UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

The Role of Top-Down Modulation in Shaping Sensory Processing Across Brain States

Implications for Consciousness

Sikkens, T.; Bosman, C.A.; Olcese, U.

DOI
10.3389/fnsys.2019.00031

Publication date
2019

Document Version
Final published version

Published in
Frontiers in Systems Neuroscience

License
CC BY

Link to publication

Citation for published version (APA):
Sikkens, T., Bosman, C. A., & Olcese, U. (2019). The Role of Top-Down Modulation in Shaping Sensory Processing Across Brain States: Implications for Consciousness. *Frontiers in Systems Neuroscience*, *13*, Article 31. https://doi.org/10.3389/fnsys.2019.00031

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:13 Nov 2026

frontiers in Systems Neuroscience

REVIEW
published: 24 July 2019
doi: 10.3389/fnsys.2019.00031

The Role of Top-Down Modulation in Shaping Sensory Processing Across Brain States: Implications for Consciousness

Tom Sikkens[1,2], *Conrado A. Bosman*[1,2] *and Umberto Olcese*[1,2]*

[1]*Cognitive and Systems Neuroscience Group, Swammerdam Institute for Life Sciences, University of Amsterdam, Amsterdam, Netherlands,* [2]*Research Priority Area Brain and Cognition, University of Amsterdam, Amsterdam, Netherlands*

Top-down, feedback projections account for a large portion of all connections between neurons in the thalamocortical system, yet their precise role remains the subject of much discussion. A large number of studies has focused on investigating how sensory information is transformed across hierarchically-distributed processing stages in a feedforward fashion, and computational models have shown that purely feedforward artificial neural networks can even outperform humans in pattern classification tasks. What is then the functional role of feedback connections? Several key roles have been identified, ranging from attentional modulation to, crucially, conscious perception. Specifically, most of the major theories on consciousness postulate that feedback connections would play an essential role in enabling sensory information to be consciously perceived. Consequently, it follows that their efficacy in modulating target regions should drastically decrease in nonconscious brain states [non-rapid eye movement (REM) sleep, anesthesia] compared to conscious ones (wakefulness), and also in instances when a given sensory stimulus is not perceived compared to when it is. Until recently, however, this prediction could only be tested with correlative experiments, due to the lack of techniques to selectively manipulate and measure the activity of feedback pathways. In this article, we will review the most recent literature on the functions of feedback connections across brain states and based on the presence or absence of perception. We will focus on experiments studying mismatch negativity, a phenomenon which has been hypothesized to rely on top-down modulation but which persists during nonconscious states. While feedback modulation is generally dampened in nonconscious states and enhanced when perception occurs, there are clear deviations from this rule. As we will discuss, this may pose a challenge to most theories of consciousness, and possibly require a change in how the level of consciousness in supposedly nonconscious states is assessed.

Keywords: brain states, sensory processing, top-down modulation, feedback projections, consciousness, mismatch negativity

OPEN ACCESS

Edited by:
Maria V. Sanchez-Vives, August Pi i Sunyer Biomedical Research Institute (IDIBAPS), Spain

Reviewed by:
Roman Shusterman, University of Oregon, United States
Paul Miller, Brandeis University, United States

***Correspondence:**
Umberto Olcese
u.olcese@uva.nl

Received: *01 February 2019*
Accepted: *05 July 2019*
Published: *24 July 2019*

Citation:
Sikkens T, Bosman CA and Olcese U (2019) The Role of Top-Down Modulation in Shaping Sensory Processing Across Brain States: Implications for Consciousness. Front. Syst. Neurosci. 13:31. doi: 10.3389/fnsys.2019.00031

INTRODUCTION

Our brain is capable of sustaining all the functions necessary for life, from the most basic ones (breathing, autonomic regulation) to the most complex ones (language, social behavior). One function which stands out for being at the same time extremely complex to grasp, yet seamless in its presence, is consciousness. Even defining consciousness is extremely difficult, and has been a long-lasting subject of debate for philosophers and cognitive scientists (Dennett, 1991, 2018; Chalmers, 1995; Crick and Koch, 2003; Tononi and Koch, 2015; Dehaene et al., 2017; Storm et al., 2017; Lamme, 2018). In spite of this, in the last two decades significant progress has been made in the scientific study of consciousness, and in particular in the search of the neural correlates of consciousness (NCC, see **Table 1** for a list of abbreviations): the neural signature of brain processes underlying consciousness (Aru et al., 2012; Koch et al., 2016). For scientists to be able to measure the NCC, however, an operational definition of consciousness is necessary.

Chalmers (1995) famously proposed an ontology of cognitive phenomena associated with consciousness—the easy problems of consciousness: mechanisms controlling wakefulness and sleep, the ability to report mental states, the control of behavior, etc. These phenomena—termed easy because experiments to study them may be technically challenging but pose no conceptual difficulty—are inextricably associated with conscious processing, but are not consciousness. What really defines consciousness is the subjective experience that is inextricably associated with all the cognitive phenomena mentioned above. This has been defined by Chalmers as the hard problem of consciousness because of the so-called explanatory gap (Levine, 1983) between neuron-level mechanisms and subjective experience.

TABLE 1 | List of abbreviations.

Abbreviation	Definition
Cg1	Anterior Cingulate Cortex
DD	Deviance Detection
ECoG	Electrocorticogram
EEG	Electroencephalography
ERP	Evoked Response Potential
fMRI	Functional Magnetic Resonance Imaging
GABA	γ-Aminobutyric Acid
GNW	Global Neural Workspace
IIT	Integrated Information Theory
L1–6	Cortical layers 1–6
LFP	Local Field Potential
MCS	Minimally-conscious state
MMN	Mismatch Negativity
MMNr	Mismatch Negativity response
N1	Negative peak in visual ERPs occurring 100 ms after stimulus onset and typically associated to the MMNr
NCC	Neural Correlates of Consciousness
P300	Typical ERP elicited within a decision-making process and during oddball paradigms, with a peak occurring about 300 ms after stimulus onset
P3b	Subcomponent of the P300 ERP which has been linked to aware processing
PC	Predictive Coding
REM	Rapid Eye Movement Sleep
SSA	Stimulus-Specific Adaptation
V1-V4	Visual cortical areas 1, 4
VS	Vegetative State

The ontology first proposed by Chalmers has inspired the development of theories of consciousness aimed at addressing the nature of conscious experience, such as those proposed by Crick and Koch (1990, 2003) and Tononi and Edelman (1998a,b), and, more recently, the Integrated Information Theory (IIT; Tononi, 2004; Tononi et al., 2016). Nevertheless, Chalmer's proposal remains highly controversial and is firmly rejected by some philosophers, chief among them Daniel Dennett (Dennett, 1991, 2018; Cohen and Dennett, 2011). According to Dennett, there is no hard problem of consciousness, but rather consciousness can be understood by studying the functions associated with it. Dennett's approach can be seen as a theoretical foundation for another of the most influential theories of consciousness, the Global Neural Workspace (GNW) theory (Sergent and Dehaene, 2004; Baars, 2005; Dehaene et al., 2017). GNW is a theory of conscious access (Lamme, 2018; Naccache, 2018). What the theory aims to explain are the neuronal mechanisms which allow the brain to access (and subsequently report) information (Baars, 2002). This ultimately corresponds to an attempt to understand consciousness *via* addressing one of Chalmer's easy problems, which in Dennett's framework is all that is needed to uncover the mystery of consciousness.

