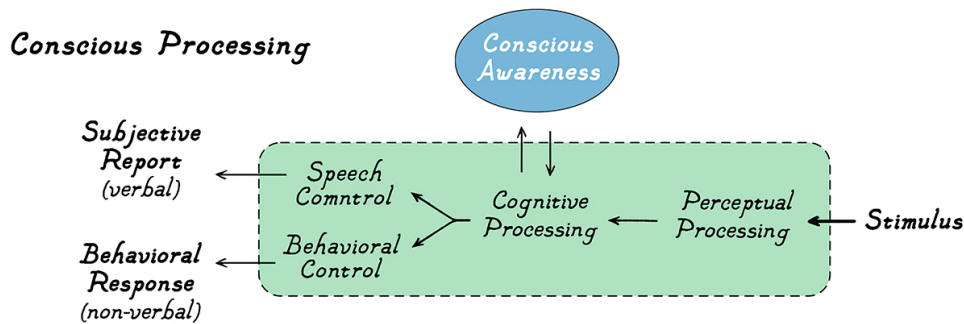
BOXES

Box 1: Phenomenal Consciousness in Animals

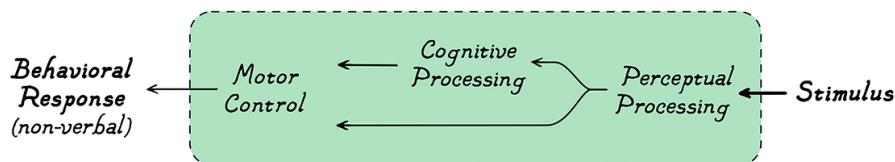
In human research, the standard practice for testing whether a process is conscious is to contrast conditions involving conscious vs. nonconscious stimulus presentations [105, 106]. However, critics of non-conscious processing try to find possible ways that the stimuli might have slipped through and had some effect on consciousness in nonconscious conditions. Researchers are constantly refining methods in the effort to evaluate conscious vs. non-conscious processing (e.g. [105, 107-113]). The result is that the criteria for demonstrating nonconscious processing become more and more stringent.

Given this, the way to demonstrate consciousness in animals should likewise be to compare conditions under which behavior is reasonably explained as being dependent on consciousness, and that are explained less convincingly in terms of non-conscious process. But in animal research, such contrasting of conditions is not standard fare because it is very difficult to obtain evidence that rules out non-conscious explanations. Most experiments end up being about amassing evidence consistent with support the intuition that consciousness was involved claiming that if animals behave similar to humans in similar kinds of situations, the animals must be similarly conscious [105].

When humans are conscious of a stimulus, they can report on their knowledge either verbally (by saying “I see an apple”) or non-verbally (by manually picking the apple from a bowl that also has a pear and orange in it) (**Box 1 Figure**). But when responding to a subliminal (unconsciously processed) stimulus, humans can only respond non-verbally to indicate the presence of the stimulus. The fact that animals can only respond non-verbally means there is no second response to help separate conscious from non-conscious processes. Lacking verbal report to drive a wedge between conscious and nonconscious processes, it is more difficult to parse when a behavioral response reflects a conscious state as opposed to being controlled non-consciously in animals (or in preverbal humans or robots) [37].



Non-conscious Processing
 (subliminal or unattended stimuli and blindsight)



Box 1 Figure: Conscious processing can be responded to verbally or non-verbally but non-conscious processing can only be responded to non-verbally.

Given this, one strategy may be to collect multiple non-verbal measures from the non-human agent, e.g. simple stimuli discrimination response, detection response, confidence, etc, and analyze them in multiple ways (sensitivity vs bias, both for the simple responses as well as for confidence). From there, we can compare how these various measures overall match those in humans. For example, in the case of blindsight in humans, following lesion to the primary visual cortex it is known that the impact is relatively greater on detection sensitivity than on two-alternative forced-choice discrimination sensitivity [114], and that it may also affect detection bias and confidence in specific ways [115, 116]. If following lesion to the same brain region, an animal shows the same profile of behavioral changes, evidence that the phenomena may be similar would be obtained [117, 118]. That is, it would look as if following lesion, the animal has become 'less' visually conscious. However, we are not advocating that this strategy can be used to unequivocally establish if an animal is conscious in the first place. Without verbal confirmation, we are unsure if we can firmly establish - or deny - consciousness based on behavioral assessment alone.

