

ScienceDirect



Toward a computational theory of conscious processing Stanislas Dehaene^{1,2,3,4}, Lucie Charles^{2,3,4}, Jean-Rémi King^{2,3,4} and Sébastien Marti^{2,3,4}

The study of the mechanisms of conscious processing has become a productive area of cognitive neuroscience. Here we review some of the recent behavioral and neuroscience data, with the specific goal of constraining present and future theories of the computations underlying conscious processing. Experimental findings imply that most of the brain's computations can be performed in a non-conscious mode, but that conscious perception is characterized by an amplification, global propagation and integration of brain signals. A comparison of these data with major theoretical proposals suggests that firstly, conscious access must be carefully distinguished from selective attention; secondly, conscious perception may be likened to a non-linear decision that 'ignites' a network of distributed areas; thirdly, information which is selected for conscious perception gains access to additional computations, including temporary maintenance, global sharing, and flexible routing; and finally, measures of the complexity, long-distance correlation and integration of brain signals provide reliable indices of conscious processing, clinically relevant to patients recovering from coma.

Addresses

 ¹ Collège de France, F-75005 Paris, France
 ² Cognitive Neuroimaging Unit, Institut National de la Santé et de la Recherche Médicale, U992, F-91191 Gif/Yvette, France
 ³ NeuroSpin Center, Institute of BioImaging Commissariat à l'Energie Atomique, F-91191 Gif/Yvette, France
 ⁴ Université Paris 11, Orsay, France

Corresponding author: Dehaene, Stanislas (stanislas.dehaene@gmail.com)

Current Opinion in Neurobiology 2014, 25:76-84

This review comes from a themed issue on **Theoretical and** computational neuroscience

Edited by Adrienne Fairhall and Haim Sompolinsky

0959-4388/\$ – see front matter, \odot 2013 Elsevier Ltd. All rights reserved.

http://dx.doi.org/10.1016/j.conb.2013.12.005

Introduction

Consciousness is the only real thing in the world and the greatest mystery of all

Vladimir Nabokov, Bend Sinister (1947)

What brain mechanisms underlie our capacity to become aware of a specific piece of information, while many others remain non-conscious? Considerable empirical and theoretical progress has been made lately in answering this deceivingly simple question. This research gained leverage when it was recognized that visual illusions [1-3] and a great variety of other normal and pathophysiological conditions such sleep, anesthesia, blindsight or hemineglect provided empirical windows into this phenomenon, by providing minimal contrasts between conscious and non-conscious brain states [4]. Here we review the recent advances made possible by this contrastive approach. We specifically focus on how these findings inform present-day theories of conscious processing. At present, there is no accepted computational theory of this function. Our hope is that the present review may point to the key ingredients that will lead to one.

Defining the terms

It is useful to start by separating the diversity of concepts that the everyday term of 'consciousness' can refer to. The *content of consciousness* refers to the specific information that I am aware of at a given moment. For instance, I am currently aware of reading these words, but not of the music playing in the background (until I attend to it). Conscious access is the process by which a piece of information becomes a conscious content. Conscious processing refers to the various operations that can be applied to a conscious content (as when multiplying two numbers mentally). Conscious report is the process by which a conscious content can be described, verbally or by various gestures. Such reportability remains the main criterion for whether a piece of information is or is not conscious: by hypothesis, I can report something if and only if I am aware of it.

A great variety of representations can be consciously accessed, including perceptual states, abstract knowledge, memories, plans, and other internal states (e.g. feelings, confidence, and errors). *Self-consciousness* is a particular instance of conscious access where the conscious 'spotlight' is oriented toward internal states.

The *state of consciousness*, associated with fluctuations in *wakefulness* or *vigilance*, finally, refers to the brain's very ability to entertain a stream of conscious contents. During normal wakefulness, any information may be consciously accessed, but this ability is continuously modulated according to the level of vigilance, and ultimately vanishes during coma, vegetative state, anesthesia or deep sleep. Although this review concentrates primarily on the mechanisms of conscious access and conscious

processing, in a final section, we consider how what has been learned about conscious access in normal subjects generalizes to the detection of the state of consciousness in brain-lesioned patients.

The boundaries of non-conscious processing

To clarify the nature of conscious processing, a first step consists in delineating what it is *not*. Using masking [5], crowding [6], inattention [7] or binocular rivalry [8], images can be presented under conditions such that they remain strictly invisible. Behavioral priming and brain imaging can then reveal how deep these stimuli are processed. Studies of non-conscious processing play an instrumental role in refuting specific theories of consciousness. The logic is simple: if a cognitive computation or neural marker, proposed by some theory to be uniquely associated with conscious processing, can be observed under demonstrably non-conscious conditions, then that theory is severely undermined.

Twenty years of research indicates that subliminal processing can be quite deep. Many cortical areas can be activated by an unseen stimulus, including areas of the visual ventral [9] and dorsal pathways [10]. The brain nonconsciously recognizes the abstract identity of pictures, words and faces [9,11,12[•]], the quantity attached to a number symbol [10,13], the fact that two words are related or synonymous [6,14,15], the emotional meaning of a word [16[•],17], or the reward value of a coin or an arbitrary symbol [18,19,20^{••}].

In recent years, the frontiers of non-conscious processing have been pushed further. For instance, in chess experts, a brief non-conscious flash of a chessboard suffices to determine whether the king is in check [21]. Within the language domain, the grammatical fit of a masked word with the preceding sentence can be determined nonconsciously [22[•]]. Transitive inferences can also be deployed non-consciously: after non-conscious exposure to arbitrary word pairs such as 'winter-red' and 'redcomputer', word association effects generalize to nonadjacent pairs ('winter-computer'), a transitive link mediated by the hippocampus [23]. As another example of high-level computation, the approximate average of four masked numbers can be extracted non-consciously [13]. There is even a suggestion that multi-step operations such 9-5+2 may be mediated non-consciously [24], although this conclusion will require better control over the stimuli and the degree of non-consciousness.

