Trends in Cognitive Sciences

CellPress

Opinion

What does decoding from the PFC reveal about consciousness?

Ned Block [1,*]

Disputes between rival theories of consciousness have often centered on whether perceptual contents can be decoded from the prefrontal cortex (PFC). Failures to decode from the PFC are taken to challenge 'cognitive' theories of consciousness such as the global workspace theory and higher-order monitoring theories, and decoding successes have been taken to confirm these theories. However, PFC decoding shows both too much and too little. Too much because cognitive theories of consciousness do not need PFC rerepresentation of perceptual contents since pointers to perceptual representations suffice. Too little because there is evidence that PFC decoding of perceptual content reflects postperceptual cognitive representation, such as thoughts that have those perceptual contents rather than conscious percepts.

Highlights

A number of recent articles have shown intracranial decoding in prefrontal cortex (PFC) for high-level perceptual features but not low-level features. Furthermore, sustained perception results only in momentary PFC decoding. These results have been taken to both confirm and challenge aspects of 'prefrontalist' theories of consciousness.

New 'inattentional blindness' results have been used to argue for PFC representation of perceptual contents.

Pointer versions of prefrontalist theories of consciousness provide a way of accommodating the decoding results but have been thought to be challenged by inattentional blindness results.

Pointer theories run into problems with intracranial stimulation.

Bifurcation dynamics yields a new no-report paradigm.

The global playground provides a better account of pure access than the global workspace.

PFC decoding may depend on postperceptual cognition.

Decoding from the PFC

Recent work on consciousness has emphasized intracranial **decoding** (see Glossary) of perceptual contents in the PFC. In binocular rivalry, subjects are shown different stimuli in each eye. If the stimuli 'conflict' ([1], pp. 191–192) subjects will see first one stimulus, then the other, alternating on and on. Researchers have decoded conscious PFC content in monkeys intracranially (i.e., using electrode arrays penetrating the cortex) during binocular rivalry of perception of up/down or right/left moving grids without any task (using eye movements as a substitute for report). This has been taken to support **global workspace** and **higher-order theories of consciousness** [2–4], both of which hold that consciousness requires PFC representation. Confirming the relevance of the decoding to consciousness, decoders that work for binocular rivalry also work for 'replay' in which subjects are shown first grids moving one way, then grids moving the other way, in a manner that is supposed to replicate the experience of binocular rivalry.

Furthermore, one research group demonstrated that perceptual contents can be decoded from the PFC even when stimuli are presented quite rapidly in succession, 10 per second [5]. While the authors acknowledge that the stimuli may have been perceived unconsciously, they argue that this finding provides some support for cognitive theories of consciousness since each stimulus masks the previous one, discouraging **postperceptual processing**, which, as we will see later, provides an alternative explanation for PFC decoding. However, another research group [6] found no difference between conscious and unconscious processing in the PFC, concluding (p. 295), 'our results challenge the pre-frontal theories in that prefrontal connectivity was not modulated by conscious awareness.' Again, the success or failure of PFC decoding is taken to be crucial.

Templeton World Charity Foundation has funded several 'adversarial collaborations' in which advocates of rival theories of consciousness have devised experiments that could challenge each of the rival theories. The first results [7] found PFC decoding for faces and letters but not

[1]New York University, 5 Washington Place, New York, NY 10003, USA

*Correspondence: ned.block@nyu.edu (N. Block).

for orientations of faces; the former was classified as a win for global workspace theories, and the latter was classified as a challenge for them [7] (even though the face orientations were never task-relevant). As a news report in *Science* put it, 'When it came to decoding different categories of objects, the data provided strong support for global neuronal workspace theory (GNWT). But when it came to decoding the orientation of faces, IIT was the better fit' [5]. (IIT is **Integrated Information Theory**, an approach that emphasizes posterior visual areas.)

Given the fact that both conscious and unconscious contents can be decoded from posterior visual areas even in anesthetized animals [8,9], decoding from the PFC has seemed pivotal to the fortunes of cognitive theories of perceptual consciousness even when it conflicts with reports (Box 1). This article argues that the emphasis on decoding from the PFC is a mistake even for prefrontalist theories. I will argue that there are advantages to prefrontalist theories that do not require decoding from the PFC and further that decoding from the PFC does not show what many of its adherents have supposed.

Why cognitive theories of consciousness do not require PFC decoding

Let us next consider why advocates of cognitive theories of consciousness have emphasized decoding from the PFC, asking whether decoding from the PFC is truly important to the aims of these theories.

Cognitive versus localist theories of consciousness

Consciousness research has been dominated by a conflict between two approaches. According to 'cognitive' theories of consciousness (notably, global workspace and higher-order thought theories), consciousness of a stimulus is based in processing involving thought, reasoning, noticing, and other forms of cognitive 'access' to the perceptible properties of the stimulus based in the PFC; these cognitions can be (but need not be) automatic, involuntary, and implicit. These theorists are sometimes called the 'prefrontalists.' (Please note that I use the terms 'prefrontalist' and 'PFC' to indicate the commitments of cognitive theories of consciousness. However, global workspace theories also emphasize parietal cortex, and, even when it comes to the PFC, what is really at issue is certain regions in front of the central sulcus, notably dorsolateral, ventrolateral, medial prefrontal, anterior cingulate, and orbitofrontal parts of the PFC [3,10–12].)

