Conscious, preconscious, and subliminal processing: a testable taxonomy

Stanislas Dehaene1,2, Jean-Pierre Changeux2,3, Lionel Naccache1, Jérôme Sackur1 and Claire Sergent1

1INSERM-CEA Cognitive Neuroimaging Unit, Service Hospitalier Frédéric Joliot, Orsay, France
2Collège de France, Paris, France
3CNRS Unit, Receptors and Cognition, Institut Pasteur, Paris, France

Of the many brain events evoked by a visual stimulus, which are specifically associated with conscious perception, and which merely reflect non-conscious processing? Several recent neuroimaging studies have contrasted conscious and non-conscious visual processing, but their results appear inconsistent. Some support a correlation of conscious perception with early occipital events, others with late parieto-frontal activity. Here we attempt to make sense of these dissenting results. On the basis of the global neuronal workspace hypothesis, we propose a taxonomy that distinguishes between vigilance and access to conscious report, as well as between subliminal, preconscious and conscious processing. We suggest that these distinctions map onto different neural mechanisms, and that conscious perception is systematically associated with surges of parieto-frontal activity causing top-down amplification.

Introduction
Understanding the neuronal mechanisms of consciousness is a major challenge for cognitive neuroscience. Recently, great progress has been achieved by contrasting brain activation images obtained during minimally different experimental conditions, one of which leads to conscious perception while the other does not. Surprisingly, however no coherent picture has emerged from those experiments. On the contrary, a controversy has arisen, as some studies suggest that consciousness depends mostly on the thalamus and brain stem [1], others on early visual areas [2,3], and yet others on higher prefrontal and parietal association areas [4–9].

Here, we propose that those apparent contradictions can be resolved by a relevant theorizing of the physiological conditions for conscious processing of sensory stimuli. Based on the recent proposal of a large-scale thalamo-cortical formal network and its simulations [4,5], we tentatively propose a plausible and testable taxonomy of brain activity states associated with conscious and non-conscious processing. In particular, within non-conscious processing, we distinguish a transient ‘preconscious’ state of activity in which information is potentially accessible, yet not accessed.

An enabling condition: vigilance
The term ‘consciousness’ has multiple meanings, one of them intransitive (e.g. ‘the patient regained consciousness’), and the other transitive (e.g. ‘consciousness of color’). To avoid further confusion, we abandon the term and use ‘states of vigilance’ to refer to the non-transitive meaning, i.e. a continuum of states which encompasses wakefulness, sleep, coma, anesthesia, etc.

Being in an appropriate state of vigilance (e.g. awake rather than asleep) is an obvious enabling condition for conscious processing of sensory stimuli. Empirically, awakening into the vigilant state correlates with a progressive increase in regional cerebral blood flow, first in the brainstem and thalamus, then in the cortex with a particularly important increase in prefrontal-cingulate activation and functional connectivity [10]. Anesthesia, sleep, vegetative state and coma [1,11] are all associated with modulations of the activity of this large-scale thalamocortical network which also shows high baseline activity during vigilant rest [12] and encompasses prefrontal, cingulate and inferior parietal nodes.

These observations may be captured by a recent implementation of the neural workspace model [4] in which ascending brain stem nuclei (e.g. cholinergic among others) send globally depolarizing neuromodulatory signals to a thalamic and cortical hierarchy. Simulations show a progressive increase in spontaneous firing as a function of neuromodulator release, which evolves into what is known in dynamical systems theory as a Hopf bifurcation: spontaneous firing increases continuously in intensity, but high-frequency oscillations appear suddenly in the gamma band (20–80 Hz). By increasing spontaneous activity, and thus bringing a broad thalamo-cortical network closer to firing threshold, vigilance lowers the threshold for external sensory inputs.

In summary, vigilance is a graded variable, and a minimum level is essential for placing thalamo-cortical systems into a receptive state.
Early visual activation is not sufficient for conscious report
We now consider the neural bases of the second, transitive meaning of consciousness, which we term ‘access to conscious report’. How do we consciously perceive a visual stimulus? Many neuroimaging experiments have demonstrated a tight correlation between the conscious visual perception and the activation of striate and extrastriate visual areas [13–18]. For instance, unmasking of a visual stimulus increases activity in extrastriate areas in tight correlation with subjective reports of stimulus visibility [18]. Furthermore, extrastriate regions clearly play a causal role in conscious visual perception, because their selective lesioning eliminates the corresponding contents from experience – for instance a lesion of area V4 can destroy color perception in the contralateral hemifield [19].