The debate between these (and other) philosophical frameworks to explain consciousness, and between the different theoretical models to explain why the brain—and possibly artificial systems—is conscious is lively and at times heated, as attested by the vast recent literature on the topic (Boly et al., 2017; Odegaard et al., 2017; Dennett, 2018; Lamme, 2018; Naccache, 2018; Olcese, 2018; Olcese et al., 2018b). Nevertheless, in spite of the fundamental differences between the various frameworks to study consciousness, most theories aimed at explaining the neural mechanisms of consciousness agree on a key ingredient which makes our brain conscious: feedback connectivity. Feedforward processing consists in a progressive processing and transmission of sensory information from sensory organs up to higher-order cortical and motor regions (Lamme and Roelfsema, 2000; Lamme, 2018). While very complex forms of processing can be achieved by purely feedforward networks, as exemplified by the performance of deep artificial neural networks in the field of computer vision (LeCun et al., 2015), this is commonly agreed to occur non-consciously (Lamme, 2018). For example, both IIT and GNW concur that feedforward processing is *per se* not conscious. In IIT this can even be quantified, and purely feedforward networks (such as deep artificial neural networks) achieve a Φ value of 0 (Oizumi et al., 2014), where Φ quantifies the "level" of consciousness in IIT. Recurrent processing is, conversely, seen as a pre-requisite for consciousness by most if not all modern theories of consciousness. In IIT, as previously said, only systems with some level of integration (i.e., recursive connections) possess $\Phi > 0$, and therefore can be considered conscious (Oizumi et al., 2014). In GNW, the feedback flow of information from frontal to posterior brain regions gives rise to the so-called global ignition which is seen as essential for consciousness

(Sergent and Dehaene, 2004; Baars, 2005). Similarly, other theories, such as the recurrent processing hypothesis (Lamme, 2006, 2018), and the Predictive Coding (PC) framework (Pennartz, 2015) indicate feedback processing as a key ingredient for consciousness.

The neuroscientific evidence underlying the link between recurrent processing and consciousness has been until recently primarily limited to correlational studies done at the mesoscopic level in human subjects, *via* techniques such as electroencephalography (EEG) and functional magnetic resonance imaging (fMRI; van Gaal et al., 2011; Fahrenfort et al., 2012). These techniques, albeit powerful, lack cell-level resolution. Thus, while experiments done using EEG and fMRI can provide evidence about a generalized increase or decrease in patterns of neural activity which are compatible with modulation in feedback/recurrent coupling, they cannot discriminate what individual neurons and neuronal populations do. As an example, EEG experiments have shown that, during brain states characterized by the loss of consciousness [Non-rapid eye movement (REM) sleep, anesthesia, coma], cortical effective connectivity drops markedly (Massimini et al., 2005, 2012; Ferrarelli et al., 2010; Casarotto et al., 2016). This was thought to reflect a generalized drop in the communication between cortical areas. Recently, however, studies performed at cellular resolution in rats showed that specific forms of long-range connectivity are preserved or even enhanced in Non-REM sleep compared to wakefulness (Olcese et al., 2016, 2018a). Therefore, human studies can only give limited insight into the role of recurrent connections in conscious processing, as only average, area-level dynamics can be assessed.

Our aim is to provide an overview of the existing evidence on the role of feedback processing as a key constituent of the NCC, as provided by studies with neuron-level resolution performed in animal models. In particular, we will focus on one of the easy problems of consciousness, as defined by Chalmers (1995): the difference between wakefulness and sleep/anesthesia. We will specifically address how top-down, intra-cortical feedback varies between states of consciousness (wakefulness vs. sleep and anesthesia) and thus investigate whether the presence or absence of this form of neural dynamics can be considered a valid NCC. This will allow us to dwell into the vast literature on cortical processing in animal models, and to focus on the neocortex, i.e., the brain region which is thought to be crucial for consciousness (Koch et al., 2016). By primarily centering on differences between conscious and non-conscious brain states, we will be able to address to what extent feedback communication decreases when consciousness fades. In the first section of this manuscript, we will discuss the circuit-level mechanisms of feedforward and feedback cortical communication, and what has been reported in terms of variation across behavioral states. Next, we will zoom in on studies investigating the genesis of mismatch responses. This strongly preserved phenomenon has been shown to occur during both conscious and non-conscious states and is composed of two different components: stimulus-specific adaptation (SSA) and deviance detection (DD; Näätänen et al., 2007; Garrido et al., 2009b; Hamm and Yuste, 2016; Harms et al., 2016). These two components are, respectively, classically thought to represent feedforward and feedback forms of processing. The analysis of how these components vary across behavioral state will allow us to assess if and how feedback processing is disrupted during loss of consciousness.

THE ROLE OF FEEDBACK PROCESSING ACROSS DIFFERENT BRAIN STATES

Brain state transitions can be described as the result of dynamic changes in functional long-range neuronal networks across time. Local and global connectivity greatly differ in their dynamics during the establishment of different behavioral states. While local connectivity is less affected by brain state transitions (Townsend et al., 2015; Olcese et al., 2016, 2018b; Siclari and Tononi, 2017), behavioral states exert great influences over the modulation of long-range connections (Destexhe et al., 1999; Massimini et al., 2005). In particular, a reduction in long-range connectivity has been associated with the loss in consciousness occurring in NREM sleep and anesthesia (Olcese et al., 2016; Storm et al., 2017). During wakefulness, local activity elicited within cortical microcircuits is usually broadcasted towards different brain areas through large-scale networks interconnected through feedforward and feedback projections, enabling sustained activity. This recurrent activity—necessarily supported by feedback, reentrant connections—is acknowledged to be at the basis of the NCC by several theories of consciousness (Koch et al., 2016; Lamme, 2018). In this section, we review the contribution of large-scale feedback networks in the emergence of conscious brain states, and how these networks influence the local activity of cortical microcircuits across different states.