The alternative sensory approach is 'localist' in that it holds that perceptual consciousness is based in the circuits that process visual content in posterior sensory areas, although some localists also allow that cognitive phenomenology might be based partly in the PFC [13–15] and that perhaps some high-level perceptual contents [16,17] might be partially localized in the

Box 1. When decoding from the PFC conflicts with reports

Three different laboratories [7,43,44] found only a few hundred milliseconds of PFC/parietal activation, even though subjects were looking attentively at the stimulus for as long as a second and a half. Although reported consciousness has been the gold standard measure of awareness for prefrontalist approaches, this result has led some researchers to take seriously the idea that when you stare at something attentively for a few seconds, you are only conscious of it for a few hundred milliseconds, despite viewers' reports that they are conscious of it the whole time; that is, decoding has been taken to trump reports. In one study [43], the possibility that these transient PFC activations may be taken by some to indicate momentary consciousness despite reports to the contrary is mentioned five times (pp. 1, 2, 9, 10, 11). They consider the possibility that what seems to be continuous perception of a stimulus is composed of discrete samples, saying (p. 9), '…the introspective subjective percept is of stable continuous images with varying durations. Taken at face value, this would suggest that the sustained and stable visual representations in the ventral visual stream underpin our ongoing conscious experience. However, to the extent that perception is composed of discrete samples, each generating a transient ignition, the frontoparietal response would correspond more directly to experience.' This line of thought would fit with some theorists' interpretation of 'inattentional blindness' in which what seems to be a constant ongoing perception of details in the environment does not reflect what is really happening in conscious perception [52].

Glossary

Change blindness: when subjects fail to report changes in stimuli even when looking at them. Change blindness may be a matter of failing to apply concepts rather than a failure of perception.
Decoding: the use of brain imaging to reconstruct how the perceiver is representing the stimulus.
Dereferencing: a process by which the representation that is the object of a pointer is retrieved for further processing or use.
Global workspace theory of consciousness: holds that a perceptual representation is conscious if, and only if, it is transmitted, via long-range connections between perceptual and cognitive areas of the brain in the prefrontal cortex, so as to be globally available for processing by cognitive mechanisms.
Higher-order theories of consciousness: hold that a perception is conscious in virtue of another state that represents the perception. According to traditional 'rerepresentation' higher-order theories, for a perceptual representation of a square to be conscious, there must be a thought (usually considered to be based in the prefrontal cortex) about it to the effect that one is seeing a square, thus involving two representations of squareness.
Inattentional blindness: when subjects fail to report unusual stimuli if their attention is distracted. Inattentional blindness may be a matter of failing to apply concepts rather than a failure of perception.
Information Integration Theory (IIT): differs from global workspace and higher-order accounts of consciousness in that it can be seen as characterizing the relation '*x* is more conscious than *y*' rather than consciousness *per se*. IIT says roughly that *x* is more conscious than *y* to the extent that *x* is more differentiated than *y* and those differentiated elements are more integrated.
Postperceptual processing: processing applied to a perceptual representation, often for cognitive purposes. Conceptualization of a perceptual representation is one important type of postperceptual processing.
Pointer: a register can contain the address of another register, in which case the first register is regarded as

PFC. According to localists, decoding of perceptual contents in the PFC normally reflects post-perceptual cognitive processing. Some localists focus on what they call the 'hot zone' in posterior visual areas because their theories favor the kind of cortical organization found in perceptual areas [18]. The higher-order and global workspace camps of prefrontalists [19] recently joined forces to argue that their cognitive approach is the most promising approach to machine consciousness.

No report paradigms

Early results on the neural correlates of consciousness favored the PFC as an essential part of the neural basis of consciousness [20]. However, these early studies tended to focus on reported conscious states, thereby conflating the neural basis of consciousness with the decision and evaluation processes required to decide which response to make in order to report consciousness, as I pointed out in a series of articles [21–23]. Likewise, others [24] have pointed out that attentional activations prior to the stimulus also contaminated attempts to find the neural correlates of consciousness.

The major idea for dealing with this conflation was 'no-report' paradigms [25]. No-report paradigms rely on reports, but only indirectly. In no-report paradigms, reports are used to validate a measure of consciousness in some subjects; then, in other subjects, the measure can be used without the report. In the intracranial studies of binocular rivalry mentioned earlier [2,4], eye movements (optokinetic nystagmus) tell us whether the subject is experiencing motion in one direction or the other and are a good enough indicator of what the subject is experiencing to substitute for report. In some no-report paradigms, there is no task; in others, there are tasks that are not relevant to the variable of interest and do not affect responses in a relevant way. Another kind of no-report paradigm is discussed in the last section. Many studies have found decoding from the PFC for conscious contents using no-report paradigms [2,7,26–28].