On the basis of such data, Zeki [2] has proposed that the conscious perception of a given visual attribute resides in the extrastriate area specialized for that attribute (e.g. area MT/V5 for motion, or area V4 for color). A ‘micro-consciousness’ would be involved whenever that area receives a sufficient amount of activation.

We argue, however, that early sensory activation is necessary but not sufficient for conscious access, because activity in extrastriate visual areas is frequently observed while participants deny having seen any stimulus [14,20–23]. When invisibility is caused by masking [20] or by dichoptic stimulation [14] this stimulus-evoked activity remains weak, and one might argue that its small amplitude alone could explain the absence of conscious perception [2,14]. However, the visual activation evoked by invisible stimuli can also be very strong, for instance when invisibility is caused by neglect [21] or inattention [22,23]. In a recent study of the attentional blink, we observed that up to about 180 ms after stimulus presentation, the occipito-temporal event-related potentials evoked by a invisible word were large and essentially indistinguishable from those evoked by a visible word [23]. Yet on invisible trials, the participants’ visibility ratings did not deviate from the lowest value, used when no word was physically present. Thus, intense occipito-temporal activation can be accompanied by a complete lack of conscious report.

Top-down amplification, long-distance reverberation, and reportability
We [4–6] and others [7,8,24] have suggested that, in addition to vigilance and bottom-up activation, a third factor underlying conscious access is the extension of brain activation to higher association cortices interconnected by long-distance connections and forming a reverberating neuronal assembly with distant perceptual areas. Why would this brain state correspond to conscious access? Neurocomputational simulations show that once stimulus-evoked activation has reached highly interconnected associative areas, two important changes occur: (1) The activation can reverberate, thus holding information on-line for a long duration essentially unrelated to the initial stimulus duration; (2) Stimulus information can be rapidly propagated to many brain systems. We argue that both properties are characteristic of conscious information processing which in our view is associated with a distinct internal space, buffered from fast fluctuations in sensory inputs, where information can be shared across a broad variety of processes including evaluation, verbal report, planning and long-term memory [25].

Empirically, access of sensory stimuli to conscious report correlates with the activation of higher associative cortices, particularly parietal, prefrontal and anterior cingulate areas. In fMRI, the activation of those regions systematically separates masked versus unmasked presentations of words [20] or images [26]; undetected versus detected changes during change blindness [27,28]; extinguished versus seen visual stimuli in neglect patients [29]; or missed versus reported stimuli during the attentional blink [9,22,23,30–32]. In many of these paradigms, anterior activation is accompanied by an amplification and an increase in functional correlation with posterior stimulus-specific areas [20,26,30]. Sudden parieto-frontal activation and top-down amplification are two frequent signatures of conscious perception.

Is attention a confound or a necessity for conscious access?
Some have argued that many of the above neuroimaging paradigms are inappropriately controlled because conscious perception is confounded with increased attention and more extended stimulus processing. For instance, a conscious word can be attended, repeated or memorized while a non-conscious word cannot. Such confounds would suffice to explain the greater parieto-prefrontal activity to unmasked words [20]. For this reason, Tse et al. [18] have argued that one should prefer experimental designs in which attention is drawn away from the stimulus. They show that, in such a situation, correlates of stimulus visibility are found solely in occipital areas, not in higher associative regions, and therefore argue that the mechanisms of conscious visual perception lie in extrastriate cortex.

We obviously agree on one point: it is important to design paradigms in which conscious perception is not confounded with massive changes in overt or covert behaviour. However, this goal has been achieved in several studies. In our recent study of the attentional blink [23], for instance, subjects viewed a fixed stimulus and made similar motor gestures on seen and not-seen trials, yet those were still distinguished by strong parieto-frontal activation.