In contrast with the predominant role recurrent connectivity plays in many of the aforementioned theories of consciousness, classical theories about sensory processing have mostly relied on the role of feedforward projections to explain how sensory stimulation can be processed (Hubel and Wiesel, 1977). Under this framework, simple stimulus features are encoded by neurons located in early sensory areas. Subsequently, this information is transferred *via* feedforward projections towards hierarchically higher areas. In these higher areas, the cumulative processing of incoming stimuli can be finally integrated into one coherent perception (Riesenhuber and Poggio, 1999). However, several cognitive processes, such as sensory perception, attention and goal-oriented behavior, have been shown to require feedback modulation—see e.g., Gazzaley and Nobre (2012), Gilbert and Li (2013), Manita et al. (2015) and Kwon et al. (2016). Furthermore, a pure feedforward approach falls short to explain how percepts—processed in a hierarchical feedforward fashion—can be accessible to conscious experience (Lamme and Roelfsema, 2000). This has been extensively investigated at the mesoscopic scale in humans (Fahrenfort et al., 2007, 2008, 2012), but the underlying circuit-level dynamics are less well understood. In a recent study investigating the neuronal responses in V1, V4, and prefrontal cortex of awake macaques during near-threshold stimulation, strong stimuli—those invariably leading to a conscious report—elicited a strong neuronal response progressing through all recorded

areas, in a way compatible with the notion of inter-areal feedforward propagation (van Vugt et al., 2018). However, in case of weaker, near-threshold stimuli, signal propagation across multiple areas was highly influenced by the global pre-stimulus brain state of the animal. In detail, a computational modeling approach was used by the authors of the study to show that recurrent connectivity between hierarchically-organized areas and feedback connectivity towards sensory cortices are necessary to be able to predict neuronal responses occurring in early processing stages when stimulus detection is reported (van Vugt et al., 2018). These findings strongly suggest that feedforward activity *per se*—albeit important for perceptual processing—is not sufficient to elicit a conscious experience of a perceived stimulus. Conversely, recurrent connectivity seems to be required to support conscious responses—see also Boly et al. (2011). Nevertheless, the precise contribution of this feedback connectivity to the establishment of these networks and the mechanisms underlying local cortical circuitry modulation remains to be elucidated.

Anatomical Organization of Recurrent Neuronal Networks

To understand what precise feedback connectivity mechanisms are contributing to consciousness, it is important to revisit some anatomical principles that govern brain connectivity in general. Different anatomical studies have estimated that a vast majority of cortical connections are essentially local (Markov et al., 2014a), organized in evolutionarily conserved microcircuits (Bosman and Aboitiz, 2015). This local architecture represents a basic organizational unit for cortical computations (Douglas and Martin, 2004; Womelsdorf et al., 2014) and, as we discuss below, the way distinct local circuits communicate with each other is tightly linked to the emergence of different brain states.

Cortical microcircuits follow a columnar organization spread across all cortical layers. The prototypical microcircuit in sensory cortices features a central granular layer (L4), receiving projections from the thalamus and other cortical layers—but see Constantinople and Bruno (2013). Supragranular layers 2 and 3 (L2/3) receive presynaptic inputs from L4, while layer 1 (L1) displays cortico-cortical fibers connecting neighborhood columns. Infragranular layers include layers 5 and 6 (L5 and L6). Neurons in these layers receive local inputs from collateral projections of L2/3 neurons and provide feedback and feedforward connectivity to different thalamic structures (Hubel and Wiesel, 1977; Shepherd, 2011; Harris and Shepherd, 2015). Both feedforward and feedback projections have a strong preference for targeting areas that are close to their origin (Markov et al., 2014b). Supporting this highly redundant local connectivity, an estimated amount of 20% of the total connectivity is organized through long-range interareal connections (Markov et al., 2014a). These projections follow a well-defined pattern of connectivity (Felleman and Van Essen, 1991; Markov et al., 2014a,b). Feedforward projections preferentially originate in supragranular layers and target granular layers. Conversely, feedback connections arise from infragranular layers and avoid targeting granular layers (Felleman and Van Essen, 1991).

The quantification of this pattern of connectivity (Ercsey-Ravasz et al., 2013; Markov et al., 2014b) has suggested an important structural heterogeneity across brain areas, arranged in what Kennedy and colleagues have denominated a Bow-Tie organization (Markov et al., 2013b). Under such an organization, a subset of association areas including frontal, parietal, and temporal cortices appear to be heavily interconnected *via* recurrent connections, forming a central core of brain areas (**Figure 1A**). This highly interconnected core is characterized by a high prevalence of long-distance connections departing from

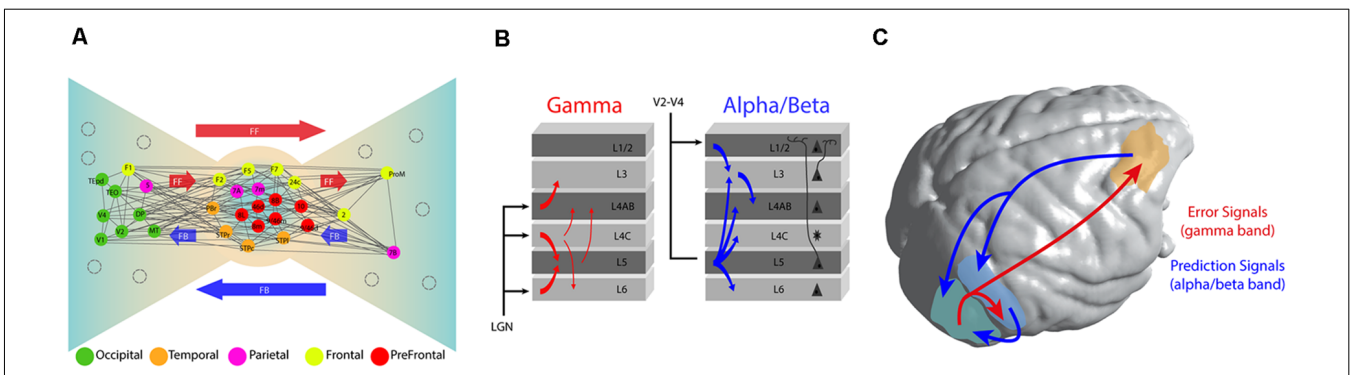

FIGURE 1 | Functional architecture of cortical connectivity. **(A)** Brain topology as a Bow-Tie organization. This brain organization arises from the connectivity profile obtained from retrograde tracer studies. It defines a heavily interconnected core of fronto-parietal areas and the ties based in feedforward (FF) and feedback (FB) connections from the core. As expected, early sensory areas are located more at the extremes of the tie. Adapted with permission from Markov et al. (2013b). **(B)** Representation of the directed oscillatory influences (measured by Granger causality) over the cortical column in area V1 of the monkey. Gamma (30–90 Hz) oscillations are predominantly feedforward and they target granular and supragranular layers. Alpha/beta oscillations (roughly 8–30 Hz, combined) are predominantly feedback. They originate at hierarchically higher areas and target predominantly infragranular layers. Adapted with permission from van Kerkoerle et al. (2014). **(C)** Long and short-range connections implementing a Predictive Coding (PC) framework of sensory processing. Under such framework, a predictive model is originated at prefrontal cortex and local areas and channeled towards early sensory areas by long and short-range connections respectively, using low frequency (alpha and beta) oscillations. Error signals, originated from sensory areas are broadcasted towards hierarchically higher brain areas using high-frequency oscillations.

this core to other brain regions (Markov et al., 2013a). At the periphery, two different sets of regions are linked to the core through feedback and feedforward connections: a first group composed of primary and secondary sensory regions, and a second group consisting of premotor areas. These two groups are interconnected from the first to the second group *via* direct feedforward connections, and from the second to the first one *via* direct feedback connections (Ercsey-Ravasz et al., 2013).