Rich versus sparse perceptual representation

Another aspect of the prefrontalist/localist controversy is that perceptual theories suggest 'rich' conscious perception because many perceptual representations can be active at once, whereas cognitive theories suggest that consciousness is only as rich as our thinking about and cognitively accessing our perceptual states. Our cognitive concepts of our perceptions are much sparser than the perceptual contents themselves. For example, there are thousands of discriminable sounds but many fewer concepts of them [29]. Even people with perfect pitch can recognize only around 100 pitches [30].

Prefrontalists argue that psychological phenomena such as **inattentional blindness** and **change blindness** support the view that consciousness is sparse. Researchers have shown subjects naturalistic pictures and videos that change in various ways outside the fixation point, that point being specified by eye trackers. The pictures and videos were normal at the fixation point but were deprived of color or normal form outside the fixation point (the latter using texturization algorithms [31]). The overall result has been that subjects often fail to notice the weirdness of the peripheral displays [32,33]. Because in these paradigms the PFC representations are assumed to be 'normal' but perceptual representations of the environment in posterior visual areas are assumed to be abnormal (fitting the actual texturized or colorless periphery of the perception) but do not result in noticing, these results have been taken to support cognitive theories of consciousness. (I have argued against such ideas [1], but I will not discuss my objections here.)

Pointer theories

This article concerns two major reasons to think this emphasis on decoding from the PFC is a mistake. First, as recent versions of the higher-order approach have emphasized, consciousness

pointing to the content of the other register. In the context of theories of consciousness, PFC pointers monitor first-order representations, allowing their contents to be retrieved.

Rich versus sparse representation: in a variety of cognitive tasks involving alphanumeric symbols or other closed class items that are easy to discriminate from one another, humans can process only about four items at once. But perception seems to be able to represent many more items. The former representation is said to be sparse, and the latter is rich.

Box 2. Why double representation is problematic

As mentioned in the main text, standard versions of the higher-order approach have postulated two representations of a conscious perceptual content. For consciousness of a first-order representation of motion, the higher-order thought would be something like, 'I myself am experiencing visual motion.'

The main problem with this 'double representation' aspect of higher-order views has to do with conflicts between the higher-order thought content and the perceptual content [53,54]. Suppose I have the thought that I am seeing an entire surface as pure red and pure green at the same time. I can have that thought, but there will be no corresponding experience or first-order representation with that content without a specialized apparatus that allows both ends of the red/green opponent channel to be activated at once [55]. (You need a device that tracks the direction of eye gaze and projects an image to the same place on the retina, no matter how the eye moves.) Perhaps the thought has to be of some special type, but, as we will see in the next paragraph, no special kind of thought will work. A second issue is that if there really are two representations of every conscious content, we would expect there to be cases in which these representations conflict, but no one has ever given convincing evidence of such a thing. Of course, there are cases in which we have conflicting perceptions; for example, in perceptual crowding, subjects report confused and changing perceptions [56]. But, as noted in [57], these reports can be explained entirely in terms of conflicting representations in perceptual areas.

One might wonder why higher-order theorists did not abandon the double representation by getting rid of the requirement of any first-order state at all. In fact, some theorists did adopt the view that because of the possibility of an 'absent' first-order state, 'we are sometimes conscious of ourselves as being in various mental states that we are not actually in' [58]. Or, alternatively, 'one can be in a conscious state even if that state does not exist, so long as one is suitably aware of oneself as being in that state' [59]. But no kind of thought that I am visually experiencing a surface as pure red and pure green all over at the same time makes it so, given that special apparatus is needed for that experience.

can be based in the PFC without representation of perceptual contents in the PFC and thus without the possibility of decoding conscious contents from the PFC. Let me explain.

According to the higher-order thought account, a conscious perception of, say, motion becomes conscious by virtue of a cognitive state about that perception of motion. Older versions of the view required the thought to rerepresent the perceptual content [34]. This double representation was problematic for reasons discussed in Box 2. However, newer versions of the higher-order approach (Figure 1) are hybrid, requiring both first- and higher-order states, with the perceptual content in the first-order states (see [35] for a brief summary of scientific advantages). According to this approach, the cognitive representations in the PFC do not rerepresent but rather are **pointers** to or indexes of perceptual activations [10,35–38]. As Hakwan Lau puts it ([10], p. 134),