We question, however, the proposal that inattention is an appropriate control. Under conditions of diverted attention, such as those studied by Tse et al. [18], even an unmasked stimulus is not guaranteed to be consciously perceived. On the contrary, considerable evidence indicates that without attention, conscious perception cannot occur. In the inattentional blindness paradigm, even a 700-ms stimulus presented in the fovea, when unattended, might fail to be seen [33]. During the attentional blink, a mildly masked stimulus, normally quite visible, becomes invisible when attention is diverted to another task [23,34].
The relations between stimulus strength, attention, and conscious perception are complex because attention mechanisms can also be activated automatically in a bottom-up manner. When the stimuli have strong energy, sharp onsets or strong emotional content, they might trigger an activation of frontal eye fields or amygdala pathways, thus causing an amplification that can lower their threshold for conscious perception [35]. Thus, both bottom-up stimulus strength and top-down attentional amplification (whether triggered voluntarily or by automatic attraction) are jointly needed for conscious perception, but they might not always be sufficient for a stimulus to cross the threshold for conscious perception. Conscious perception must therefore be evaluated by subjective report, preferably on a trial-by-trial basis. Verifying that the stimuli can be consciously perceived in a separate experimental block where they are attended, as done by Tse et al. [18], does not suffice to guarantee conscious perception in a different block where attention is diverted. One cannot simply assume that, by unmasking stimuli, one is studying the neural correlates of conscious processing.

**Distinguishing accessibility from access**
The above distinctions lead us to proposal a formal definition of two types of non-conscious processes (Figure 1):

1. **Subliminal processing.** We define subliminal processing (etymologically ‘below the threshold’) as a condition of information inaccessibility where bottom-up activation is insufficient to trigger a large-scale reverberating state in a global network of neurons with long range axons. Simulations of a minimal thalamo-cortical network [4] indicates that such a non-linear self-amplifying system possesses a well-defined dynamical threshold. A processing stream that exceeds a minimal activation level quickly grows until a full-scale ignition is seen, while a slightly

**Figure 1.** Proposed distinction between subliminal, preconscious, and conscious processing. Three types of brain states are schematically shown, jointly defined by bottom-up stimulus strength (on the vertical axis at left) and top-down attention (on the horizontal axis). Shades of color illustrate the amount of activation in local areas, and small arrows the interactions among them. Large arrows schematically illustrate the orientation of top-down attention to the stimulus, or away from it (‘task-unrelated attention’). Dashed curves indicate a continuum of states, and thick lines with separators indicate a sharp transition between states. During **subliminal processing**, activation propagates but remains weak and quickly dissipating (decaying to zero after 1–2 seconds). A continuum of subliminal states can exist, depending on masking strength, top-down attention, and instructions (see Box 1). During **preconscious processing**, activation can be strong, durable, and can spread to multiple specialized sensori-motor areas (e.g. frontal eye fields). However, when attention is oriented away from the stimulus (large black arrows), activation is blocked from accessing higher parieto-frontal areas and establishing long-distance synchrony. During **conscious processing**, activation invades a parieto-frontal system, can be maintained ad libidum in working memory, and becomes capable of guiding intentional actions including verbal reports. The transition between preconscious and conscious is sharp, as expected from the dynamics of a self-amplified non-linear system [4].
Weaker activation quickly dies out. Subliminal processing corresponds to the latter type.

Note that, under our hypothesis, subliminal processing is not confined to a passive spreading of activation, independent of the subject's attention and strategies, as previously envisaged. On the contrary, whichever task and attentional set are prepared consciously can orient and amplify the processing of a subliminal stimulus, even if its bottom-up strength remains insufficient for global ignition. In agreement with this analysis, many top-down influences on subliminal processing have now been experimentally observed (Box 1).

(2) Preconscious processing. Freud [36] noted that 'some processes [...] may cease to be conscious, but can become conscious once more without any trouble', and he proposed that 'everything unconscious that behaves in this way, that can easily exchange the unconscious condition for the conscious one, is therefore better described as "capable of entering consciousness" or as preconscious.'

Here we further specify the latter term. We propose to call preconscious (or potentially conscious, or P-conscious) a neural process that potentially carries enough activation for conscious access, but is temporarily buffered in a non-conscious store because of a lack of top-down attentional amplification (for example, owing to transient occupancy of the central workspace system). As shown by the attentional blink and inattentional blindness paradigms, even strong visual stimuli can remain temporarily preconscious. They are potentially accessible (they could quickly gain access to conscious report if they were attended), but they are not consciously accessed at the moment.