We therefore remain agnostic about animal consciousness—that is, we do not rule it out from the start, but feel that there are significant challenges to demonstrating it. Still, the distinction proposed by Tulving offers a way of conceiving animal consciousness that does not contradict that possibility that human experience is distinct. In other words, it is possible that some animals could have a noetic (semantic) form of phenomenal consciousness (consciousness of perceptual or memory representations) without having an auto-noetic one (awareness of themselves as having the phenomenal experience) [36, 37]. This idea overlaps with Shea and colleagues' notion of uniquely human supra-personal metacognition [45]. At the same time, the possibility that noetic consciousness is made possible by cognitive abilities shared with some other animals, especially non-human primates, allows speculation about perceptual awareness in these animals [119]. But the key problem still remains—that a gap exists between what the evidence can unequivocally show and what may exist in the mind of the animal. Given cellular and structural [89, 92, 96] and cognitive [38, 45, 120, 121] differences between human and non-human primate PFC, even if prefrontal regions that are activated when humans report conscious experience are also active when animals perform similar tasks, this does not necessarily mean that the animals are having the kinds of experiences the human is having. Similar considerations also apply to issues about scientifically determining whether robots or young infants may be conscious.

Box 2: Unconscious working memory and metacognition

It has been claimed that subliminal stimuli can drive or influence working memory and explicit metacognition (assessed via confidence ratings). In the main text we clarified that these claims, even if true, pose no threat to HOT per se. Just because consciousness requires some specific metacognitive-like mechanisms does not mean that metacognition always entails conscious awareness of the relevant stimuli. In other words, while metacognition and consciousness share mechanisms they are not the same.

Independently, it is also important to point out that claims of explicit metacognition with unconscious stimuli (as opposed to implicit or unconscious metacognition) are still not totally convincing – these hinge on how unconsciousness is defined. As has been pointed out by decades of psychophysics, simply asking people to label trials within a task condition as “consciously seen” vs “unconscious” is not always sufficient [108]. There is a serious concern about an arbitrary criterion's being set to fit trials into the two forced labels.

Against historical [122] and contemporary [123, 124] skepticism by some about the existence of unconscious processing at all, more rigorous standards have been proposed for the demonstration that certain perceptual information is genuinely unconsciously processed [107, 108]. For instance, to demonstrate such unconsciousness one may need to show that the relevant stimuli cannot be discriminated above chance at all, or that it can be discriminated better than

chance under forced-choice settings, but it is subjectively indistinguishable from a 'blank' stimulus containing no information. Current efforts to demonstrate 'unconscious' working memory [125] and unconscious metacognition [107] have faced difficulties and have not yet passed these more rigorous standards.

In part, the difficulties in studying unconscious processing are due to the fact that unconscious stimuli are by their very nature weakly encoded [126]. Thus, as with the discussion of animal consciousness, it is important to keep in mind that there may be a gap between what we can measure scientifically and what may actually exist in the non-conscious mental terrain in humans. It is also important to distinguish the question of whether working memory can operate on unconscious information as opposed to whether it create unconscious temporary representations [125].

BOX 3 - Subjective inflation in the unattended periphery

Our impression of experiencing details needs not be fully veridical. We have the subjective impression of seeing vivid color in the periphery despite our relatively poor early sensory processing. It is thus argued that our subjective impression may be at times inflated beyond the actual representational capacity [127], due to mis-representation at the higher-order level.

Against this view, Ned Block [128] has argued that color perception is actually not weak in the periphery - if we enlarge the stimuli for the periphery to compensate for the larger receptive field size there [18]. But we are not sure how this point is relevant because in real life rather than in calibrated psychophysics experiments, stimuli don't automatically enlarge themselves upon entering our periphery. What matters is that for the same stimuli, color processing is relatively weak in the periphery, and yet we see the world as somewhat rather uniform.

Besides color [129], inflation for other attributes are also well documented - subjects reported higher false alarm rates for detection in the periphery, or in unattended locations. Critics argue these results only reflected a bias in responding strategies, and it would beg the question to assume that such biases necessarily reflect inflated subjective experiences [128]. However, it is unclear if the charge of begging the question is fair, because proponents of HOT do provide independent reasons for believing that the inflation results reflect experiences rather than just responses. For instance, if inflation reflects biased strategy it would be difficult to explain why the results persisted even upon training with extensive feedback. The hypothesis that subjective experience itself inflates seems to account for the data as well as anecdotal experience better.