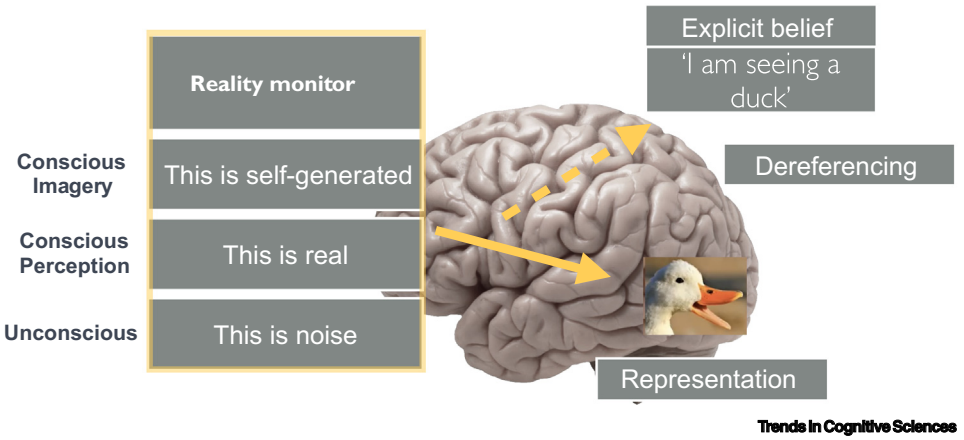

Figure 1. Cartoon depiction of the pointer version of the higher order theory. Perceptual monitoring requires a 'discriminator' that decides how perceptual contents are to be represented in cognition, as imagery ('This is self-generated'), as reliable perception ('This is real'), or unconsciously represented ('This is noise'). 'Dereferencing' is a computer science term for retrieving the content that a pointer points to. See [10,35].

'the role of the prefrontal cortex may not be to "duplicate" the sensory information. Rather, it may just monitor and redirect information in the sensory cortices, using something akin to indexing mechanisms.' The concept of a pointer is taken from computer science, where a register can contain the address of another register, sending the processing to that other register. In pointer theories, a PFC pointer is a link to a first-order representation, typically a perceptual representation of the outside world. The pointers can have contents involving the monitoring of probability or reliability (in some versions, subpersonal contents), but not perceptual contents. Pointers function to recruit first-order contents in the service of cognition concerning that content [35,39] (Figure 1).

Once one sees that monitoring pointers can have this role, it is natural to suggest that global workspace theories make the same modification [39,40]. The key concept of the global workspace viewpoint is 'ignition.' The idea of ignition is that there is a competition among neural coalitions in posterior visual areas. Only one or two of these can win out, and these winners can 'ignite' larger coalitions via long-range projections to the PFC and to parietal areas. These ignited activations are mutually reinforcing and allow the perceptual information to be accessed by any of the 'consumer' systems of reporting, decision-making, reasoning, and the like. The important point is that although previous versions have emphasized ignited rerepresentations in the PFC, what is ignited need only be a pointer to perceptual contents. The perceptual contents can be accessed by consumer systems via '**dereferencing**' the activated pointers, a process via which the contents are retrieved. As Mashour *et al.* describe it, the global workspace can act as a router 'through which information can be amplified, sustained, and made available to specialized sensory processors' [12].

It has commonly been thought that changes in conscious perception during binocular rivalry have been driven solely by a winner-takes-all competition between activations in perceptual areas in the occipital and temporal cortex, but a recent result suggests that these changes are driven instead by signals in the PFC [4,41]. These researchers show that one kind of PFC activity promoted binocular change and another promoted stability. It may be that these PFC signals are related to pointers.

My point has been that failure to decode from the PFC does not challenge the pointer view. The PFC can be pivotal even without decoding perceptual contents from the PFC. But what about the converse: does the success in decoding from the PFC described earlier challenge the pointer view? I think not, for several reasons. First and foremost, I think there is a good case that decoding from the PFC is due to postperceptual processing. That is the topic of the next section. But what if I am wrong about that? If we exclude postperceptual processing, do the decoding results from the PFC already mentioned challenge the pointer view? There are several ways rerepresentation and pointer views can coexist. Some decoding results might be explained by rerepresentation in the PFC for generic or abstract contents (e.g., red or rectangle). Results that suggest conscious representation in perceptual areas could be accommodated by combining this view with pointers for the full details of consciousness (the shade of red, the shape of the rectangle). One well-known criticism of rerepresentation views has been that they are incompatible with what we take to be the fine grain of perceptual phenomenology [42]. Combining rerepresentation for abstract contents with pointers to details defuses that objection. Alternatively, low-level contents such as orientation might be represented by PFC pointers, at least briefly, then replaced by a pointer to face representations in the temporal lobe. This might be suggested by the results of [7,43,44], but another explanation of those results, suggested by what I report in the next section, is that faces are simply of more interest to the subjects than orientations. (Another possibility is that – as in every case in which decoding fails – more fine-grained techniques could detect orientations.)

In sum, the pointer option for both higher-order and global workspace theories faces no problem of a lack of PFC decoding. See Box 3, however, for a challenge to PFC pointer views.

Why decoding from the PFC may not support cognitive theories of consciousness

I have been arguing that decoding of perceptual contents from the PFC shows too much (because cognitive theories do not require decoding); I shift now to pointing out that PFC decoding also shows too little (because it may reflect postperceptual processing rather than conscious perception).

Bored monkey problem

There is a persistent problem with the positive evidence for PFC decoding. Even no-report methodologies do not preclude subjects noticing or thinking about or otherwise cognizing the stimulus. I argued that this 'bored monkey' problem is especially pressing in the aforementioned binocular rivalry experiments in which monkeys see gratings moving one way, then another way, repeating over and over [45]. I said monkeys that are sitting in a primate chair with no task and nothing else to do might be noticing or otherwise cognizing the stimulus direction, contaminating the perceptual decoding from the PFC [15].

Other researchers [39,46] objected, arguing that reproducible stimulus-specific thoughts would be unlikely to occur as a result of boredom. However, what these replies ignore is that the cognitive categorization might be a kind of automatic cognition – or, as suggested in one of the replies [39], it could be due to a cognitive representation of confidence in the perceptual content.