At the neurocomputational level, preconscious processing is proposed to involve resonant loops within medium range connections which maintain the representation of the stimulus temporarily active in a sensory buffer for a few hundred milliseconds. A preconscious stimulus might ultimately achieve conscious access once the central workspace is freed (as exemplified by the psychological refractory period paradigm [37,38], in which one task is put on hold while another task is being processed). It might never gain access to conscious processing if the preconscious buffer is erased before the orienting of top-down attention (as achieved by masking in the attentional blink paradigm).

Box 1. Why attention and consciousness are different: top-down influences on subliminal processing

Subliminal processing is frequently thought to be automatic and independent of attention. However, the present framework implies that top-down attention and task set can have an effect on subliminal processing (see Figure 1 in main text, top row). This prediction has been verified in several recent reports.

Modulation of subliminal priming by temporal attention

In a numerical masked priming paradigm, Naccache et al. [43] first showed that subliminal priming was present when subjects could allocate attention to the prime-target pair, but vanished when stimuli could not be temporally attended. Kiefer and Brendel [44] observed a similar effect in an experiment investigating the N400 potential elicited by masked words. Unseen masked words elicited a much larger N400 when they were temporally attended than when they were not.

Modulation by spatial attention

Kentridge et al. [45,46] first reported that blindsight patient GY could use consciously perceived cues to enhance unconscious processing of visual targets. When a target was presented in his blind visual field, GY responded faster and more accurately when it was validly cued by a consciously perceptible arrow pointing to it, than when it was invalidly cued. In both cases, he still claimed that he could not see the target. Modulation of subliminal priming by spatial attention was also observed in normal subjects [47].

Modulation by strategies

Task instructions also alter the fate of subliminal stimuli. For instance, masked primes can elicit instruction-dependent activation in motor cortex [48,49], suggesting that arbitrary stimulus–response mappings conveyed by conscious instructions can also apply to non-conscious stimuli. The influence is always unidirectional: once a strategy or response mapping is consciously adopted, it extends to non-conscious primes [50,51]. Kunde et al. [51] studied the ‘Gratton effect’, a strategic increase in executive control that follows Stroop interference trials. They observed this effect following conscious conflict trials, but not following subliminal conflict trials. Once established, however, the increase in control applied to both subliminal and supraliminal trials – another instance of a top-down effect on subliminal processing.

Accounting for conflicting neuroimaging data

In experimental studies of conscious perception, preconscious processing, as an intermediate category, has sometimes been confounded with subliminal processing, and sometimes with conscious processing. We now show how this distinction can provide a simple account of conflicting neuroimaging results (Figure 2).

(1) Masking when stimuli are attended. Some experiments require participants to attend to masked stimuli which are made visible or invisible by changing the masking strength. In our taxonomy, those experiments contrast subliminal versus conscious stimuli – a major contrast which should reveal both early stimulus processing regions and a distributed parieto-frontal workspace system. Indeed, empirically, both early extrastriate and late parietal and prefrontal differences have been observed [20,26].

(2) Stimuli presented at threshold. Even when attended, stimuli presented at sensory threshold may or may not be perceived. In our theory this is again a contrast between subliminal and conscious stimuli. As predicted, neuroimaging experiments relying on this contrast have yielded both early (e.g. P100) and late (e.g. P300) correlates of conscious perception [15,16,39]. The theory can also explain why conscious access fluctuates even though the stimulus remains constant. Simulations show that the threshold for global ignition can vary both with vigilance and with the amount of spontaneous activity before stimulus presentation [4]. Several experiments confirm that the perception of near-threshold stimuli can be predicted by the prestimulus state, in both humans and monkeys [40,41].

(3) Masking when stimuli are not attended. If stimuli are not attended, then the comparison of masked and
unmasked stimuli becomes a contrast between subliminal and preconscious processing. As predicted, only the early components of occipito-temporal activation are seen [18]. According to our terminology, these are the correlates of preconscious visual processing (potential visibility, yet no conscious access).