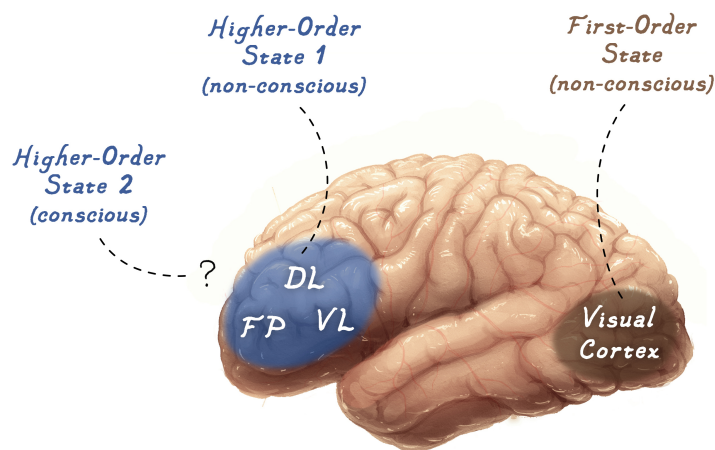
BOX 4 - Prefrontal Cortex and Higher-Order Experience

'Prefrontal cortex' subsumes a number of different sub-regions (some located laterally and some medially) with different cytoarchitectonic properties (granular, dysgranular, agranular), different patterns of connectivity with each other and with sensory, memory, and conceptual processing regions, and different phylogenetic histories--some present in all mammals, some mainly in primates (dorsolateral) and some particularly well developed in humans (frontal pole), and some uniquely human (frontal pole) [88-95]. This complexity, though at

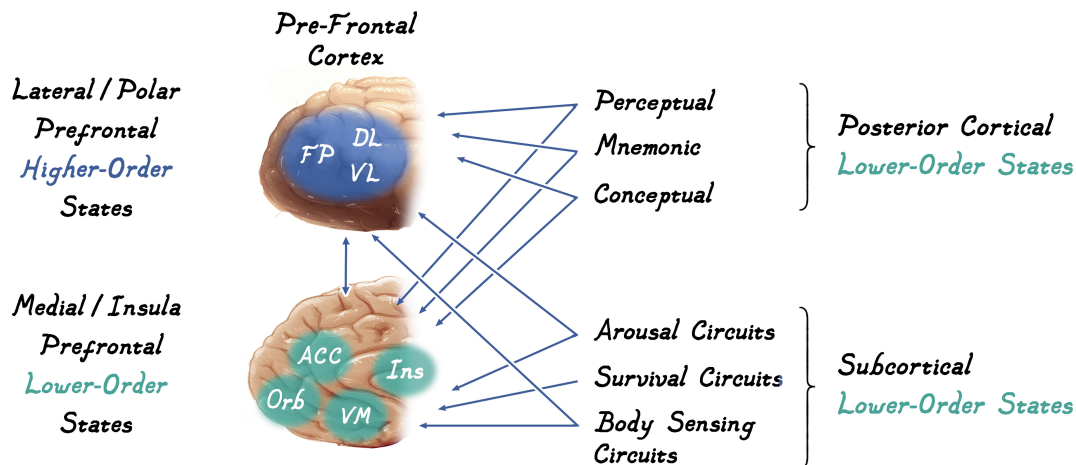
first blush seemingly daunting, might in fact offer opportunities for a more nuanced view of how PFC contributes to experience.

The most detailed proposal to date about the role of specific areas of prefrontal cortex in higher-order representations was presented by [5]. Key areas in their model were the dorsolateral and polar regions of prefrontal cortex. While their review was primarily about visual perception, adding the ventral lateral PFC allows their model to generalize to other external senses (**Box 4 Figure A**). Consistent with Rosenthal's HOTT, they suggested that higher-order representations in this network are not themselves consciously experienced—conscious experience of the higher-order representation requires an additional level of higher-order representation. Though they did not identify the source of the additional higher-order representation, the highly conceptual nature of polar regions makes it a candidate worth considering.