Perceptual contents were decoded from PFC 60 milliseconds after the stimulus in a paradigm presenting ten pictures per second, each masking the previous picture [26]. The authors argue that masking precludes post-perceptual processing and since it has been found [47] that even at 12 pictures per second, subjects are above chance at post-perceptual matching of pictures with descriptions, the study is supposed to show conscious representation based in the PFC.

There are two errors in this reasoning. First, the PFC representations 60 milliseconds after the stimulus that they find are not conscious. They are likely to be a result of the low spatial frequency

Box 3. Problems with pointer theories

Although pointer views do not predict PFC decoding, they cannot use one of the advantages claimed for prefrontalist theories concerning rich versus sparse perception. I mentioned the 'inattentional blindness' results in which the periphery of the visual field loses color and form but subjects do not notice this. Rerepresentation versions of higher-order theories purport to explain this by postulating PFC representation of a periphery without loss of color and form, but pointer views cannot use this idea, because pointers do not have perceptual contents. (See Chapter 4 of [10]) and a reply in [1], pp. 436–442.) However, inattentional blindness results can be accommodated by pointer theories if they result from the failure to apply concepts, resulting in a failure to notice the peripheral features (see Chapter 6 of [1]).

Furthermore, some recent anti-PFC results would apply to pointer views. My colleagues and I [60] noted that a survey of intracranial electrical stimulation work failed to find reproducible perceptual perturbations of ongoing experience with PFC stimulation. Others [61] responded by noting that PFC representation is highly distributed, so one would not expect that electrical stimulation would yield a coherent percept and that stimulation that works for perceptual areas might work less well for the more distributed representations of PFC. This is true. However, the key finding is that ongoing perception (e.g., of the doctor's face) is not perturbed by intracranial electrical stimulation. If consciousness is based in PFC pointers, it would not be unreasonable to expect that electrical stimulation might cause a pointer to be degraded to the extent of perturbing the patient's conscious experience or might cause the pointer to point to a different perceptual representation, despite the differences between the PFC and perceptual areas. But no indication of such perturbations appears in the literature. With some exceptions (e.g., olfactory effects) as detailed in [60] and section 3.10 of [10], from the subjects' point of view, intracranial stimulations to the PFC are treated the same as sham stimulations in which the noise of a stimulation appears without the stimulation. Of course, this is a null result that should be interpreted with caution, and, as [46] noted, the stimulation could conceivably have affected subjects' ability to report the perturbations. Michel [39] discusses this issue, noting that progress will require putting together stimulation with behavioral tasks.

'fast feed-forward sweep' that is known to be relatively immune to backward masking [49,50] and in which unconscious representations are used to focus attention signals in perceptual areas. Earlier work on these early signals in PFC acknowledged that these signals precede scene segmentation and selection of attentional targets [50]. Second, the conscious representations occur after the feed forward sweep. As [47] notes, we should think of these experiments in terms of the 'carwash' model of masking in which multiple stimuli can be processed in parallel so long as they are not in the same stage of the 'carwash' [48] showed that pictures were poor masks of previous pictures at these high rates of processing, so it is likely that there is extensive futher processing of the pictures.

At this point in the debate, it seemed that we had a clash of intuitions that required further evidence to resolve. Fortunately, that further evidence was available.

Bifurcation dynamics

Sergent *et al.* [51] used a paradigm in which the target stimulus was task-relevant in some trials but not others. They did not use binocular rivalry; rather, they presented two French vowels, /a/ or /ə/, embedded in noise with varying signal-to-noise ratios. In 'active' (task-relevant) sessions, subjects were asked to report the identity and audibility of the vowels. In 'passive' (task-irrelevant) sessions, the vowels were present, but the tasks were unrelated to the vowels. Subjects' brain activity was recorded using electroencephalogram (EEG). There were two key findings, one related to the no-report methodology, the other related to the bored monkey problem.

The no-report result was that at threshold (behaviorally determined), there was a bimodal distribution of activation: Sometimes there was widespread activation between 250 ms and 700 ms post–vowel stimulus and sometimes not, with few intermediate cases, what they called 'bifurcation dynamics.' Widespread activation accurately predicted subjects' reports in the active version of the study. That is, when there was widespread activation linked to the vowel onset, subjects reported the identity and audibility of the vowels, and when there was no widespread activation, they did not. Here is the methodological result: the authors [51] were able to use a version of bifurcation dynamics as a substitute for report in the passive (no vowel report) sessions. That is, even when subjects were not reporting the vowels, widespread activation yoked to the sounds predicted awareness of them, as shown by a 'mind-wandering' method to be described later.

The widespread activation in the passive (no-report) condition was not as widespread as global workspace activation, leaving out PFC areas involved in executive function. Sergent *et al.* [51] term the widespread activation in the passive condition 'global playground' activation, where global playground activation is a subset of global workspace activation. Global playground activations are denuded of the decision and executive processes underlying report, and this makes them a better candidate for the neural basis of access-consciousness than global workspace activations (Figure 2).