All in all, these findings refute the idea that non-conscious processing stops at an early perceptual level: meaning and value can clearly be assigned non-consciously. There is also considerable evidence that attention can be deployed and enhance processing even if its target remains non-conscious [25–27]. At the brain level, attending to a stimulus and becoming conscious of it have distinct

signatures that occur on distinct trials and at different times [28–30]. For instance, by orthogonally manipulating visibility and attention (using masked images presented at the threshold for conscious perception such that half were visible and half were invisible, and preceding them by valid or invalid attentional cues). Wyart and colleagues [29] found a double dissociation: attention, but not visibility, modulated early occipital activity, while visibility, but not attention, modulated later temporal and parietofrontal activity. Under some circumstances, greater spatial attention may even lead to a reduced visibility [31^{••}]. These findings refute theories that conflate attention and consciousness. William James' classical definition of attention ('the taking possession by the mind, in clear and vivid form, of one out of what seem several simultaneously possible objects or trains of thought') mixes up conscious access proper ('taking possession of the mind') with selection ('one out of several') which can be fully non-conscious. Selective attention may facilitate conscious access, even when the attentional cue comes long after the stimulus is gone [32[•]], but it operates largely non-consciously.

Recent findings also invalidate the idea that the central executive, which controls our strategies and inhibits unwanted behaviors, always operates consciously. A series of experiments with the go/no-go paradigm indicate that an unseen visual cue can trigger inhibitory control circuits in the pre-supplementary motor area and anterior insula [33,34,35°,36]. Error detection [37°,38°°] and task switching [39°,40], which are typical executive functions, can be triggered non-consciously. Even the maintenance of a stimulus in working memory may remain above the chance level for subliminal stimuli [41°] — although this recent finding will need to be reconciled with the more frequent observation that subliminal priming drops to chance level after a second or less [42–44].

Overall, these findings support the view that virtually any cerebral processor may operate in a non-conscious mode. They challenge theories that associate conscious processing with a specific cognitive processor. For instance, the hypothesis that conscious perception coincides with the ability to deploy higher-order thoughts or metacognition (the brain's ability to represent its own knowledge states) [45] does not bode well with evidence that self-monitoring, error detection and confidence assignment partially operate non-consciously [38°,46°,47].

Findings from subliminal research also eliminate some physiological theories of conscious processing. It is now clear that early changes in gamma band power (>30 Hz), once postulated as a marker of consciousness, can be evoked by a non-conscious stimulus $[48^{\bullet\bullet},49^{\bullet\bullet}]$ and do not faithfully track variations in subjective reports [50]. Similarly, the views that recurrent interactions [51,52] and information integration [53,54] are necessary and

sufficient markers of conscious processing, although not directly refuted, are made implausible by empirical findings of non-conscious interactions between frontal and occipital regions [55], non-conscious integration of unseen visual contours [56[•]], unseen objects in an unseen complex visual scene [57], or unseen words in the semantic or syntactic context of other words [14,22[•]]. These operations are slow (\sim 220–260 ms for contour integration, \sim 400 ms for semantic integration), clearly involve integration of multiple sources, and are unlikely to occur in a purely feedforward manner without recurrent interactions, and yet they occur non-consciously. Similarly, serial accumulation of evidence can occur without awareness [58,59].

Conscious access as an accumulation of evidence leading to an all-or-none ignition

What, if anything, remains unique to conscious processing? Although many cognitive operations can be partially launched non-consciously, these operations rarely if ever run to completion in the absence of consciousness. A subliminal stimulus may induce above-chance performance, behavioral priming, and a small amount of brain activity in narrowly defined brain networks, but these measures often increase dramatically as soon as the subject reports seeing the stimulus, especially in high-level areas [46°,60,61°]. Accumulation of evidence has been demonstrated with non-conscious stimuli [59], but only conscious stimuli cross the threshold beyond which an overt strategy can be flexibly deployed [58].

Such findings vindicate the pre-theoretical idea that consciousness possesses a threshold that separates subliminal and supraliminal stimuli (limen is the Latin word for threshold). Several theorists propose that conscious perception occurs when the stimulus allows the accumulation of sufficient sensory evidence to reach a threshold, at which point the brain 'decides' whether it has seen anything, and what it is [62,63]. The mechanisms of conscious access would then be comparable to those of other decisions, involving an accumulation toward a threshold — with the difference that conscious perception would correspond to a global high-level 'decision to engage' many of the brain's internal resources, not just a single effector [63]. The mathematical frameworks of signal detection theory and Bayesian decision making have been used to model subjective reports of seeing in normal subjects and blindsight patients [64,65]. Neural network models have also been proposed for how highorder cortices might accumulate metacognitive evidence about the state of other cortices, rather than about the external world, leading to a confident feeling of seeing [66].

Recurrent thalamo-cortical networks provide a simple and generic implementation of elementary stimulus categorization processes [67–69]. Recurrent NMDA connections impose slow accumulation dynamics and multi-stable 'all-or-none' behavior, whereby the incoming evidence either quickly dies out (corresponding to subliminal processing) or is accumulated and amplified non-linearly into a full-blown state of high-level activity. This global 'ignition' has been proposed as a marker of conscious perception [70]. Indeed, empirically, when stimulus strength is varied, the early stages of non-conscious processing typically show a linear variation in activation, whereas conscious access is often characterized by a late non-linear amplification of activation which invades a distributed set of parietal, prefrontal and cingulate areas [58,60,61, for extensive review, see 70-72,73°,74]. In behavior, perceptual processing is continuous for subliminal stimuli, but becomes categorical when the stimulus is seen [75,76]. In EEG, MEG, and intracranial recordings, conscious stimuli, compared to matched non-conscious ones, induce a late (~300 ms) and sudden increase in slow event-related potentials (inducing a P3 wave on the scalp), gamma power and long-range beta and gamma synchrony [48^{••},49^{••},77]. Specific components such as the error-related negativity evoked after an erroneous motor response also follow this 'all-or-none' non-linear pattern [46,78].

A direct relation between evidence accumulation and conscious visibility was demonstrated in a recent MEG experiment with gratings presented at threshold [79]. The subjective reports of seeing or not-seeing could be predicted on a single trial basis as a sum of gamma power present before the presentation of the stimulus (-300 to -100 ms) and long after it (+250 to +450 ms). Thus, whether a stimulus is detected seemed to be determined by an accumulation of pre-stimulus bias ('prior') and stimulus-evoked activation ('evidence') [see also 80].

Late ignition seems to provide a robust signature of conscious access. The contrast between an early linear variation in brain activity and a very late non-linear ignition has even been observed in 5, 12 and 15-month-old infants [81], leading to the tentative suggestion that infants too enjoy a conscious perception of visual stimuli, albeit at a much slower pace.