(4) Stimuli made invisible by inattention. Some experiments have contrasted consciously perceived stimuli with stimuli made invisible by diverting top-down attention (attentional blink, change blindness, inattentional blindness). This is a contrast between preconscious and conscious processing. As expected, the difference appears late (200–300 ms after the stimulus) and involves parieto-prefrontal activation as well as late amplification of posterior activity [22,23,26–30,32].

Conclusion

Instead of the classical binary separation between non-conscious and conscious processing, we introduce here a tripartite distinction between subliminal, preconscious, and conscious processing. The key idea is that, within non-conscious states, it makes a major difference whether stimuli invisibility is achieved by a limitation in bottom-up stimulus strength, or by the temporary withdrawal of top-down attention. The first case corresponds to subliminal processing, the second to preconscious processing. We have shown how this distinction is theoretically motivated and helps make sense of neuroimaging data.

Is our taxonomy complete? Box 2 briefly discusses three other types of non-conscious knowledge in the nervous system: latent connectivity patterns, distributed firing patterns, and functionally disconnected brain systems. Altogether, these distinctions might suffice to capture the known experimental conditions in which information escapes conscious reportability. The proposed taxonomy is testable, not only with neuroimaging tools, but also using electrophysiological techniques in the awake monkey, provided that tasks similar to the attentional blink and psychological refractory period can be developed in these species (see Box 3).

Our proposal could also lead to a reconciliation of several major theories of conscious perception. The distinction between preconscious and conscious processing is consistent with Lamme's proposal of a progressive build-up of recurrent interactions, first locally within the visual system, and second more globally into parieto-frontal regions [3]. It is also consistent with Zeki's hypothesis of an asynchronous construction of visual perception in multiple distributed sites before binding into a
‘macro-consciousness’ [2]. Our only source of disagreement – but an important one – resides in their attribution of ‘phenomenal consciousness’ or ‘micro-consciousness’ to what we have termed pre-conscious processing. Remember that, in such a state, only a few hundreds of milliseconds after a stimulus was presented and yielding strong visual activity, participants deny perceiving anything [34]. Whether they actually had a conscious phenomenal experience but no possibility of reporting it, does not seem to be, at this stage, a scientifically addressable question. The only rationale for attributing phenomenal-consciousness to preconscious processing seems to be the intuition that visual experience involves a richness of content that goes beyond what we can report [42]. However, this intuition itself can be explained as a kind of illusion – we think that we see more than we actually do (Box 4).

To further explore these difficult issues in the future, it will be crucial to develop better methods for the formal collection and quantification of introspective reports [23,34], as well as for the study of the spontaneous flow of conscious processes [4,12]. Meanwhile, the proposed distinction between subliminal, preconscious and conscious processing, and the identification of conscious workspace hypothesis states that it must be represented by small groups of neurons whose firing provides an unambiguous index of the relevant attribute, and which would be amplified by top-down attention. For instance, although the collective firing of V1 neurons encodes all aspects of the visual scene, including the presence of faces or color, those attributes would not be consciously seen unless the extrastriate areas involved in their extraction are intact. At a higher cognitive level, when we gain conscious access to a previously subliminal distinction (e.g. development of ‘phonemic awareness’ in children), neuronal populations selective for this learned distinction should be found.

(i) Information is not encoded in neuronal firing
Knowledge stored in a latent form as synaptic efficacies remains inaccessible until it is used to recreate evoked patterns of neural firing. This constraint may explain instances of implicit learning, and why we do not have conscious access to most of our mental algorithms. In the few cases where we do (e.g. when we describe the steps needed for long division), the model predicts that each step should be explicitly coded in the firing of workspace neurons. Indeed, experimentally, prefrontal neurons coding for intentions, plans, ordinal steps, evaluations, intermediate decisions, and errors have been identified.

(ii) Information is not represented in explicit firing form [52]
For an aspect of the visual scene to be consciously accessible, the information might remain permanently non-conscious for at least three reasons [6]:

(iii) Information is coded by neurons functionally disconnected from the workspace
Even information in explicit firing form can remain non-conscious if the relevant neurons lack the bidirectional projections appropriate to establish a reverberant assembly with parietal and prefrontal cortices. This functional disconnection hypothesis might explain why we have no conscious access to the state of activity of subcortical systems sustaining basic maintenance processes (respiration, ingestion, posture, etc). Patients with white matter lesions, including callosal lesions, can also lose conscious access to word, color or object information that is still extracted, yet functionally disconnected.