This organizational principle encompasses several advantages. It reduces brain wiring and volumetric surface, yet increases the efficiency of information transfer and computational speed (Markov et al., 2013b). It reveals the importance of a fronto-parietal network, fundamentally characterized by a high prevalence of long-range connections, which some authors have suggested to be compatible with a GNW organization (Ercsey-Ravasz et al., 2013; Markov et al., 2013a), and provides an evolutionary advantage for long-range communication and cortical coordination of brain dynamics (Bosman and Aboitiz, 2015).

Functional Dynamics of Feedback and Feedforward Connections: A Possible Role of Brain Oscillations

A bow-tie brain topology strongly relies on the functionality of long-range feedback connections to support sensory processing. As stated before, classical (and mainly feedforward) theories about perception are insufficient to explain sensory processing under such organizational principles. Conversely, the PC framework offers important insights into sensory perception in a view that is compatible with the notion of neuronal entrainment. PC is rooted in the tradition of inferential models of brain perception (Friston, 2010; Clark, 2013). Under this formulation, feedback projections transmit to hierarchically lower areas a generative model of sensory perception. In turn, feedforward projections transfer a signal error from the model, derived from the comparison between the existing model and the incoming sensory signals (Rao and Ballard, 1999; Bastos et al., 2012). Accordingly, recurrent (i.e., local) interactions between models and error signals at the level of cortical microcircuits are thought to improve the generation of statistical inferences about sensory perception. It has been argued that cortical microcircuit architectures can effectively implement PC computations. Under such scenario, local cortical microcircuits effectively integrate data and models originated from long-range feedforward and feedback networks, respectively (Bastos et al., 2012, 2015; Fontolan et al., 2014; Chao et al., 2018). Nevertheless, it is important to note that a PC framework can infer statistical regularities, but it cannot specify why data from different types of sensors would be consciously experienced (Pennartz, 2009). In other words, a PC framework helps to explain the importance of feedback connectivity but does not offer a solution to the hard problem of consciousness.

Experimentally, the study of neuronal dynamics underlying cortical computations across areas requires the utilization of techniques able to record several areas and spatial levels simultaneously. From a myriad of emerging techniques (Adesnik and Naka, 2018), ensemble recordings and surface local field potentials (LFPs), recorded—respectively—by high-density laminar probes and electrocorticograms (ECoGs), are considered a primary choice due to their capacity to simultaneously record multiple local neuronal assemblies through cortical layers (Lewis et al., 2015; Pesaran et al., 2018). Laminar recordings in early visual areas have shown that LFP oscillations—a prominent feature of field recordings which has been associated with several perceptual and cognitive functions (Bosman et al., 2014; Fries, 2015)—are compartmentalized across cortical layers (**Figure 1B**). High-frequency oscillations (e.g., in the gamma frequency band, between 30–90 Hz) are observed mostly in supragranular layers. Conversely, lower frequency bands, such as alpha (8–12 Hz) and beta (13–30 Hz), are observed mostly in infragranular layers (von Stein et al., 2000; Buffalo et al., 2011; van Kerkoerle et al., 2014).

Brain rhythm compartmentalization appears to have a functional role during inter-areal communication (Fries, 2015). Recent studies enabling simultaneous recordings across multiple brain areas have demonstrated that directed long-range interactions can be exerted across different frequency band channels, a scenario compatible with the PC framework (Bastos et al., 2012, 2015; Fontolan et al., 2014; Michalareas et al., 2016; Chao et al., 2018). A study using high-density laminar profile in V1 of awake monkeys has shown that, while gamma band responses initiate in L4 and propagate through supra and infragranular layers of higher visual areas, alpha oscillations travel in the opposite direction (van Kerkoerle et al., 2014, **Figure 1B**). These relationships between frequencies and cortical layers are consistently observed through the cortical hierarchy. Using ECoGs in awake monkeys, Bastos et al. (2015) showed that, while feedforward directed influences (measured in terms of Granger causality) are observed through gamma oscillations across eight cortical regions recorded simultaneously, feedback influences across the visual hierarchy are consistently carried out by beta oscillations, and a similar functional connectivity pattern has been observed in humans (Michalareas et al., 2016). Furthermore, new studies have shown that feedback modulation can be updated continuously using error signals broadcasted from sensory areas to prefrontal cortex (Chao et al., 2018; **Figure 1C**). Neurons located in prefrontal cortex have the ability to integrate such error signals and continuously submit updated versions of the model through feedback projections to temporal areas (Chao et al., 2018). In early sensory cortices, the efficiency of feedforward frequency coupling—measured as an increment of gamma frequency band coherence between V1 and V4—increases after augmented top-down beta frequency band modulation (Richter et al., 2017). Importantly, optimized feedforward efficiency improves conscious access of perceived stimuli (Lamme and Roelfsema, 2000; Boly et al., 2011), and behavioral responses during attentional tasks (Rohenkohl et al., 2018), indicating that the extent of long-range feedback modulation correlates with conscious accessibility and enhanced behavioral performance.

Thus, simultaneous recordings across multiple areas using LFP rhythmic fluctuations have consistently shown that long-range feedback signals, possibly originating from

highly interconnected hubs comprising anterior regions of the brain, effectively modulate the activity of early sensory regions. This feedback modulation ultimately facilitates feedforward communication, behavioral performance and conscious reporting of perceived stimuli. A PC architecture, implemented across local-to-global anatomical connections and dynamical LFP oscillatory phase relationships, may support this organization (Olcese et al., 2018b).

However, while several studies have shown the importance of PC architectures and their oscillatory dynamics during stimulus processing, the exact contribution of these processes during different brain states still needs to be elucidated (Olcese et al., 2018b). The observation of local and global network dynamics across wide-brain areas across different behavioral states is a crucial step to understand what the mechanisms underlying these behavioral and network transitions are. Feedforward gamma connectivity globally decreases during deep sleep states, as compared to awake states, in intracranial recordings in epilepsy patients (Mikulan et al., 2018), but there are no consistent reports about low-frequency (and putatively feedback) phase relationships during sleep, in spite of the reduced role of feedback during non-conscious states (Boly et al., 2011). During anesthesia, a pharmacologically-induced brain state, quantitative EEG studies have shown that sedation with propofol is accompanied by a decreased posterior alpha and increased frontal/central beta power (Gugino et al., 2001; Akeju and Brown, 2017). This shift has been extensively studied and is thought to emerge from a disruption of prefrontal circuits created by a strong low-frequency thalamocortical synchronization (Vijayan et al., 2013; Flores et al., 2017). Yet, the detailed effects of anesthesia on cortico-cortical synchronization remain elusive. Finally, a promising line of research about the effects of feedback connectivity in patients with consciousness disorders has been developed in recent years. Feedback influences, but not feedforward ones, seem to be compromised in these patients (Boly et al., 2011), in agreement with the notion that feedback connections are important to sustain conscious activity. Intriguingly, a temporary recovery in the consciousness state of a subgroup of patients has been observed following the administration of the GABA modulator Zolpidem (Hall et al., 2010; Williams et al., 2013). Zolpidem administration reduces EEG power and coherence at 6–10 Hz frequencies (Williams et al., 2013). Surprisingly, however, Zolpidem seems to elevate beta frequency power at anterior regions following its administration. These effects seem to be mediated by thalamocortical interactions but is not clear whether direct cortico-cortical communication is also affected (Hall et al., 2010; Williams et al., 2013). Future studies might help to unveil the exact nature of these effects.