It remains debated, however, whether ignition is a unified process or whether it can be decomposed into a series of stages that correspond to pre-conscious, conscious and post-conscious processes [82°]. The P3 wave may partly reflect processes that unfold after conscious access, such as executive attention, working memory updating, or the preparation of a behavioral report. When these processes are eliminated by making the stimulus irrelevant to the current task, its conscious perception may correlate solely with a transient posterior negativity of moderate size, peaking around \sim 300 ms [56°,83], although other studies continue to observe a large and long-lasting effect [84,85].

Conscious processing as global information sharing

Global Neuronal Workspace (GNW) theory [2,86,87] proposes that conscious access stems from a cognitive architecture with an evolved function: the flexible sharing of information throughout the cortex [4]. While non-conscious stimuli are processed in parallel by specialized cortical processors, conscious perception would be needed in order to flexibly route a selected stimulus through a series of non-routine information processing stages. Global information sharing and routing would rely on a set of interconnected high-level cortical regions forming a 'global workspace' and involving primarily the dorsolateral prefrontal cortex, but also additional hubs in inferior parietal cortex, mid-temporal cortex, and precuneus, and now described as forming a 'rich club' network [88,89].

Behavioral research supports this idea in various ways. A subliminal prime often facilitates performance in a single task, but this non-conscious performance drops to chance level when the task requires a series of novel operations that involve 'piping' the output of one process to the input of another [90,91]. Likewise, a series of subliminal

primes can have cumulative non-conscious effects on a behavioral decision, but only conscious primes allow for the development of subsequent serial strategies [58,92,93].

The brain's routing system is capacity-limited, and this feature may explain the frequent failure of conscious perception in a dual-task setting. Conscious processing of a first target T1 causes a bottleneck on the routing of a subsequent target T2, either by dramatically postponing its processing (a phenomenon known as the 'psychological refractory period', PRP) or by preventing its conscious perception altogether ('attentional blink', AB). Recent evidence confirms that PRP and AB are tightly related phenomena that may occur within the same experiment [94]. Like AB, PRP causes a loss of conscious perception: the second target T2 is not only delayed, but also temporarily unperceived, such that its subjective onset is displaced to the moment when T1 processing finishes [95[•]]. The minimal condition for creating these effects is that T1 is consciously perceived [96,97]. These effects have been related to a global parietal and prefrontal network [94,98], and have been partially captured in simulations of spiking neurons [69,99] (Figure 1).



Flexible information routing and conscious processing in large-scale models of the cortex. (left) Original depiction of the Dehaene-Changeux model of a Global Neuronal Workspace [GNW; Ref. 87]. The GNW model proposes that what we subjectively experience as a conscious state is the global availability of the corresponding information. Conscious access would occur when a piece of information enters a distributed network of cortical areas tightly interconnected by long-distance axons, the GNW, which allows its flexible broadcasting to any of the brain's many specialized processors. (right) Spiking-neuron simulation of a flexible routing system [Ref. 99]. While a first stimulus is processed and routed to an arbitrary response (left column, top to bottom), a second stimulus (right column) is also processed perceptually but is then blocked at the routing stage. This model captures in great detail two neuropsychological phenomena, the psychological refractory period and the attentional blink.

In order to be globally shared, conscious information should be represented by a stable and reproducible brain-scale assembly for a minimal duration. This stability criterion was explicitly tested in an fMRI study where brain activity patterns were more reproducible across trials for perceived than for unperceived stimuli [100]. Electro-encephalography and magneto-encephalography confirm that conscious processing causes sustained brain activity, often extending for several hundreds of milliseconds [29,84,85,101,102]. In intracranial recordings, conscious stimuli, but not non-conscious ones, trigger a sustained activation and the formation of a metastable state of long-distance phase synchrony in the beta band [48^{••}]. Nevertheless, a debate remains, as some data suggest that local synchrony and reverberation may suffice for conscious perception [49^{••},103], while other experiments indicate that long-distance synchrony between prefrontal and occipital cortex may exist even under non-conscious conditions [55].

Consciousness as integrated information

According to Information Integration Theory (IIT) [53,54], global synchrony and re-entry may be needed, not just to globally share or broadcast a conscious stimulus, but, more essentially, to create an integrated representation of its various features. A precise mathematical formula (Φ is proposed to quantify the amount of integration of a system composed of multiple parts). High levels of Φ would be indicative of a conscious device (whether biological or artificial). Any system would possess a small amount of Φ and therefore some degree of consciousness (panpsychism). This formal framework is however limited in its ability to make specific behavioral and biological predictions. Indeed, Φ is impossible to compute in practice (only approximations exist [104]). Furthermore, this theory does not offer any neurophysiological mechanisms for why conscious perception follows a non-linear profile or why highly integrative semantic processes can be triggered non-consciously, as reviewed above.

A more modest proposal is that Φ and related quantities provide one of many possible signatures of the state of consciousness [104,105], simply because they reflect the brain's capacity to broadcast information in the global neuronal workspace, and therefore to entertain a ceaseless stream of episodes of conscious access and conscious processing [70]. Experimentally, mathematical measures of the complexity and global integration of brain signals do provide solid markers of the state of consciousness, particularly when contrasting wakefulness with sleep or anesthesia. Intracranial recordings in humans undergoing propofol anesthesia indicate a dramatic and sudden fragmentation of neural activity, which remains locally organized but globally disintegrated [106**, see also 107], possibly because prefrontal cortices are invaded by an alpha-like rhythm [108]. Some of these effects of

anesthetics are captured by a simple neuronal network model [109,110].

Most importantly, integration and long-distance cortical communications provide signatures of residual consciousness that are clinically applicable to patients recovering from coma. From behavior alone, the presence of consciousness may be quite difficult to detect, and functional MRI has revealed that a few patients in apparent vegetative state may, in fact, be fully conscious and 'locked-in' [111]. An exciting study indicates that the complexity of EEG waves evoked by a single TMS pulse to the cortex provides a strictly quantitative measure of the state of consciousness, with a bimodal distribution separating the awake state from sleep, anesthesia, coma or vegetative state [112^{••}]. Similarly, an EEG measure of the amount of information shared by distant cortical sites provides a highly sensitive discrimination of patients in vegetative versus minimally conscious states, regardless of etiology and time elapsed since injury [113^{••}]. Both observations suggest that global cortical communication provides an excellent index of conscious processing, and are in agreement with both GNW and IIT theories.