<table>
<thead>
<tr>
<th>Information encoding</th>
<th>Main features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-conscious</td>
<td>Information is encoded in latent form as matrices of synaptic weights</td>
</tr>
<tr>
<td></td>
<td>Information is encoded in the distributed firing of many neurons, not condensed in small specialized groups of neurons</td>
</tr>
<tr>
<td></td>
<td>Information is encoded in the firing of neurons functionally disconnected from the workspace</td>
</tr>
<tr>
<td></td>
<td>Processing is confined to a brief travelling pulse of firing</td>
</tr>
<tr>
<td></td>
<td>Processing involves local resonant firing loops, but top-down attention is focussed on another stimulus or task set.</td>
</tr>
<tr>
<td>Conscious processing</td>
<td>Processing receives top-down amplification and expands into a global parieto-frontal reverberant state.</td>
</tr>
</tbody>
</table>

Box 3. Questions for further research
- Can one design attentional blink, psychological refractory period, and partial report paradigms for non-human primates? Can they be used to disrupt the neural mechanisms of the ‘preconscious buffer’? Does this preconscious state engage solely local occipito-temporal loops?
- Do all demonstrations of non-conscious information processing in humans fall into one of the categories of the proposed taxonomy? In particular, can one identify model cases where the non-conscious information is demonstrably encoded in synaptic weights, or in neural systems functionally disconnected from parieto-frontal areas?
- Can transcranial magnetic stimulation (TMS) be used to disrupt parieto-frontal circuits and probe their causal involvement in conscious visual perception? Would occipito-temporal TMS similarly disrupt the preconscious buffer during the psychological refractory period?
- Can one find experimental means of testing whether any subjective content is associated with preconscious states? Or is the existence of non-reportable conscious states untestable by definition?
- Can better non-verbal methods be developed for the quantification of introspective reports, both in humans and in non-human primates?
- What type of neural activity patterns underlies introspective reports, as opposed to other more direct sensory-motor decisions?
Box 4. ‘Phenomenal consciousness’ without reportability?

Following Weiskrantz [53], we consider that subjective reports are the primary criterion that can establish whether a percept is conscious or not. Such reports need not be verbal. Many neuroimaging experiments rely on manual reports of conscious perception, which can be made more precise by the use of a continuous visibility measure [23,34].

The philosopher Ned Block, however, has suggested that the reportability criterion underestimates conscious contents [42]. When we view a complex visual scene, we experience a richness of content that seems to go beyond what we can report. This intuition led Block to propose a distinct state of ‘phenomenal consciousness’ prior to global access. This proposal receives an apparent confirmation in Sperling's iconic memory paradigm. When an array of letters is flashed, viewers claim to see the whole array, although they can later report only one subsequently cued row or column. One might conclude that the initial processing of the array, before attentional selection of a row or column, is already phenomenally conscious [3,42].

However, these intuitions are questionable, because viewers are known to be over-confident and to suffer from an ‘illusion of seeing’ [54]. The change-blindness paradigm demonstrates this ‘discrepancy between what we see and what we think we see’ [55]. In this paradigm, viewers who claim to perceive an entire visual scene fail to notice when an important element of the scene changes. This suggests that, at any given moment, very little of the scene is actually consciously processed. Interestingly, changes that attract attention or occur at an attended location are immediately detected. Thus, the illusion of seeing might arise because viewers know that they can, at will, orient attention to any location and obtain conscious information from it.

In summary, according to the present terminology, the whole visual scene is processed preconsciously. Subjects report (correctly) that all of its details are visible – but the change-blindness paradigm reveals that they are not necessarily consciously seen.

... contents with globally accessible ones, appear to be productive avenues for scientific research.

References

26 Haynes, J.D. et al. (2005) Visibility reflects dynamic changes of effective connectivity between V1 and fusiform cortex. Neuron 46, 811–821