As it stands, the model may not extend well to other, more complex kinds of experiences (for example, of memories, emotions, and self). LeDoux [36] proposes nature of the relevant lower-order representations might need to be reconceived to account for these. For example, medial and insula prefrontal areas receive inputs, not only from perceptual, but also from conceptual, memory, and subcortical circuits, and interconnect with the lateral-polar prefrontal higher-order network. Medial and insula areas might therefore construct complex, nevertheless lower-order, representations used by the lateral-polar areas in the construction of non-conscious higher-order representations and phenomenally conscious experiences. That prefrontal areas could be a source of lower-order representations in higher-order conscious experiences emphasizes that a give prefrontal area, while higher-order anatomically and cognitively, is not necessarily higher-order in the sense meant in HOT (**Box 4 Figure B**).



Box 4 Figure 1: Cortical areas underlying conscious experience in First-Order and Higher-Order theories.



Box 4 Figure 2: Lower-order inputs to the higher-order network involving lateral and polar prefrontal areas.

1. Block, N. (1995) On a confusion about a function of consciousness. *Behavioral and Brain Sciences* 18, 227-247.
2. Nagel, T. (1974) What is it like to be a bat? *Philosophical Review* 83, 4435-4450.
3. Rosenthal, D.M. (2005) *Consciousness and mind*, Oxford University Press.
4. Rosenthal, D. (2004) Varieties of higher-order theory. In *Higher-Order Theories of Consciousness: An Anthology* (Gennaro, R.J. ed), John Benjamins.
5. Lau, H. and Rosenthal, D. (2011) Empirical support for higher-order theories of conscious awareness. *Trends Cogn Sci* 15 (8), 365-73.
6. Rosenthal, D. and Weisberg, J. (2008) Higher-order theories of consciousness. *Scholarpedia* 3 (5), 4407.
7. Armstrong, D.M. (1968) *A Materialist Theory of the Mind*, Routledge.
8. Lycan, W.G. (2004) The superiority of Hop to HOT. In *Higher-Order Theories of Consciousness: An Anthology* (Gennaro, R.J. ed), p. 93-114, John Benjamins.
9. Lycan, W.G. (1996) *Consciousness and Experience*, MIT Press.
10. Sauret, W. and Lycan, W.G. (2014) Attention and internal monitoring: A Farewell to HOP. *Analysis* 74 (3), 363-370.
11. Carruthers, P. (2000) *Phenomenal consciousness: A naturalistic theory*, Cambridge University Press.

12. Carruthers, P. (2018) Comparative psychology without consciousness. *Conscious Cogn* 63, 47-60.
13. Metzinger, T. (2003) *Being no one*, MIT Press.
14. Cleeremans, A. (2011) The Radical Plasticity Thesis: How the Brain Learns to be Conscious. *Front Psychol* 2, 86.
15. Kriegel, U. (2009) *Subjective Consciousness: A self-representational theory*, Oxford University Press.
16. Gennaro, R.J. (2011) *The Consciousness Paradox*, MIT Press.
17. Brown, R. (2015) The HOROR theory of phenomenal consciousness. *Philosophical Studies* 172 (7), 1783-1794.
18. Lau, H. and Brown, R. (2019) The Emperor's New Phenomenology? The Empirical Case for Conscious Experience without First-Order Representations. In *Blockheads! Essays on Ned Block's philosophy of mind and consciousness* (Pautz, A. and Stoljar, D. eds), MIT Press.
19. LeDoux, J.E. and Brown, R. (2017) A higher-order theory of emotional consciousness. *Proc Natl Acad Sci U S A* 114 (10), E2016-E2025.
20. Gottlieb, J. (2015) Presentational Character and Higher Order Thoughts. *Journal of Consciousness Studies* 22 (7-8), 103-123.
21. Farrell, J. (2018) Higher-order theories of consciousness and what-it-is-likeness. *Philos Stud* 175 (11), 2743-2761.
22. Picciuto, V. (2017) Keeping it Real: Intentional Inexistents, Fineness-of-Grain, and the Dilemma for Extrinsic Higher-Order Representational Theories. *Pacific Philosophical Quarterly* 98 (4), 555-575.
23. Sebastián, M.Á. (2018) Drop it like it's HOT: a vicious regress for higher-order thought theories. *Philosophical Studies*, 1-10.
24. Adams, F. and Shreve, C. (2016) What Can Synesthesia Teach Us About Higher Order Theories of Consciousness? Symposium: Theoretical and Applied Inquiries in Philosophy and Social Sciences 3 (3), 251-257.
25. Brinck, I. and Kirkeby-Hinrup, A. (2017) Change Blindness in Higher-Order Thought: Misrepresentation or Good Enough? *Journal of Consciousness Studies* 24 (5-6), 50-73.
26. Block, N. (2011) The higher-order approach to consciousness is defunct. *Analysis* 71 (3), 419-431.
27. Malach, R. (2011) Conscious perception and the frontal lobes: comment on Lau and Rosenthal. *Trends Cogn Sci* 15 (11), 507; author reply 508-9.
28. Coleman, S. (2019) The Merits of Higher-Order Thought Theories. *Trans/Form/Ação*, (in press).
29. Rosenthal, D. (2019) Misrepresentation and Mental Appearance. *Trans/Form/Ação*, (in press).
30. Dretske, F. (1995) *Naturalizing the Mind*, MIT Press.
31. Seager, W.E. (2004) A cold look at HOT theory. In *Higher-Order Theories of Consciousness: An Anthology* (Gennaro, R.J. ed), John Benjamins.
32. Jack, A.I. and Shallice, T. (2001) Introspective physicalism as an approach to the science of consciousness. *Cognition* 79 (1-2), 161-96.