Conclusion

Consciousness research has truly come of age. Empirically, several candidate markers of conscious processing are now available. Theoretically, we reviewed three specific theoretical proposals that tentatively relate conscious processing, respectively, to global ignition, longdistance broadcasting, and information integration. These ideas are not necessarily incompatible. On the contrary, considerable convergence exists to suggest that firstly, conscious access triggers an all-or-none change in the state of distributed cortical networks; secondly, conscious processing corresponds to a massive cortico-cortical exchange of information, allowing flexible routing and therefore the slow serial performance of novel and arbitrary tasks; and finally, the state of consciousness, that is the brain's very ability to host a ceaseless stream of such all-or-none conscious episodes, rests upon the integrity of long-distance cortico-cortical exchanges, which can be continuously modulated by lesions or anesthetics and is reflected by electrophysiological indices of brain-wide information sharing.

Future research should investigate whether the proposed markers of conscious processing are generic and valid in all conditions, or whether some are more diagnostic than others. Above all, more detailed computational theories, framed as large-scale simulations of spiking neurons, will be needed to understand the conditions of their emergence in experimental recordings.

Acknowledgements

We gratefully acknowledge extensive discussions with Jean-Pierre Changeux, Lionel Naccache, and Aaron Schurger. This work was supported by a Direction Générale de l'Armement grant (to IRK) an European Research Council grant "NeuroConsc" (to SD), and by the European Human Brain Project.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- .. of outstanding interest
- Crick F, Koch C: Some reflections on visual awareness. Cold 1. Spring Harb Symp Quant Biol 1990. 55:953-962
- Dehaene S, Naccache L: Towards a cognitive neuroscience of 2 consciousness: basic evidence and a workspace framework. Cognition 2001, 79:1-37.
- Kim CY, Blake R: Psychophysical magic: rendering the visible 3. 'invisible'. Trends Cogn Sci 2005, 9(8):381-388.
- Baars BJ: A Cognitive Theory of Consciousness. Cambridge, MA: 4. Cambridge University Press; 1989, .
- Kouider S, Dehaene S: Levels of processing during non-5 conscious perception: a critical review of visual masking. Philos Trans R Soc Lond B: Biol Sci 2007, 362:857-875.
- Yeh SL, He S, Cavanagh P: Semantic priming from crowded 6. words. Psychol Sci 2012, 23:608-616.
- 7. Chun MM, Marois R: The dark side of visual attention. Curr Opin Neurobiol 2002, 12:184-189.
- Pitts MA. Britz J: Insights from intermittent binocular rivalry and 8. EEG. Front Hum Neurosci 2011, 5:107.
- Dehaene S, Naccache L, Cohen L, Le Bihan D, Mangin JF, Poline JB, Rivière D: Cerebral mechanisms of word masking 9. and unconscious repetition priming. Nat Neurosci 2001, 4:752-758
- 10. Naccache L, Dehaene S: The priming method: imaging unconscious repetition priming reveals an abstract representation of number in the parietal lobes. Cereb Cortex 2001, 11:966-974.
- 11. Nakamura K, Kuo WJ, Pegado F, Cohen L, Tzeng OJ, Dehaene S: Universal brain systems for recognizing word shapes and handwriting gestures during reading. Proc Natl Acad Sci USA 2012, 109:20762-20767.
- 12. Qiao E, Vinckier F, Szwed M, Naccache L, Valabregue R,
 Dehaene S, Cohen L: Unconsciously deciphering handwriting: subliminal invariance for handwritten words in the visual word form area. *Neuroimage* 2010, **49**:1786-1799.

This study uses subliminal priming and functional MRI to prove that even hard-to-read handwritten words cause repetition enhancement in the visual word form area, suggesting that invisible visual stimuli are processed non-consciously up to a high level of perceptual invariance.

- 13. Van Opstal F, de Lange FP, Dehaene S: Rapid parallel semantic processing of numbers without awareness. Cognition 2011, 120(1):136-147.
- 14. Luck SJ, Vogel EK, Shapiro KL: Word meanings can be accessed but not reported during the attentional blink. Nature 1996. 383:616-618.
- 15. Devlin JT, Jamison HL, Matthews PM, Gonnerman LM: Morphology and the internal structure of words. Proc Natl Acad Sci U S A 2004, **101**:14984-14988.
- Gaillard R, Del Cul A, Naccache L, Vinckier F, Cohen L, Dehaene S: 16. Nonconscious semantic processing of emotional words
- modulates conscious access. Proc Natl Acad Sci U S A 2006, 103:7524-7529

Using subliminal words, this study shows that, during masking, the threshold for conscious perception is lowered firstly, for masked words that convey an emotional meaning (e.g. rape) compared to words that convey a neutral meaning (e.g. cape); and secondly, for masked words that have been previously seen, compared to words that have never been seen before. Thus, subliminal processing is modulated by both emotion and prior knowledge.

- 17. Naccache L, Gaillard R, Adam C, Hasboun D, Clémenceau S, Baulac M, Dehaene S, Cohen L: A direct intracranial record of emotions evoked by subliminal words. Proc Natl Acad Sci USA 2005, 102:7713-7717.
- 18. Schmidt L, Palminteri S, Lafargue G, Pessiglione M: Splitting motivation: unilateral effects of subliminal incentives. Psychol Sci 2010 21:977-983
- Pessiglione M, Petrovic P, Daunizeau J, Palminteri S, Dolan RJ, 19. Frith CD: Subliminal instrumental conditioning demonstrated in the human brain. Neuron 2008, 59:561-567.
- 20. Pessiglione M, Schmidt L, Draganski B, Kalisch R, Lau H, Dolan RJ. Frith CD: How the brain translates money into force: a neuroimaging study of subliminal motivation. Science 2007, 316:904-906

Participants were presented with masked images of coins serving as incentives while they performed a hand-grip force task. Recordings of skin conductance, hand-grip force and brain activity indicated that subliminal incentives modulated the brain's reward and motivation circuits, as well as the participants' behavioral responses.