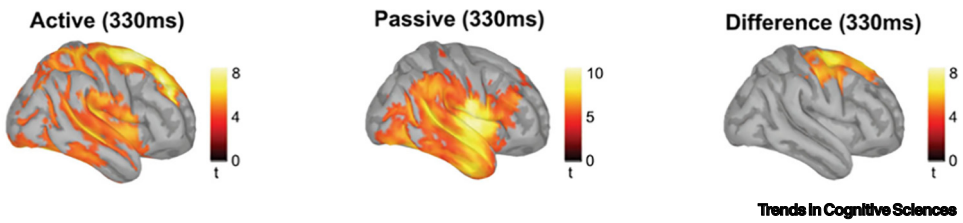

Figure 2. Reconstruction from electroencephalogram (EEG) data of activations 330 ms from stimulus onset. The leftmost figure depicts global workspace activation, the middle depicts global playground activation, and the rightmost figure represents the difference. EEG data are not very locationally precise, so these reconstructions must be regarded as very approximate, as emphasized by Sergent and colleagues [51].

Mind wandering

This result appears to support the prefrontalists since the global playground activation included a PFC component. However, there was another key result. Recall that half the sessions were 'active' in which subjects were asked to identify and report the audibility of the vowel. In the 'passive' sessions, there were four different tasks, one or another of which was substituted for the vowel task, although the vowels were still presented. The four tasks were (i) a visual task (detecting a large green circle); (ii) a multiple-choice task on arithmetic, general information, and other topics; (iii) pressing a 'click to continue' button; and (iv) finally, the task of interest to us, a mind wandering probe.

In the mind wandering probe, subjects were asked (in French), 'What is on your mind just now?' with four options: (i) 'the sound,' (ii) 'my thoughts,' (iii) 'the task,' and (iv) 'nothing/I feel sleepy.' Most of the subjects' responses in the mind-wandering task reflected the visual task or 'my thoughts.' Only 19% of the responses were 'the sound,' with 16% responding 'the sound' at the lowest intensity (signal-to-noise ratio) and 33% at the highest. So, even at the highest audibility level, most reports were not 'the sound.'

Here is the result of interest: when there was global playground activation linked to the sound, subjects tended to give 'the sound' answer to the mind-wandering probe. The authors conclude (p. 11), '…with the bifurcation model of conscious access, we could use neural activity to predict whether participants were spontaneously aware of sounds in a passive listening condition.' This result shows that global playground activation predicted awareness of the sound because participants would not be thinking about the sound if they had not heard it. But this result also suggests that it may be cognition of the sound that was responsible for the global playground activation rather than the consciousness of the sound itself. Global playground activation predicts mind-wandering reports of the sound on their mind. Reports most directly reflect cognition, whereas conscious perception of the sound is inferred from cognition of the sound. Given that the reports most directly reflect cognition of the sound, and we know that the PFC is the home of cognition, we must accept the possibility that the global playground activations reflect, entirely or in part, cognition rather than conscious perception.

Some cognitions – for example, automatic cognitions perhaps including thoughts about the sounds – may be regarded by cognitive theorists as a component of conscious experience, but the bored monkey problem does challenge them to defend that view.

This result is enough to show that the bored monkey problem is real, but we can go further. Note that subjects could also have been aware of the sound when they reported 'other things' as on their minds, such as the arithmetic or the green circle. Those of us who teach in New York City have had the experience of teaching a class or attending a lecture while right outside the window there is a loud garbage truck. Of course, some may be able to gate out the sounds, but others of us have had the experience that we would report as being painfully aware of the garbage truck while thinking about how to answer a question. If I choose to report thinking about how to answer a question instead of the sound of the garbage truck, it may be because the answering task seems more respectable or is more salient to me at the moment, not because I am unaware of the garbage truck. Needless to say (but I will say it anyway), we can all think of many ways in which the reports could be wrong or could mislead us. But we should take reports seriously. If we do, we have a further reason for thinking the bored monkey problem is real.

Importantly, the signal-to-noise ratio of the sounds was varied in the experiment, and, when it was very low, subjects were only 50% accurate (i.e., at chance) in the active condition. However, in a

crucial band of signal-to-noise ratios, subjects were roughly 95% accurate in identifying the vowels in the active condition, even though at this signal-to-noise range, they were mostly not giving 'the sound' response. So, subjects were reporting that arithmetic or the green circle or the answers to questions were on their minds, even when the sounds were clearly identifiable at least when they were task-relevant.

Thus, reports of arithmetic or the green circle as being 'on my mind' in the mind-wandering task may be accompanied by background awareness of the sounds, just as awareness of the garbage truck can remain part of the conscious background of my thoughts about answering the question. So, there may have been awareness of the sound even without global playground activation.

Concluding remarks

Returning to the topic of decoding perceptual content from the PFC, we cannot rule out the possibility that the impressive intracranial decoding from the PFC might reflect in part the contents of postperceptual cognition rather than conscious contents. This fact indicates the 'bored monkey' problem is real, suggesting caution about overinterpreting decoding binocular rivalry contents in monkeys (but see Outstanding questions). Decoding from the PFC has played an outsize role in debates about the neural basis of consciousness, but it is doubly mistaken. First, failures of decoding from the PFC are harmless for theories of consciousness to the extent that pointer versions of those theories work. Second, successes in decoding from the PFC can reflect thought processes based on the stimuli rather than conscious perception of them.