33. Nelson, T.O. and Narens, L. (1994) Why investigate metacognition? In *Metacognition: Knowing about knowing* (Metcalf, J. and Shimamura, A.P. eds), pp. 1-25, The MIT Press.
34. Rosenthal, D. (2002) Explaining Consciousness. In *Philosophy of Mind: Classical and contemporary readings* (Chalmers, D.J. ed), pp. 406-17, Oxford University Press.
35. Lau, H. and Rosenthal, D. (2011) The higher-order view does not require consciously self-directed introspection: response to Malach. *Trends Cogn Sci* 15 (11), 508-509.
36. LeDoux, J. (2019) *The Deep History of Ourselves: How ancient microbes became conscious brains*, Viking.
37. LeDoux, J.E. (2015) *Anxious: Using the brain to understand and treat fear and anxiety*, Viking.
38. Tulving, E. (2005) Episodic memory and autoevidence: Uniquely human? In *The Missing Link in Cognition* (Terrace, H.S. and Metcalfe, J. eds), pp. 4-56, Oxford University Press.
39. Edelman, G.M. and Tononi, G. (2000) *A Universe of Consciousness: How matter becomes imagination*, Basic Books.
40. Damasio, A.R. (1999) *The feeling of what happens: Body and emotion in the making of consciousness*, Harcourt Brace.
41. Panksepp, J. (1998) *Affective Neuroscience*, Oxford U. Press.
42. Lewis, M. (2013) *The rise of consciousness and the development of emotional life*, The Guilford Press.
43. Maniscalco, B. and Lau, H. (2012) A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. *Conscious Cogn* 21 (1), 422-30.
44. Fleming, S.M. and Lau, H.C. (2014) How to measure metacognition. *Front Hum Neurosci* 8, 443.
45. Shea, N. et al. (2014) Supra-personal cognitive control and metacognition. *Trends Cogn Sci* 18 (4), 186-93.
46. Lau, H.C. and Passingham, R.E. (2007) Unconscious activation of the cognitive control system in the human prefrontal cortex. *J Neurosci* 27 (21), 5805-11.
47. Rounis, E. et al. (2010) Theta-burst transcranial magnetic stimulation to the prefrontal cortex impairs metacognitive visual awareness. *Cogn Neurosci* 1 (3), 165-75.
48. Ruby, E. et al. (2018) On a 'failed' attempt to manipulate visual metacognition with transcranial magnetic stimulation to prefrontal cortex. *Conscious Cogn* 62, 34-41.
49. Shekhar, M. and Rahnev, D. (2018) Distinguishing the roles of dorsolateral and anterior PFC in visual metacognition. *J Neurosci*.
50. Baars, B.J. and Franklin, S. (2007) An architectural model of conscious and unconscious brain functions: Global Workspace Theory and IDA. *Neural Netw* 20 (9), 955-61.
51. Dehaene, S. (2014) *Consciousness and the Brain: Deciphering how the brain codes our thoughts*, Penguin Books.