- 21. Kiesel A, Kunde W, Pohl C, Berner MP, Hoffmann J: Playing chess unconsciously. J Exp Psychol Learn Mem Cogn 2009, 35.292-298
- 22. Batterink L, Neville HJ: The human brain processes syntax in the absence of conscious awareness. J Neurosci 2013, 33:8528-8533

In this study, syntactic violations in a stream of words were made undetectable by a brief auditory distraction. The undetected violations still evoked a early left anterior negativity, indicating that syntactic processing can proceed without awareness. Only detected violations evoked a later P600 response, confirming that late positive event-related potentials are a frequent marker of conscious processing.

- 23. Reber TP, Luechinger R, Boesiger P, Henke K: Unconscious relational inference recruits the hippocampus. J Neurosci 2012, 32:6138-6148.
- 24. Sklar AY, Levy N, Goldstein A, Mandel R, Maril A, Hassin RR: Reading and doing arithmetic nonconsciously. Proc Natl Acad Sci U S A 2012, 109:19614-19619.
- 25. Kentridge RW, Nijboer TC, Heywood CA: Attended but unse visual attention is not sufficient for visual awareness. Neuropsychologia 2008, 46:864-869.
- 26. Kentridge RW, Heywood CA, Weiskrantz L: Spatial attention speeds discrimination without awareness in blindsight. Neuropsychologia 2004, 42:831-835.
- 27. Naccache L, Blandin E, Dehaene S: Unconscious masked priming depends on temporal attention. Psychol Sci 2002, **13**:416-424
- 28. Wyart V, Tallon-Baudry C: Neural dissociation between visual awareness and spatial attention. J Neurosci 2008, 28:2667-2679.
- 29. Wyart V, Dehaene S, Tallon-Baudry C: Early dissociation between neural signatures of endogenous spatial attention and perceptual awareness during visual masking. Front Hum Neurosci 2012, 16:1-14.
- 30. Watanabe M, Cheng K, Murayama Y, Ueno K, Asamizuya T Tanaka K, Logothetis N: Attention but not awareness modulates the BOLD signal in the human V1 during binocular suppression. Science 2011, 334:829-831
- 31. Rahnev D, Maniscalco B, Graves T, Huang E, de Lange FP, Lau H: Attention induces conservative subjective biases in visual ... perception. Nat Neurosci 2011, 14:1513-1515.

When attended and unattended visual stimuli were matched for objective performance, the unattended stimuli were paradoxically rated as being more visible than the attended ones on a subjective visibility task. This study therefore provides one of the clearest piece of evidence for a dissociation between attention and conscious perception.

- Sergent C, Wyart V, Babo-Rebelo M, Cohen L, Naccache L, 32.
- Tallon-Baudry C: Cueing attention after the stimulus is gone can retrospectively trigger conscious perception. Curr Biol 2013, 23:150-155.

This study shows that a visual stimulus presented at threshold shows enhanced detection when it is 'post-cued' by a subsequent attentional cue presented several hundreds of milliseconds after the initial stimulus. The authors conclude that 'this retroperception effect demonstrates that postcued attention can retrospectively trigger the conscious perception of a stimulus that would otherwise have escaped consciousness'.

- 33. van Gaal S, Scholte HS, Lamme VA, Fahrenfort JJ, Ridderinkhof KR: Pre-SMA graymatter density predicts individual differences in action selection in the face of conscious and unconscious response conflict. J Cogn Neurosci 2010, 23:382-390.
- **34.** van Gaal S, Ridderinkhof KR, Scholte HS, Lamme VA: **Unconscious activation of the prefrontal no-go network**. *J Neurosci* 2010, **30**:4143-4150.
- 35. van Gaal S, Lamme VA, Fahrenfort JJ, Ridderinkhof KR:
- Dissociable brain mechanisms underlying the conscious and unconscious control of behavior. J Cogn Neurosci 2010, 23:91-105.

In this experiment, metacontrast masking was used to present subliminal stop signals. These non-conscious signals slowed down the participants' responses and strongly modulated early electrophysiological markers of frontal inhibitory control (N2). However, the subsequent P3 event was much larger on conscious than on non-conscious trials. Together with Refs. [34,36], this research suggests that part, but not all of the prefrontal mechanisms of inhibitory control can be activated non-consciously.

- van Gaal S, Ridderinkhof KR, Fahrenfort JJ, Scholte HS, Lamme VA: Frontal cortex mediates unconsciously triggered inhibitory control. J Neurosci 2008, 28:8053-8062.
- Nieuwenhuis S, Ridderinkhof KR, Blom J, Band GP, Kok A: Error related brain potentials are differentially related to awareness of response errors: evidence from an antisaccade task. *Psychophysiology* 2001, 38:752-760.

This classic study used an occulomotor anti-saccade task to show that the error-related negativity (ERN), a known brain marker of error detection, remained present on trials in which subjects failed to consciously detect their eye-movement errors. Only the later Pe component varied with subjective error awareness. Thus, performance monitoring can occur in the absence of conscious perception.

 38. Logan GD, Crump MJ: Cognitive illusions of authorship reveal
 hierarchical error detection in skilled typists. Science 2010, 330:683-686

While skilled typists were asked to type a target word, a computer program surreptitiously inserted errors or, conversely, corrected the errors made by the subjects. Subjects often did not detect these manipulations and claimed responsibility for the words as they appeared on screen. However, their typing rates revealed post-error slowing after genuine errors, even when they were computer-corrected and therefore were not detected consciously, but not after pseudo-errors inserted by the program. These findings support a hierarchical model of error detection, in which action monitoring may partially unfold non-consciously.

 Lau HC, Passingham RE: Unconscious activation of the
 cognitive control system in the human prefrontal cortex. J Neurosci 2007, 27:5805-5811.

The authors used subliminal priming to evaluate if task switching could be partially launched without awareness. When presented with subliminal cues coding for a switch to the alternative task, subjects were less accurate in performing the current task, and activity in regions associated with the current task decreased while activity in regions associated with the alternative task increased. This study therefore suggests that subliminal cues can reach higher-order stages of cognitive control.

- 40. Reuss H, Kiesel A, Kunde W, Hommel B: Unconscious activation of task sets. Conscious Cogn 2011, 20:556-567.
- 41. Soto D, Mantyla T, Silvanto J: Working memory without
 consciousness. Curr Biol 2011, 21:R912-R913.