Outstanding questions

How can the pointers postulated by pointer theories be found in the PFC?

If sustained perception of high-level properties starts with a brief burst of rerepresentation in the PFC and then switches to a pointer, how does that happen and why?

How can we adjudicate between the pointer versions of cognitive theories and the relational theories of Rafi Malach and colleagues?

How can we develop more sophisticated mind-wandering probes to ascertain whether bifurcation dynamics reflects postperceptual cognitive processing?

Acknowledgments

Thanks to Matthias Michel for Figure 1 and Claire Sergent for Figure 2.

Declaration of interests

The author has no interests to declare.

References

1. Block, N. (2023) *The Border between Seeing and Thinking*, Oxford University Press
2. Kapoor, V. *et al.* (2022) Decoding internally generated transitions of conscious contents in the prefrontal cortex without subjective reports. *Nat. Commun.* 13, 1535
3. Michel, M. and Morales, J. (2020) Minority reports: consciousness and the prefrontal cortex. *Mind Lang.* 35, 493–513
4. Dwarakanath, A. *et al.* (2023) Bistability of prefrontal states gates access to consciousness. *Neuron* 111, 1666–1683
5. Bellet, J. *et al.* (2022) Decoding rapidly presented visual stimuli from prefrontal ensembles without report nor post-perceptual processing. *Neurosci. Conscious.* 2022, niac005
6. Rowe, E.G. *et al.* (2024) Feedforward connectivity patterns from visual areas to the front of the brain contain information about sensory stimuli regardless of awareness or report. *Cortex* 172, 284–300
7. Ferrante, O. *et al.* (2023) An adversarial collaboration to critically evaluate theories of consciousness. *bioRxiv*, 2023.2006.2023.546249
8. Schmolesky, M.T. *et al.* (1998) Signal timing across the macaque visual system. *J. Neurophysiol.* 79, 3272–3278
9. Xu, H. *et al.* (2016) Rivalry-like neural activity in primary visual cortex in anesthetized monkeys. *J. Neurosci.* 36, 3231
10. Lau, H. (2022) *In Consciousness We Trust: The Cognitive Neuroscience of Subjective Experience*, Oxford University Press
11. Lau, H. and Brown, R. (2019) The emperor's new phenomenology? The empirical case for conscious experiences without first-order representations. In *Themes from Block* (Pautz, A. and Stoljar, D., eds), pp. 171–197, MIT Press
12. Mashour, G.A. *et al.* (2020) Conscious processing and the global neuronal workspace hypothesis. *Neuron* 105, 776–798
13. Malach, R. (2022) The role of the prefrontal cortex in conscious perception: the localist perspective. *J. Conscious. Stud.* 29, 93–114
14. Lamme, V.A.F. (2020) Visual functions generating conscious seeing. *Front. Psychol.* 11, 83
15. Block, N. (2020) Finessing the bored monkey problem. *Trends Cogn. Sci.* 24, 167–168
16. Tsao, D.Y. *et al.* (2008) Patches of face-selective cortex in the macaque frontal lobe. *Nat. Neurosci.* 11, 877–879
17. Kornblith, S. and Tsao, D.Y. (2017) How thoughts arise from sights: inferotemporal and prefrontal contributions to vision. *Curr. Opin. Neurobiol.* 46, 208–218
18. Tononi, G. and Koch, C. (2015) Consciousness: here, there and everywhere? *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 370, 20140167
19. Dehaene, S. *et al.* (2017) What is consciousness, and could machines have it? *Science* 358, 486–492
20. Rees, G. *et al.* (2002) Neural correlates of consciousness in humans. *Nat. Rev. Neurosci.* 3, 261–270
21. Block, N. (2005) Two neural correlates of consciousness. *Trends Cogn. Sci.* 9, 46–52
22. Block, N. (1996) How can we find the neural correlate of consciousness? *Trends Neurosci.* 19, 456–459
23. Block, N. (2001) Paradox and cross purposes in recent work on consciousness. *Cognition* 79, 197–220
24. Aru, J. *et al.* (2012) Distilling the neural correlates of consciousness. *Neurosci. Biobehav. Rev.* 36, 737–746
25. Tsuchiya, N. *et al.* (2015) No-report paradigms: extracting the true neural correlates of consciousness. *Trends Cogn. Sci.* 19, 757–770