52. Naccache, L. and Dehaene, S. (2007) Reportability and illusions of phenomenality in the light of the global neuronal workspace model. *Behav Brain Sci* 30 (5-6), 518-20.
53. Naccache, L. (2018) Why and how access consciousness can account for phenomenal consciousness. *Philos Trans R Soc Lond B Biol Sci* 373 (1755).
54. Rosenthal, D.M. (2008) Consciousness and its function. *Neuropsychologia* 46 (3), 829-40.
55. LeDoux, J. and Daw, N.D. (2018) Surviving threats: neural circuit and computational implications of a new taxonomy of defensive behaviour. *Nat Rev Neurosci* 19 (5), 269-282.
56. Weiskrantz, L. (1997) *Consciousness lost and found: A neuropsychological exploration*, Oxford University Press.
57. Block, N. (2007) Consciousness, accessibility, and the mesh between psychology and neuroscience. *Behav Brain Sci* 30 (5-6), 481-99; discussion 499-548.
58. Lamme, V.A. (2006) Towards a true neural stance on consciousness. *Trends Cogn Sci* 10 (11), 494-501.
59. Zeki, S. (2001) Localization and globalization in conscious vision. *Annu Rev Neurosci* 24, 57-86.
60. Fisch, L. et al. (2009) Neural "ignition": enhanced activation linked to perceptual awareness in human ventral stream visual cortex. *Neuron* 64 (4), 562-74.
61. Dehaene, S. et al. (2017) What is consciousness, and could machines have it? *Science* 358 (6362), 486-492.
62. Pearson, J. et al. (2015) *Mental Imagery: Functional Mechanisms and Clinical Applications*. *Trends Cogn Sci* 19 (10), 590-602.
63. Dehaene, S. et al. (2006) Conscious, preconscious, and subliminal processing: a testable taxonomy. *Trends Cogn Sci* 10 (5), 204-11.
64. Sperling, G. (1960) The information available in brief visual presentations. *Psychological Monographs* 74 (11), 1-29.
65. Block, N. (2008) Consciousness and cognitive access. *Proceedings of the Aristotelian Society* 108 (1pt3), 289-317.
66. Tsuchiya, N. et al. (2015) No-Report Paradigms: Extracting the True Neural Correlates of Consciousness. *Trends Cogn Sci* 19 (12), 757-70.
67. Block, N. (2014) Rich conscious perception outside focal attention. *Trends Cogn Sci* 18 (9), 445-7.
68. Kouider, S. et al. (2007) Cerebral bases of subliminal and supraliminal priming during reading. *Cereb Cortex* 17 (9), 2019-29.
69. Lumer, E.D. and Rees, G. (1999) Covariation of activity in visual and prefrontal cortex associated with subjective visual perception. *Proc Natl Acad Sci U S A* 96 (4), 1669-73.
70. Tse, P.U. et al. (2005) Visibility, visual awareness, and visual masking of simple unattended targets are confined to areas in the occipital cortex beyond human V1/V2. *Proc Natl Acad Sci U S A* 102 (47), 17178-83.
71. Odegaard, B. et al. (2017) Should a Few Null Findings Falsify Prefrontal Theories of Conscious Perception? *J Neurosci* 37 (40), 9593-9602.