This provocative study shows that an undetected masked grating, rated as subjectively invisible, may nevertheless lead to better-than-chance performance in a delayed match-to-sample task. Working memory performance remained above even when the sample appeared after a delay of 5 s, and even when the delay period was briefly interrupted by an irrelevant distractor. The authors conclude that working memory can partially operate without conscious perception.

- 42. Mattler U: Inhibition and decay of motor and nonmotor priming. Percept Psychophys 2005, 67:285-300.
- Greenwald AG, Draine SC, Abrams RL: Three cognitive markers of unconscious semantic activation. Science 1996, 273:1699-1702.

- Dupoux E, de Gardelle V, Kouider S: Subliminal speech perception and auditory streaming. Cognition 2008, 109:267-273.
- 45. Lau H, Rosenthal D: Empirical support for higher-order theories of conscious awareness. *Trends Cogn Sci* 2011, 15:365-373.
- Charles L, Van Opstal F, Marti S, Dehaene S: Distinct brain
 mechanisms for conscious versus subliminal error detection. Neuroimage 2013, 73:80-94.

This study indicates a dissociation between two metacognitive mechanisms of error detection. When performing a forced-choice task with masked stimuli, subjects performed above chance level in guessing whether they made an error or not, even when the stimulus was subjectively invisible. Contrasting with this graded non-conscious metacognitive performance, however, the error-related negative behaved in an allor-none manner and was only present on visible trials. The authors conclude that metacognition can partially operate non-consciously, but that consciousness is required norder to achieve near-certainty of having made an error.

- Kanai R, Walsh V, Tseng CH: Subjective discriminability of invisibility: a framework for distinguishing perceptual and attentional failures of awareness. *Conscious Cogn* 2010, 19(4):1045-1057.
- 48. Gaillard R, Dehaene S, Adam C, Clemenceau S, Hasboun D,
 Baulac M, Cohen L, Naccache L: Converging intracranial markers of conscious access. *PLoS Biol* 2009, 7:e61.

Using intracranial recordings in 10 epilepsy patients during a masking paradigm, the authors report that subliminal stimuli evoked easily detectable early event-related potentials and gamma-band activity that propagate in a feed-forward manner with the first few hundred milliseconds after stimulus onset. Only conscious stimuli, however, evoked late and long-lasting potentials over frontal electrodes as well as back into the visual cortex. Conscious perception was also associated with strong and late increases in gamma-band power, long-distance beta synchrony and bidirectional Grainger causality. All of these markers may index the late ignition of a global neuronal workspace associated with conscious processing.

- 49. Fisch L, Privman E, Ramot M, Harel M, Nir Y, Kipervasser S,
 Andelman F, Neufeld MY, Kramer U, Fried I *et al.*: Neural
- Andelman F, Neufeld MY, Kramer U, Fried I et al.: Neural "ignition": enhanced activation linked to perceptual awareness in human ventral stream visual cortex. Neuron 2009, 64:562-574.

Using electro-corticograms (ECOG) in 11 epilepsy patients, the authors identified several occipito-temporal sites where gamma-band power was selectively evoked by one category of visual stimuli (e.g. pictures of faces). When the stimuli were masked near the visibility threshold, only the pictures reported as visible led to a sharp and long-lasting 'ignition', a strong increase in gamma-band power. Invisible stimuli only led to a modest and short-lived gamma burst (see Figure 6c). Thus, gamma-band ignition, particularly after 200–300 ms, is a candidate marker of conscious perception.

- Aru J, Axmacher N, Do Lam AT, Fell J, Elger CE, Singer W, Melloni L: Local category-specific gamma band responses in the visual cortex do not reflect conscious perception. J Neurosci 2012, 32:14909-14914.
- van Gaal S, Lamme VA: Unconscious high-level information processing: implication for neurobiological theories of consciousness. *Neuroscientist* 2012, 18:287-301.
- 52. Lamme VA: Towards a true neural stance on consciousness. *Trends Cogn Sci* 2006, **10**:494-501.
- 53. Tononi G: An information integration theory of consciousness. BMC Neurosci 2004, 5:42.
- 54. Tononi G: Consciousness as integrated information: a provisional manifesto. *Biol Bull* 2008, **215**:216-242.
- Cohen MX, van Gaal S, Ridderinkhof KR, Lamme VA: Unconscious errors enhance prefrontal-occipital oscillatory synchrony. Front Hum Neurosci 2009, 3:54.
- 56. Pitts MA, Martinez A, Hillyard SA: Visual processing of contour
 patterns under conditions of inattentional blindness. J Cogn Neurosci 2012, 24:287-303.

This careful event-related potential study uses inattentional blindness to present above-threshold stimuli with or without conscious perception. The authors also attempt to separate conscious perception from further stimulus processing by making the perceived stimulus task-relevant or task-irrelevant on distinct blocks. The findings suggest that conscious perception per secorrelates with a late posterior negativity. A P300 component only appears when the visual stimulus is both seen and task-relevant.

- 57. Mudrik L, Breska A, Lamy D, Deouell LY: Integration without awareness. *Psychol Sci* 2011, 22:764-770.
- de Lange FP, van Gaal S, Lamme VA, Dehaene S: How awareness changes the relative weights of evidence during human decision-making. *PLoS Biol* 2011, 9:e203-e1001.
- Vorberg D, Mattler U, Heinecke A, Schmidt T, Schwarzbach J: Different time courses for visual perception and action priming. Proc Natl Acad Sci U S A 2003, 100:6275-6280.
- Del Cul A, Baillet S, Dehaene S: Brain dynamics underlying the nonlinear threshold for access to consciousness. *PLoS Biol* 2007, 5:e260.
- 61. Melloni L, Schwiedrzik CM, Muller N, Rodriguez E, Singer W:
- Expectations change the signatures, timing of electrophysiological correlates of perceptual awareness. *J Neurosci* 2011, **31**:1386-1396. This important event-related potential study extends an earlier study by

This important event-related potential study extends an earlier study by Del Cul *et al.* (2007, Ref. [60]). It replicates the finding that conscious visual perception of brief unexpected stimuli is associated primarily with a late P300 potential, but also shows that the neural correlates of conscious perception shift to an earlier time period (the P2) when the identity of the stimulus is fully expected.