26. Finkel, E. (2023) Consciousness hunt yields results but not clarity: two rival theories of how it emerges went head-to-head in brain-scanning experiments. *Science* 380, 1309–1310
27. Panagiotaropoulos, T.I. *et al.* (2012) Neuronal discharges and gamma oscillations explicitly reflect visual consciousness in the lateral prefrontal cortex. *Neuron* 74, 924–935
28. Hatamimajoumerd, E. *et al.* (2022) Decoding perceptual awareness across the brain with a no-report fMRI masking paradigm. *Curr. Biol.* 32, 4139–4149.e4
29. Raffman, D. (1995) On the persistence of phenomenology. In *Conscious Experience* (Metzinger, T., ed.), pp. 293–308, Ferdinand Schoningh
30. Levitin, D. (2005) Absolute pitch: perception, coding, and controversies. *Trends Cogn. Sci.* 9, 26–33
31. Freeman, J. and Simoncelli, E. (2011) Metamers of the visual stream. *Nat. Neurosci.* 14, 1195–1201
32. Cohen, M.A. and Rubenstein, J. (2020) How much color do we see in the blink of an eye? *Cognition* 200, 104268
33. Cohen, M.A. *et al.* (2021) Characterizing a snapshot of perceptual experience. *J. Exp. Psychol. Gen.* 150, 1695–1709
34. Rosenthal, D. (1997) A Theory of Consciousness. In *The Nature of Consciousness: Philosophical Debates* (Block, N. and Flanagan, O. and Güzeldere, G., eds), pp. 729–754, MIT Press
35. Michel, M., The perceptual reality monitoring theory, In *Scientific Theories of Consciousness: The Grand Tour* (Herzog, M. et al., eds), Cambridge University Press (in press).
36. Brown, R. *et al.* (2019) Understanding the higher-order approach to consciousness. *Trends Cogn. Sci.* 23, 754–768
37. Cleeremans, A. *et al.* (2020) Learning to be conscious. *Trends Cogn. Sci.* 24, 112–123
38. Lau, H. (2008) A higher order Bayesian decision theory of consciousness. *Prog. Brain Res.* 168, 35–48
39. Michel, M. (2022) Conscious perception and the prefrontal cortex: a review. *J. Conscious. Stud.* 29, 115–157
40. Wu, W. (2014) Being in the workspace, from a neural point of view: comments on Peter Carruthers, 'On central cognition'. *Philos. Stud.* 170, 163–174
41. Jensen, O. and Ferrante, O. (2023) Prefrontal cortex oscillations update and stabilize consciousness during binocular rivalry. *Neuron* 111, 1519–1520
42. 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), pp. 199–213, MIT Press
43. Vishne, G. *et al.* (2023) Distinct ventral stream and prefrontal cortex representational dynamics during sustained conscious visual perception. *Cell Rep.* 42, 112752
44. Broday-Dvir, R. *et al.* (2023) Perceptual stability reflected in neuronal pattern similarities in human visual cortex. *Cell Rep.* 42, 112614
45. Block, N. (2019) What is wrong with the no-report paradigm and how to fix it. *Trends Cogn. Sci.* 23, 1003–1013
46. Panagiotaropoulos, T.I. *et al.* (2020) Prefrontal cortex and consciousness: beware of the signals. *Trends Cogn. Sci.* 24, 343–344
47. Potter, M. *et al.* (2014) Detecting meaning in RSVP at 13 ms per picture. *Atten. Percept. Psychophys.* 76, 270–279
48. Maguire, J. and Howe, P. (2016) Failure to detect meaning in RSVP at 27 ms per picture. *Atten. Percept. Psychophys.* 78, 1405–1413
49. Vanrullen, R. (2008) The power of the feed-forward sweep. *Adv. Cogn. Psychol.* 3, 167–176
50. Kirchner, H. and Thorpe, S. (2006) Ultra-rapid object detection with saccadic eye movements: Visual processing speed revisited. *Vis. Res.* 46, 1762–1776
51. Sergent, C. *et al.* (2021) Bifurcation in brain dynamics reveals a signature of conscious processing independent of report. *Nat. Commun.* 12, 1149
52. Cohen, M. and Dennett, D. (2011) Consciousness cannot be separated from function. *Trends Cogn. Sci.* 15, 358–364
53. Block, N. (2011) The higher order approach to consciousness is defunct. *Analysis* 71, 419–431
54. Neander, K. (1995) Misrepresenting and malfunctioning. *Philos. Stud.* 79, 109–141
55. Billock, V.A. *et al.* (2001) Perception of forbidden colors in retinally stabilized equiluminant images: an indication of softwired cortical color opponency? *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* 18, 2398–2403
56. Lettvin, J.Y. (1976) On seeing sidelong. *The Sciences* 16, 10–20
57. Pelli, D. and Tillman, K. (2008) The uncrowded window of object recognition. *Nat. Neurosci.* 11, 1129–1135
58. Rosenthal, D. (2009) Concepts and definitions of consciousness. In *Encyclopedia of Consciousness* (Banks, W., ed.), pp. 157–169, Academic Press
59. Weisberg, J. (2011) Abusing the notion of what-it's-like-ness: a response to Block. *Analysis* 71, 438–443
60. Raccah, O. *et al.* (2021) Does the prefrontal cortex play an essential role in consciousness? Insights from intracranial electrical stimulation of the human brain. *J. Neurosci.* 41, 2076–2087
61. Naccache, L. *et al.* (2021) Why intracranial electrical stimulation of the human brain suggests an essential role for prefrontal cortex in conscious processing. *J. Neurosci.*, Published online October 21, 2021. https://doi.org/10.31219/osf.io/zrqp8