72. Vidal, J.R. et al. (2014) Intracranial spectral amplitude dynamics of perceptual suppression in fronto-insular, occipito-temporal, and primary visual cortex. *Front Psychol* 5, 1545.
73. Kapoor, V. et al. (2018) Parallel and functionally segregated processing of task phase and conscious content in the prefrontal cortex. *Communications Biology*.
74. Prinz, J.J. (2012) *The Conscious Brain: How Attention Engenders Experience*, Oxford University Press.
75. Knotts, J.D. et al. (2018) Neuroscience: The Key to Consciousness May Not Be under the Streetlight. *Curr Biol* 28 (13), R749-R752.
76. Macknik, S.L. and Livingstone, M.S. (1998) Neuronal correlates of visibility and invisibility in the primate visual system. *Nat Neurosci* 1 (2), 144-9.
77. Moutard, C. et al. (2015) Spontaneous Fluctuations and Non-linear Ignitions: Two Dynamic Faces of Cortical Recurrent Loops. *Neuron* 88 (1), 194-206.
78. Silvanto, J. (2015) Why is "blindsight" blind? A new perspective on primary visual cortex, recurrent activity and visual awareness. *Conscious Cogn* 32, 15-32.
79. LeDoux, J.E. and Pine, D.S. (2016) Using Neuroscience to Help Understand Fear and Anxiety: A Two-System Framework. *Am J Psychiatry* 173 (11), 1083-1093.
80. Fanselow, M.S. and Pennington, Z.T. (2018) A return to the psychiatric dark ages with a two-system framework for fear. *Behav Res Ther* 100, 24-29.
81. Pine, D.S. and LeDoux, J.E. (2017) Elevating the Role of Subjective Experience in the Clinic: Response to Fanselow and Pennington. *Am J Psychiatry* 174 (11), 1121-1122.
82. Miller, E.K. and Cohen, J.D. (2001) An integrative theory of prefrontal cortex function. *Annu Rev Neurosci* 24, 167-202.
83. Nee, D.E. and D'Esposito, M. (2018) The Representational Basis of Working Memory. *Curr Top Behav Neurosci* 37, 213-230.
84. D'Esposito, M. and Postle, B.R. (2015) The cognitive neuroscience of working memory. *Annu Rev Psychol* 66, 115-42.
85. Brown, R. (2014) Consciousness doesn't overflow cognition. *Front Psychol* 5, 1399.
86. Sebastián, M.Á. (2013) Not a HOT Dream. In *Consciousness Inside and Out: Phenomenology, Neuroscience, and the Nature of Experience*. Springer Studies in Brain and Mind.
87. Carhart-Harris, R.L. et al. (2014) The entropic brain: a theory of conscious states informed by neuroimaging research with psychedelic drugs. *Front Hum Neurosci* 8, 20.
88. Carlen, M. (2017) What constitutes the prefrontal cortex? *Science* 358 (6362), 478-482.
89. Teffer, K. and Semendeferi, K. (2012) Human prefrontal cortex: evolution, development, and pathology. *Prog Brain Res* 195, 191-218.
90. Fuster, J. (2008) *The Prefrontal Cortex*, 4th edn., Academic Press.
91. Barbas, H. and Pandya, D.N. (1989) Architecture and intrinsic connections of the prefrontal cortex in the rhesus monkey. *J Comp Neurol* 286 (3), 353-75.
92. Koechlin, E. et al. (2003) The architecture of cognitive control in the human prefrontal cortex. *Science* 302 (5648), 1181-5.

93. Passingham, R.E. and Wise, S.P. (2012) *The Neurobiology of the Prefrontal Cortex*, Oxford University Press.
94. Petrides, M. et al. (2012) The prefrontal cortex: comparative architectonic organization in the human and the macaque monkey brains. *Cortex* 48 (1), 46-57.
95. Yeterian, E.H. et al. (2012) The cortical connectivity of the prefrontal cortex in the monkey brain. *Cortex* 48 (1), 58-81.
96. Preuss, T.M. (2011) The human brain: rewired and running hot. *Ann N Y Acad Sci* 1225 Suppl 1, E182-91.
97. Koch, C. (2018) What Is Consciousness? *Nature* 557 (7704), S8-S12.
98. LeDoux, J. et al., *Know Thyself: Well-Being and Subjective Experience*, Cerebrum, The Dana Foundation, New York, 2018.
99. Miller, G. (2010) Is pharma running out of brainy ideas? *Science* 329 (5991), 502-4.
100. Tamietto, M. and de Gelder, B. (2010) Neural bases of the non-conscious perception of emotional signals. *Nat Rev Neurosci* 11 (10), 697-709.
101. Ohman, A. (2002) Automaticity and the amygdala: nonconscious responses to emotional faces. *Curr Dir Psychol Sci* 11, 62-66.
102. Inman, C.S. et al. (2018) Human amygdala stimulation effects on emotion physiology and emotional experience. *Neuropsychologia*.
103. Anderson, A.K. and Phelps, E.A. (2002) Is the human amygdala critical for the subjective experience of emotion? Evidence of intact dispositional affect in patients with amygdala lesions. *J Cogn Neurosci* 14 (5), 709-20.
104. Feinstein, J.S. et al. (2013) Fear and panic in humans with bilateral amygdala damage. *Nat Neurosci* 16 (3), 270-2.
105. Heyes, C. (2008) Beast machines? Questions of animal consciousness. In *Frontiers of Consciousness: Chichelle lectures* (Weiskrantz, L. and Davies, M. eds), pp. 259-274, Oxford University Press.
106. Heyes, C. (2016) Blackboxing: social learning strategies and cultural evolution. *Philos Trans R Soc Lond B Biol Sci* 371 (1693).
107. Peters, M.A. and Lau, H. (2015) Human observers have optimal introspective access to perceptual processes even for visually masked stimuli. *Elife* 4, e09651.
108. Macmillan, N.A. and Creelman, C.D. (2004) *Detection Theory: A user's guide*, 2nd edn., Lawrence Erlbaum Associates, Inc.
109. Frith, C. et al. (1999) The neural correlates of conscious experience: an experimental framework. *Trends in Cognitive Sciences* 3 (3), 105-114.
110. Maniscalco, B. and Lau, H. (2016) The signal processing architecture underlying subjective reports of sensory awareness. *Neurosci Conscious* 2016 (1).
111. Seth, A.K. et al. (2008) Measuring consciousness: relating behavioural and neurophysiological approaches. *Trends Cogn Sci* 12 (8), 314-21.
112. Dehaene, S. and Changeux, J.P. (2011) Experimental and theoretical approaches to conscious processing. *Neuron* 70 (2), 200-27.
113. Giles, N. et al. (2016) What Type of Awareness Does Binocular Rivalry Assess? *Trends Cogn Sci* 20 (10), 719-20.
114. Azzopardi, P. and Cowey, A. (1997) Is blindsight like normal, near-threshold vision? *Proc Natl Acad Sci U S A* 94 (25), 14190-4.