- 62. Dehaene S: Conscious and nonconscious processes: distinct forms of evidence accumulation? In Better Than Conscious? Decision Making, the Human Mind, and Implications For Institutions. Strüngmann Forum Report. Edited by Engel C, Singer W. MIT Press; 2008.
- Shadlen MN: Consciousness as a decision to engage. In Characterizing Consciousness: From Cognition to the Clinic?. Edited by Dehaene S, Christen Y.. Research and Perspectives in NeurosciencesSpringer Verlag; 2011:27-46.
- Lau HC: A higher order Bayesian decision theory of consciousness. Prog Brain Res 2007, 168:35-48.
- 65. Ko Y, Lau H: A detection theoretic explanation of blindsight suggests a link between conscious perception and metacognition. *Philos Trans R Soc Lond B: Biol Sci* 2012, 367:1401-1411.
- Pasquali A, Timmermans B, Cleeremans A: Know thyself: metacognitive networks and measures of consciousness. Cognition 2010, 117:182-190.
- Wong KF, Wang XJ: A recurrent network mechanism of time integration in perceptual decisions. J Neurosci 2006, 26:1314-1328.
- Dehaene S, Sergent C, Changeux JP: A neuronal network model linking subjective reports and objective physiological data during conscious perception. Proc Natl Acad Sci U S A 2003, 100:8520-8525.
- Dehaene S, Changeux JP: Ongoing spontaneous activity controls access to consciousness: a neuronal model for inattentional blindness. *PLoS Biol* 2005, 3:e141.
- Dehaene S, Changeux JP: Experimental and theoretical approaches to conscious processing. Neuron 2011, 70:200-227.
- Sergent C, Baillet S, Dehaene S: Timing of the brain events underlying access to consciousness during the attentional blink. Nat Neurosci 2005, 8:1391-1400.
- Schwiedrzik CM, Ruff CC, Lazar A, Leitner FC, Singer W, Melloni L: Untangling perceptual memory: hysteresis and adaptation map into separate cortical networks. *Cereb Cortex* 2012. [Epub ahead of print].
- 73. Rounis E, Maniscalco B, Rothwell JC, Passingham R, Lau H:
- Theta-burst transcranial magnetic stimulation to the prefrontal cortex impairs metacognitive visual awareness. Cogn Neurosci 2010.

Repetitive transcranial magnetic stimulation was applied to bilateral dorsolateral prefrontal cortex while subjects performed a discrimination

task with masked stimuli. Subjects rated the subjective visibility of each target ('clear' or 'unclear'). TMS reduced subjective visibility and secondorder metacognitive d-prime, while first-order performance remained unaffected. Thus, prefrontal cortex seems to play a prominent role in secondorder subjective visibility judgments, even for simple visual stimuli.

- Quiroga RQ, Mukamel R, Isham EA, Malach R, Fried I: Human single-neuron responses at the threshold of conscious recognition. Proc Natl Acad Sci U S A 2008, 105: 3599-3604.
- de Gardelle V, Charles L, Kouider S: Perceptual awareness and categorical representation of faces: evidence from masked priming. Conscious Cogn 2011, 20:1272-1281.
- de Gardelle V, Kouider S, Sackur J: An oblique illusion modulated by visibility: non-monotonic sensory integration in orientation processing. *J Vis* 2010, 10:6.
- Melloni L, Molina C, Pena M, Torres D, Singer W, Rodriguez E: Synchronization of neural activity across cortical areas correlates with conscious perception. *J Neurosci* 2007, 27:2858-2865.
- Woodman GF: Masked targets trigger event-related potentials indexing shifts of attention but not error detection. *Psychophysiology* 2010, 47(3):410-414.
- Wyart V, Tallon-Baudry C: How ongoing fluctuations in human visual cortex predict perceptual awareness: baseline shift versus decision bias. J Neurosci 2009, 29:8715-8725.
- 80. Summerfield C, Egner T, Greene M, Koechlin E, Mangels J, Hirsch J: Predictive codes for forthcoming perception in the frontal cortex. *Science* 2006, **314**:1311-1314.
- Kouider S, Stahlhut C, Gelskov SV, Barbosa L, Dutat M, de Gardelle V, Christophe A, Dehaene S, Dehaene-Lambertz G: A neural marker of perceptual consciousness in infants. *Science* 2013, 340(6130):376-380.
- 82. Aru J, Bachmann T, Singer W, Melloni L: Distilling the neural
 correlates of consciousness. Neurosci Biobehav Rev 2012, 36:737-746.

This position paper contains an important critique of the contrastive method, used in most empirical studies of conscious processing, and which consists in contrasting behavior and brain activity evoked by conscious versus non-conscious stimuli. The authors argue that the differences between conscious and non-conscious conditions may not solely reflect the genuine neural correlates of conscious processing, but also include the 'prerequisites for and consequences of conscious processing of the particular content'.

- Railo H, Koivisto M: The electrophysiological correlates of stimulus visibility and metacontrast masking. *Conscious Cogn* 2009, 18(3):794-803.
- Wacongne C, Labyt E, van Wassenhove V, Bekinschtein T, Naccache L, Dehaene S: Evidence for a hierarchy of predictions and prediction errors in human cortex. Proc Natl Acad Sci U S A 2011, 108:20754-20759.
- King JR, Faugeras F, Gramfort A, Schurger A, El Karoui I, Sitt JD, Wacongne C, Labyt E, Bekinschtein P, Naccache L et al.: Singletrial decoding of auditory novelty responses facilitates the detection of residual consciousness. *Neuroimage* 2013, 83:726-738.
- 86. Shanahan M, Baars B: Applying global workspace theory to the frame problem. *Cognition* 2005, **98**:157-176.
- Dehaene S, Kerszberg M, Changeux JP: A neuronal model of a global workspace in effortful cognitive tasks. Proc Natl Acad Sci U S A 1998, 95:14529-14534.
- van den Heuvel MP, Kahn RS, Goni J, Sporns O: High-cost, highcapacity backbone for global brain communication. Proc Natl Acad Sci U S A 2012, 109(28):11372-11377.
- van den Heuvel MP, Sporns O: Rich-club organization of the human connectome. J Neurosci 2011, 31:15775-15786.
- 90. Sackur J, Dehaene S: The cognitive architecture for chaining of two mental operations. *Cognition* 2009, **111**:187-211.