In conclusion, long-range feedback connections are an essential component of conscious brain states and determine how sensory information is processed at the local level (*via* both feedforward and recurrent mechanisms). Yet, the underlying circuit-level mechanisms remain poorly understood. Several studies have outlined the organizational principles—both anatomical and functional—of feedback connections in the context of sensory processing. However, contrasting results have been presented on how feedforward and feedback functional connectivity vary as a function of the state of consciousness. Oscillatory dynamics during sensory processing may provide a framework to understand the role of feedback projections, but these phenomena involve a wide range of brain circuits, operating both at feedforward and feedback level and both at local and global scales. Consequently, one possible solution to understand the role of feedback projections and the underlying circuit mechanisms in conscious processing is to focus on processes that are purely feedback in nature, and which are thought to be exclusively dependent on long-range, top-down mechanisms.

MISMATCH NEGATIVITY: PRESERVED TOP-DOWN MODULATION DURING NON-CONSCIOUS BRAIN STATES?

The mismatch negativity response (MMNr) is a well-studied electrophysiological phenomenon that occurs in the human brain after violation of a rule, which is established by a sequence of repeated stimuli (Näätänen et al., 2007; Garrido et al., 2009b). The MMNr reflects the brain's capacity to automatically detect unpredicted sensory changes in our environment, without the need for attention (Tiitinen et al., 1994; Näätänen et al., 2001; Stefanics et al., 2014), and is impaired in several psychological afflictions such as schizophrenia, attention-deficit hyperactive disorder and psychosis (Erickson et al., 2016; Näätänen et al., 2016). In recent years the interest for the MMNr has grown considerably, especially in relation to the hierarchical PC framework, bringing a new plethora of insights from both human and rodent experiments.

The MMNr is often studied using an oddball paradigm, in which a particular stimulus (standard) is repeatedly presented to a subject, but sometimes unpredictably changed to a different, unexpected stimulus (deviant, see **Figure 2A**). The idea behind this paradigm is that the repetition of a particular stimulus or event leads to the formation of a prediction related to the frequency or probability at which this event occurs. Stimuli that deviate from the frequently presented one elicit a strong mismatch response. Thus, the MMNr can be interpreted as a prediction error to a stimulus that does not match the statistical regularities of sensory stimuli being perceived, thus updating our internal representation of the world (Friston, 2005; Garrido et al., 2009b; Stefanics et al., 2014).

Different Components of the Mismatch Negativity Response

The MMNr can be split into two functionally distinct components (Hamm and Yuste, 2016; Harms et al., 2016). The first component arises as an effect of stimulus repetition. Repetition of a particular event leads to a decreased neuronal response to that event, a phenomenon often referred to as repetition suppression or SSA. SSA has been found in many different neural systems and at various levels: from single

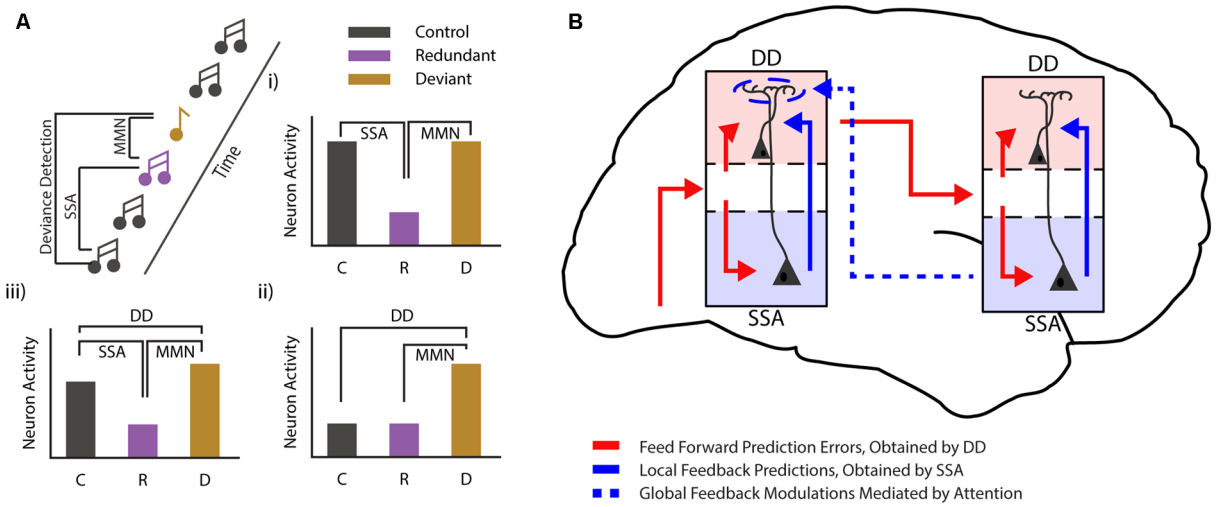

FIGURE 2 | Mismatch negativity during different states of consciousness. **(A)** Cartoon showing the commonly used auditory oddball paradigm and the different comparisons that can be used to differentiate between stimulus-specific adaptation (SSA) and deviance detection (DD). Three bar graphs (i—iii) show the hypothetical cases in which classical MMN van be observed. In one case (i) MMN can be fully explained by SSA. Another example (ii) shows a pure deviance detecting neuron, while the final example (iii) shows MMN as it is most likely found in awake subjects, with both SSA and DD. **(B)** Schematic showing how local DD may be maintained during loss of consciousness, while global predictions are lost. Global, long-range feedback projections are reduced during loss of consciousness, while local connectivity is maintained. This suggests that low-level predictions might arise from local connectivity (for example from deep to superficial layers), while more complex, global prediction requires top-down modulation. These top-down projections mostly target interneurons in the most superficial layer of the cortex, where they are in a prime position to modulate neurons from both the superficial and the deep layers of the cortex.

neurons in the cortex of primates (Miller and Desimone, 1994), cats (Ulanovsky et al., 2003) and rodents (Taaseh et al., 2011), to field and imaging recordings in rodents (Chen et al., 2015; Hamm and Yuste, 2016; Parras et al., 2017; Hamm et al., 2018) and humans (Garrido et al., 2009a). While SSA closely resembles the MMNr originally described in humans (Ulanovsky et al., 2003; Nelken and Ulanovsky, 2007) it does not inherently have the second crucial component of the MMNr, true deviance detection (DD). For DD to occur the increased response to the deviant compared to the standard stimulus should not be fully explained by the decrease in neuronal firing caused by SSA (i.e., it should not simply be equal to the response elicited when the standard stimulus is not presented in a sequence). Rather, the increase in this response should have added value in detecting the occurrence of a deviant or irregular event (Hamm and Yuste, 2016; Harms et al., 2016). In this paragraph, we will discuss the relationship between SSA and DD, and how they are related to the MMNr.