115. Ko, Y. and Lau, H. (2012) A detection theoretic explanation of blindsight suggests a link between conscious perception and metacognition. *Philos Trans R Soc Lond B Biol Sci* 367 (1594), 1401-11.
116. Persaud, N. et al. (2011) Awareness-related activity in prefrontal and parietal cortices in blindsight reflects more than superior visual performance. *Neuroimage* 58 (2), 605-11.
117. Yoshida, M. et al. (2008) Striate cortical lesions affect deliberate decision and control of saccade: implication for blindsight. *J Neurosci* 28 (42), 10517-30.
118. Schmid, M.C. et al. (2010) Blindsight depends on the lateral geniculate nucleus. *Nature* 466 (7304), 373-7.
119. Panagiotaropoulos, T.I. et al. (2012) Neuronal discharges and gamma oscillations explicitly reflect visual consciousness in the lateral prefrontal cortex. *Neuron* 74 (5), 924-35.
120. Penn, D.C. et al. (2008) Darwin's mistake: explaining the discontinuity between human and nonhuman minds. *Behav Brain Sci* 31 (2), 109-30; discussion 130-178.
121. Tomasello, M. and Rakoczy, H. (2003) What Makes Human Cognition Unique? From Individual to Shared to Collective Intentionality. *Mind & Language* 18 (2), 121-147.
122. Brentano, F. (1874/1924) *Psychologie vom Empirischen Standpunkt*, Felix Meiner.
123. Vadillo, M.A. et al. (2016) Underpowered samples, false negatives, and unconscious learning. *Psychon Bull Rev* 23 (1), 87-102.
124. Phillips, I. (2018) Unconscious perception reconsidered. *Analytic Philosophy* 59 (4).
125. Persuh, M. et al. (2018) Working Memory and Consciousness: The Current State of Play. *Front Hum Neurosci* 12, 78.
126. Seth, A.K. (in press) From unconscious inference to the Beholder's share: Predictive perception and human experience. *European Review*.
127. Odegaard, B. et al. (2018) Inflation versus filling-in: why we feel we see more than we actually do in peripheral vision. *Philos Trans R Soc Lond B Biol Sci* 373 (1755).
128. Block, N. (2019) Empirical science meets higher order views of consciousness: reply to Hakwan Lau and Richard Brown. In *Blockheads! Essays on Ned Block's philosophy of mind and consciousness* (Pautz, A. and Stoljar, D. eds), MIT Press.
129. Li, M.K. et al. (2018) An investigation of detection biases in the unattended periphery during simulated driving. *Atten Percept Psychophys*.