- Fan Z, Singh K, Muthukumaraswamy S, Sigman M, Dehaene S, Shapiro K: The cost of serially chaining two cognitive operations. *Psychol Res* 2012, 76:566-578.
- 92. de Lange FP, Jensen O, Dehaene S: Accumulation of evidence during sequential decision making: the importance of top-down factors. *J Neurosci* 2010, **30**:731-738.
- Schwiedrzik CM, Singer W, Melloni L: Subjective and objective learning effects dissociate in space and in time. Proc Natl Acad Sci U S A 2011, 108:4506-4511.
- Marti S, Sigman M, Dehaene S: A shared cortical bottleneck underlying Attentional Blink and Psychological Refractory Period. Neuroimage 2012, 59:2883-2898.
- 95. Marti S, Sackur J, Sigman M, Dehaene S: Mapping
- introspection's blind spot: reconstruction of dual-task phenomenology using quantified introspection. Cognition 2010, 115:303-313.

Although introspection is usually considered with much suspicion in cognitive psychology, this study shows that human subjects can accurately report on the trial-by-trial variations in their own processing time. They only suffer from a specific illusion: in a dual-task setting, when they are distracted by a primary task, their introspection of the onset of a second stimulus is delayed and diffused — a phenomenon analogous to the attentional blink.

- Asplund CL, Todd JJ, Snyder AP, Gilbert CM, Marois R: Surpriseinduced blindness: a stimulus-driven attentional limit to conscious perception. J Exp Psychol Hum Percept Perform 2011, 36:1372-1381.
- Nieuwenstein MR, Potter MC, Theeuwes J: Unmasking the attentional blink. J Exp Psychol Hum Percept Perform 2009, 35:159-169.
- Tombu MN, Asplund CL, Dux PE, Godwin D, Martin JW, Marois R: A unified attentional bottleneck in the human brain. Proc Natl Acad Sci U S A 2011, 108:13426-13431.
- Zylberberg A, Fernandez Slezak D, Roelfsema PR, Dehaene S, Sigman M: The brain's router: a cortical network model of serial processing in the primate brain. *PLoS Comput Biol* 2010, 6:e765.
- 100. Schurger A, Pereira F, Treisman A, Cohen JD: Reproducibility distinguishes conscious from nonconscious neural representations. *Science* 2009, 327:97-99.
- 101. Chennu S, Noreika V, Gueorguiev D, Blenkmann A, Kochen S, Ibanez A, Owen AM, Bekinschtein TA: Expectation and attention in hierarchical auditory prediction. J Neurosci 2013, 33: 11194-11205.
- 102. Bekinschtein TA, Dehaene S, Rohaut B, Tadel F, Cohen L, Naccache L: Neural signature of the conscious processing of auditory regularities. Proc Natl Acad Sci U S A 2009, 106:1672-1677.
- 103. Fahrenfort JJ, Snijders TM, Heinen K, van Gaal S, Scholte HS, Lamme VA: Neuronal integration in visual cortex elevates face category tuning to conscious face perception. Proc Natl Acad Sci U S A 2012, 109:21504-21509.
- 104. Barrett AB, Seth AK: Practical measures of integrated information for time-series data. *PLoS Comput Biol* 2011, 7:e0521001.

- 105. Seth AK, Izhikevich E, Reeke GN, Edelman GM: Theories and measures of consciousness: an extended framework. *Proc Natl Acad Sci U S A* 2006, 103:10799-10804.
- 106 Lewis LD, Weiner VS, Mukamel EA, Donoghue JA, Eskandar EN,
- Madsen JR, Anderson WS, Hochberg LR, Cash SS, Brown EN et al.: Rapid fragmentation of neuronal networks at the onset of propofol-induced unconsciousness. Proc Natl Acad Sci U S A 2012, 109:E3377-E3386.

In this rare study, single-neuron recordings were obtained while three epilepsy patients underwent propofol anesthesia. The loss of consciousness was associated with a sudden onset of slow oscillations (<1 Hz) in the local field potential. Local neuronal networks continued to exhibit spiking activity, but this activity became fragmented in time by the slow oscillations, and decorrelated in space as the oscillations were asynchronous across distant sites. The suggestion is that temporal and spatial disintegration of cortical activity may cause a loss of consciousness.

- 107. Schroter MS, Spoormaker VI, Schorer A, Wohlschlager A, Czisch M, Kochs EF, Zimmer C, Hemmer B, Schneider G, Jordan D et al.: Spatiotemporal reconfiguration of large-scale brain functional networks during propofol-induced loss of consciousness. J Neurosci 2012, 32:12832-12840.
- 108. Supp GG, Siegel M, Hipp JF, Engel AK: Cortical hypersynchrony predicts breakdown of sensory processing during loss of consciousness. Curr Biol 2011, 21:1988-1993.
- 109. Vijayan S, Ching S, Purdon PL, Brown EN, Kopell NJ: Thalamocortical mechanisms for the anteriorization of alpha rhythms during propofol-induced unconsciousness. J Neurosci 2013, 33:11070-11075.
- 110. Ching S, Cimenser A, Purdon PL, Brown EN, Kopell NJ: Thalamocortical model for a propofol-induced alpha-rhythm associated with loss of consciousness. *Proc Natl Acad Sci U S* A 2010, 107:22665-22670.
- 111. Owen AM, Coleman MR, Boly M, Davis MH, Laureys S, Pickard JD: Detecting awareness in the vegetative state. *Science* 2006, 313:1402.
- 112 Casali A, Gosseries O, Rosanova M, Boly M, Sarasso S,
- Casali KR, Casarotto S, Bruno MA, Lauréys S, Tononi G et al.: A theoretically based index of consciousness independent of sensory processing and behavior. Sci Transl Med 2013, 5:198ra105.

This study proposes a quantitative assessment of the state of consciousness. By quantifying the complexity of the event-related response to a TMS pulse, the authors obtain a numerical index that shows a complete bimodal separation of conscious and non-conscious normal subjects (during sleep or anesthesia). When applied to a small group of patients with disorders of consciousness, this index also separates vegetative and minimally conscious patients.

113 King JR, Sitt JD, Faugeras F, Rohaut B, El Karoui I, Cohen L,

Naccache L, Dehaene S: Information sharing in the brain indexes consciousness in non-communicative patients. Curr Biol 2013, 23:1914-1919.

This study introduces a novel mathematical measure, weighted symbolic mutual information, which can accurately quantify the amount of information exchanged between two recording sites. Applied to a large set of high-resolution EEG recordings from patients with disorders of conscious patients, regardless of etiology and delay since insult.