Repetition of a particular stimulus leads to neuronal adaptation in the response to this stimulus, which can be characterized by the decrease of the neuronal response over repetitions. There are several hypotheses on which neuronal mechanisms could underlie this adaptation. Some models suggest that relatively simple feedforward mechanisms, such as neuronal fatigue, may fully explain the effects of stimulus repetition (Grill-Spector et al., 2006). Indeed, it is possible that SSA can be explained, at least in part, by purely feedforward mechanisms (Garrido et al., 2009a; Farley et al., 2010; May and Tiitinen, 2010), where (synapses of) neurons that are responsive to the standard stimulus adapt over time, resulting in responsive depression, while the neurons responding to the deviant stimulus are "fresh" and are activated in full. A more likely case, however, is that multiple mechanisms contribute to repetition suppression under different conditions (Grill-Spector et al., 2006).

Another possible view on SSA comes from the PC framework. According to this framework, the decrease in neuronal activation upon repeated stimulation reflects the brain's ability to predict that stimulus (Friston, 2005, 2008, 2010). This is achieved by neuronal processes optimized to probabilistically represent causes of sensory inputs. These processes can be seen as the building of an internal model representing our external environment. Keeping this model up-to-date requires a constant interaction between top-down predictions and bottom-up prediction errors (Bastos et al., 2012). An increased ability of this generative model to predict sensory inputs is reflected in a decrease of neuronal activity necessary to update the model. In other words, SSA can be considered as a product of perceptual learning where the predicted part of a sensory input is "explained away."

DD, on the other hand, is an increased neuronal response to deviant stimuli, and can be considered a neural correlate of error signaling or memory update, where the prediction does not match the actual input and an adjustment of the generative model is called for.

Though SSA has been extensively studied in animal models, until recently it remained unclear whether MMN and more specifically DD could also be found in rodents (Taaseh et al., 2011; Chen et al., 2015; Hamm and Yuste, 2016;

Parras et al., 2017; Hamm et al., 2018). These studies, paving the way for investigating the microcircuit-level mechanisms of DD, not only found that true DD is present in the rodent neocortex, but that SSA and DD can be observed within different time-windows of the whole MMNr (Chen et al., 2015; Hamm and Yuste, 2016). While SSA is mainly found in earlier time components (40–80 ms after stimulus onset), DD is predominantly visible in a later time window (120–240 ms after stimulus onset). Furthermore, these separate components possibly involve distinct cortical networks and neuronal populations (Chen et al., 2015; Natan et al., 2015; Hamm and Yuste, 2016; Parras et al., 2017; Hamm et al., 2018). Though questions remain on how experimental results regarding SSA and DD in animal models compare to MMNr in human subjects, it can be assumed that they are all part of the same PC process.

MMN and the Canonical Microcircuits for Predictive Coding

The generation of the MMNr is a hierarchical process. While forms of SSA can already be found in the early stages of sensory processing such as the thalamus (Natan et al., 2015; Parras et al., 2017; Hamm et al., 2018), DD seems to be absent in these early processing steps and appears later in the hierarchy, starting from (primary) cortical areas. Moreover, the proportion of neurons showing true DD increases when going up in the cortical hierarchy (Parras et al., 2017; Hamm et al., 2018). This has been taken as an indication that, while SSA can arise from purely feedforward processing streams, DD is more likely dependent on recurrent or feedback processing. This notion fits very well with the connectivity scheme proposed by the PC theory (Bastos et al., 2012; Auksztulewicz and Friston, 2016)—see "The Role of Feedback Processing Across Different Brain States" section.

In the PC framework, top-down predictions play an important role in eliciting mismatch or error-signals. Excitatory connections from (higher-order) cortical areas send predictions to predominantly inhibitory neuronal population in earlier sensory areas (Bastos et al., 2012). These inhibitory neurons enable local pyramidal cells to compare sensory inputs with these predictions, resulting in processes such as SSA (Bastos et al., 2012; Auksztulewicz and Friston, 2016; Yarden and Nelken, 2017). Indeed, it has been shown that deactivating auditory cortex reduces SSA in the superior colliculus of rats, though SSA is not completely extinguished (Anderson and Malmierca, 2013). Furthermore, inhibiting projections from the anterior cingulate cortex (Cg1)—a high-order cortical area involved in modulating visual responses (Zhang et al., 2014; Fiser et al., 2016)—to the primary visual cortex (V1) abolishes DD at a population level in V1 itself (Hamm et al., 2018). This would support a primary role for long-range feedback projections—which are also thought to be essential for conscious processing—in enabling DD. It is important to note, however, that, when splitting neuronal activity into populations that, are "adapting," "deviance detecting" or "non-modulated," the activity of "deviance detecting" neurons in V1 remains intact (Hamm et al., 2018). In other words, while DD seems to disappear when looking at the gross activity of V1, it is preserved at the single-neuron level. This raises the question of how important long-range feedback connections are for facilitating mismatch responses in early sensory cortices, and what their exact role is. Is it possible for example that mismatches caused by low-complexity stimulus features such as visual orientation can be solved within the visual cortex through local (recurrent) connectivity?

To answer this question, we should make a distinction between local/low-level expectations that can operate independently of active cognitive processing (the pre-attentive part of the MMNr), and global/higher-order generalizations that require cognitive or attentional modulation. One way to make this dissociation is by using an adapted version of the oddball paradigm. The local-global paradigm dissociates between two types of predictions, based on local probabilities vs. global rules (**Figure 3A**). It uses blocks of stimuli where a sequence of for instance five stimuli ("xxxxx") is presented interspersed with infrequent blocks in which the final stimulus is changed to a rare/deviant stimulus ("xxxxY"; Bekinschtein et al., 2009; Wacongne et al., 2011). Importantly, in other blocks, the frequent sequence is of the "xxxxY" type, while the deviant sequence consists of five identical stimuli ("xxxxx"). Evoked response potential (ERP) recordings reveal that the fifth, locally deviant tone of the standard "xxxxY" type, although fully predictable, still elicits a MMNr. However, only the rare violation sequence (or global deviant), which contains the five identical tones "xxxxx," elicits a distinct and later novelty response, the "late positive complex" or P3b wave (Bekinschtein et al., 2009, **Figure 3B**). This P3b component has been associated with consciousness through processes such as attention and awareness—see Chennu and Bekinschtein (2012) for an overview. In the next paragraph, we will discuss how the MMNr and related ERP components, such as the P3b wave, are influenced by different states of awareness and how these changes are supported by (long-range) feedback connections.

Mismatch Negativity in Relation to Consciousness

The MMNr has been investigated extensively in relation to different states of consciousness. It has been of interest in cognitive research because of its robustness: MMN can be observed in different states of consciousness, be it awake, asleep or anesthetized (Atienza et al., 2002; Koelsch et al., 2006). This has led to the belief that MMN is an automated response not dependent on active, conscious processing (Näätänen et al., 2001; Stefanics et al., 2014). However, ERP components related to the MMNr, especially the P3b wave, can be altered by conscious processes such as attention and awareness (Woldorff et al., 1998; Näätänen et al., 2007). Consequently, it has been hypothesized that, as the MMNr may be a key NCC, changes in its properties can be used to track different states of awareness, such as sedation or loss of consciousness. In a clinical setting, the estimation of a patient's level of consciousness may be of paramount importance for the correct diagnosis of disorders of consciousness.

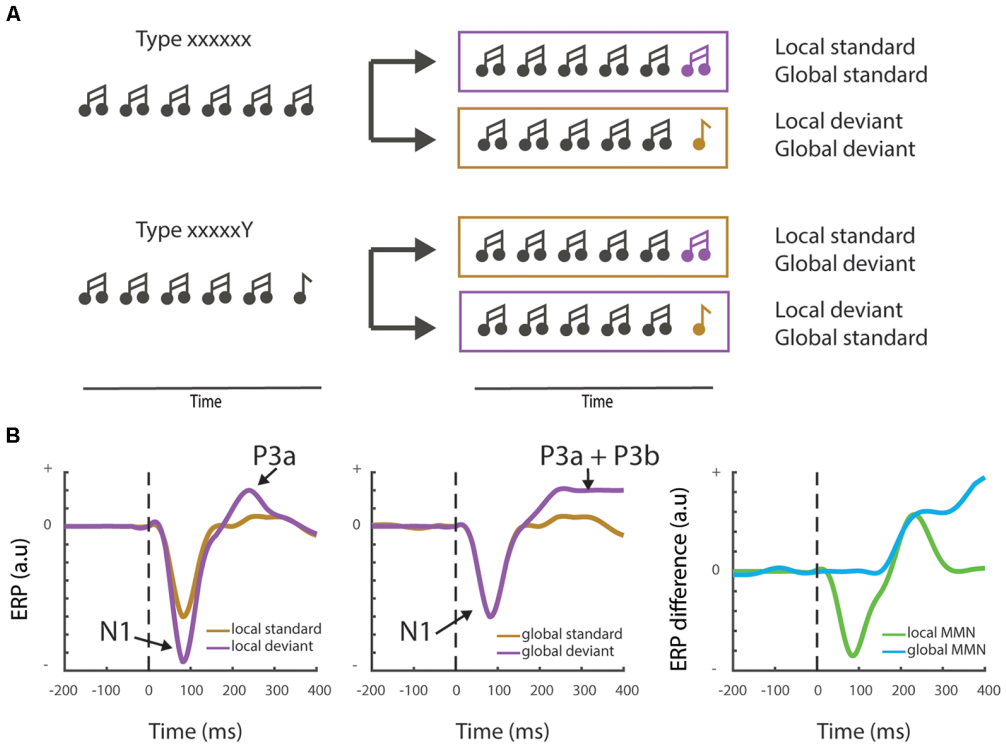

FIGURE 3 | Local vs. global mismatch responses. **(A)** Cartoon showing the local-global oddball paradigm and the distinct stimulus trains that differentiate between local and global deviants. Note that in the "xxxxxY" block the change of the final "Y" to an "x" stimulus induces a global deviant ("xxxxxx"), whereas it remains a standard stimulus at the local level. **(B)** Schematic examples (not actual data—see e.g., Bekinschtein et al., 2009 or Strauss et al., 2015 for measured electrophysiological traces) showing the typically observed field responses during a local-global mismatch paradigm. The left panel shows the classical Mismatch negativity response (MMNr) where the N1 component is more negative for the deviant response compared to the standard response. The local MMNr also generally show an early P3a response. Global deviants (middle panel) may not elicit a change in the N1 response, but induce a maintained sensory novelty or P3b response. The right panel depicts the difference waves (deviant—standard) generally used to visualize the MMNr.

In cognitive electrophysiology, the P3b wave is a part of the most widely studied P300 ERP component. It has been proposed as a marker of conscious perception of salient events or stimuli (Sutton et al., 1965) and more specifically as a sign of the "top-down" deployment of selective attention to task-relevant stimuli. While the P3b response, similarly to the MMNr, is often studied using a form of the oddball paradigm, the processes underlying the two responses may be very different. For example, the processing of statistical irregularities that elicits the MMNr is distinct from, and may not necessarily result in, conscious awareness of a stimulus (Chennu and Bekinschtein, 2012). Furthermore, the MMNr has been reported in humans over various states of consciousness (Atienza et al., 2002; Koelsch et al., 2006; Bekinschtein et al., 2009), whereas the P3b component seems dependent on aware processing of the stimulus (Sergent et al., 2005) and is diminished in states of impaired consciousness (Boly et al., 2011; Strauss et al., 2015; Nourski et al., 2018). Nevertheless, although the P3b response is usually considered an event separate from the MMNr, the two are often temporally overlapping. For example, recent studies in mice assess oddball and mismatch responses in temporal windows up to 400 ms after the onset of sensory stimuli (Chen et al., 2015; Hamm and Yuste, 2016), beyond the typical N1 response which peaks 100 ms after stimulus onsets and is associated with the MMNr in human EEG studies. Here, we will consider the P3b as an ERP component related to the MMNr. Future experiment will need to assess whether the P3b can be considered as a long-latency DD component of the MMNr or rather a separate neuronal event.

Upon loss of consciousness, detection of local deviants remains, at least partially, present in the human brain (Strauss et al., 2015; Nourski et al., 2018). Detection of global deviants is, on the other hand, absent during states of unconsciousness and deep sleep (Strauss et al., 2015; Nourski et al., 2018). Thus, global MMNr and the related P3b component are dependent on cognitive processes that are unavailable during loss of consciousness, whereas local DD remains partially present. While both the local DD component of the MMNr and the P3b response have been hypothesized to be dependent on temporal/frontal feedback connections (Friston, 2005, 2010; Garrido et al., 2009b; Buschman and Kastner, 2015), these differences indicate that distinct forms of feedback projections may be involved in the detection of local vs. global deviants.

As mentioned before (local) MMN has been reported under different states of awareness in both human and animal

experiments (Atienza et al., 2002; Koelsch et al., 2006; Parras et al., 2017). However, questions have been raised on whether these reports truly reflect the whole MMN process, including its deviance detecting component, or whether MMN under low conscious conditions can be fully explained by SSA. In a recent study, Strauss et al. (2015) showed disruption of the local MMNr at intermediate time components during non-REM sleep. These mid-range components have previously been associated with the error-signaling or deviance detecting processes underlying the MMNr (Wacongne et al., 2012). These results indicate that the feedback modulation that is necessary to support not only global but also local DD might be missing during non-conscious brain states. Indeed, another study showed that patients in a vegetative state (VS) display reductions in feedback connectivity, while feedforward connections are maintained (Boly et al., 2011)—but see King et al. (2011). Moreover, intermediate (DD-related) components of the MMNr are not present in these patients, whereas they are still reported in patients in a minimally conscious state (MCS). Thus, the main difference between MCS and VS patients seems to be the reduction of feedback connectivity in VS patients, leading to a disruption in DD. Finally, a recent animal study that supports the previous findings showed that inactivation of feedback connections from frontal cortex to visual cortex disrupts local DD processes in the latter (Hamm et al., 2018).

Altogether, the studies mentioned in the previous paragraphs indicate that: (1) feedback connections are necessary to maintain the (local) DD component of MMN, and (2) feedback connectivity is strongly reduced during non-conscious brain states. Interestingly, however, the same study showed that, while DD disappears upon inhibiting feedback projections when looking at the population level (i.e., in terms of multi-unit activity), at the single-neuron level DD is unchanged (Hamm et al., 2018). Moreover, other studies have reported both MMNr and (local) DD at the single-neuron level in anesthetized animals (Taaseh et al., 2011; Parras et al., 2017). Thus, while DD seems to disappear in population-level recordings during non-conscious states or when prefrontal cortex is inhibited (Boly et al., 2011; Strauss et al., 2015; Hamm et al., 2018), it is preserved in the cortex when looking at the single-neuron level (Parras et al., 2017; Hamm et al., 2018). These discrepancies highlight some of the knowledge gaps in our current understanding of the role of feedback projections in MMN generation and, potentially, their involvement in conscious responses.

Several distinct mechanisms may explain the aforementioned discrepancies. First, local DD might primarily depend—as classically hypothesized and similarly to global DD—on long-range feedback projections. During anesthesia/loss of consciousness such feedback is reduced (but not completely eliminated), and this dampens the extent of local DD in such a way that it is no longer visible at the population level. A second possibility is that, at least for local mismatches, DD is primarily due to feedforward or local (recurrent) mechanisms (**Figure 2B**). Feedback projections, in a manner similar to top-down attention, amplify and synchronize DD across neurons. This makes DD visible at the population level and allows the "error" signal to be more effectively transmitted to higher-order areas. Current experimental evidence does not allow us to disambiguate between these two scenarios. In fact, no experiment has yet been able to either monitor or abolish (or even identify) all sources of feedback that might play a role in the generation of the local DD response. Consequently, only hypotheses can be made on its genesis. Addressing this problem is essential not only to provide an understanding of the mechanisms underlying MMN (which is clinically relevant for the diagnosis of several neuropsychiatric disorders), but also to better understand the neural basis of consciousness.

DISCUSSION: ARE FEEDBACK CONNECTIONS ESSENTIAL FOR CONSCIOUSNESS?

Long-range feedback projections have been hypothesized to play an essential role in the generation of consciousness, yet conclusive evidence for this is lacking. First ("The Role of Feedback Processing Across Different Brain States" section), while it has been shown that long-range feedback is dampened in non-conscious brain states (Boly et al., 2011), some studies found exceptions to this general trend (Chennu et al., 2014), even for cortico-cortical top-down projections. Other studies even showed that, during Non-REM sleep, some forms of inter-areal communication may even be enhanced compared to wakefulness (Olcese et al., 2018a). One possible solution to this discrepancy between theoretical frameworks and experimental evidence may be that not all long-range, cortico-cortical feedback projections are equally important in the generation of consciousness. Identifying which feedback pathways support conscious processing is crucial to develop better markers for the level of consciousness. State-of-the-art diagnostic tools for linking brain dynamics to the level of consciousness, such as the perturbational complexity index (Casali et al., 2013), are able to characterize inter-areal communication in general but lack the specificity to probe specific connections. Second ("Mismatch Negativity: Preserved Top-Down Modulation During Non-conscious Brain States?" section), the role of feedback is being questioned even for fundamental brain mechanisms which have long been held to depend on it and to be hallmarks of conscious states, such as the DD component of the MMNr. Not only may the DD component still be present in non-conscious states, but also it may not be fully dependent on feedback projections (Hamm et al., 2018). In particular, a key question is whether feedback projections contribute to the generation of the DD response, or only facilitate it *via* synchronizing neuronal activity. In other words, is the DD response an error signal computed in higher-order areas (as the PC framework suggests), or is it generated locally and top-down modulation only amplifies it? And, if the latter proves to be the case, what are the local/recurrent mechanisms underlying DD, and is this consciously perceived? Finally, it remains to be addressed whether the feedback functional dynamics responsible for the

DD component of the MMNr are merely a correlate of consciousness (NCC) or rather represent a key mechanism underlying the emergence of consciousness. Addressing these questions is crucial for us to understand the neural bases of consciousness, yet remains challenging and requires an interdisciplinary approach.

First, an experimental paradigm in which DD responses are preserved in non-conscious states at the single-neuron—but not population—level needs to be developed for both humans and animals. Correlational human experiments involving intracranial recordings are needed to investigate whether single-neuron DD is consciously perceived even in the absence of long-range feedback and population-level signals. This can be done *via* retrospective reports (Koch et al., 2016; Siclari et al., 2017), and will allow us to understand if feedback projections are involved in generating conscious percepts, or rather enable to report those. The answer to this question is critical to reveal which cortical regions are involved in the genesis of conscious perception (Koch et al., 2016; Boly et al., 2017; Lamme, 2018). Second, causal explanations provided by animal experiments are subsequently crucial to go beyond the correlational level and reveal what the exact role of feedback projections is in the generation of DD responses. By leveraging and integrating techniques such as high-channel count electrophysiology, two-photon calcium imaging and optogenetics, it will be possible to uncover whether (and which) feedback projections play a role in the computation of the DD response, or whether they only have a modulatory nature. One additional intriguing outcome of such experiments (which can be only achieved by combining human and animal experiments) would also be to finally reveal the differential role of top-down projection from frontal to posterior cortices, or within the posterior cortex (parietal, temporal and occipital regions), and address the debate on which cortical regions support conscious processing (Koch et al., 2016; Boly et al., 2017; Odegaard et al., 2017; Lamme, 2018).

CONCLUSIONS

How does the brain transform sensory stimuli into perceptual experiences? And how does it generate an internal model of the world? These are questions that have puzzled philosophers since the dawn of civilization. Neuroscientists have in the past decades started to formulate theories on how the brain might generate consciousness, yet all theories are lacking in terms of specific, circuit-level mechanisms. Theories focused on addressing the hard problem of consciousness (e.g., IIT) provide a foundational framework of the properties a conscious system needs to possess, but do not specify how this is reflected in terms of neuron-level (micro)circuits. The same is true for neurophysiology-based theories (recurrent processing, GNW, PC framework). General, mesoscopic-scale principles have been outlined, yet very limited insight is provided about the corresponding single-neuron correlates. Irrespective of one's philosophical or theoretical stance (see the debate between Chalmers and Dennett), this is a key question which was until recently not addressable, but that developments in neurotechnology make within reach.

While functional long-range feedback projections are generally agreed to be closely associated to the presence of consciousness, recent studies have begun to indicate that not all feedback (not even within the cortex) may play the same role. The precise function of feedback has proven to be elusive even for phenomena (such as DD) that have historically been considered dependent on feedback. In order to eventually understand consciousness, therefore, two questions must be addressed: which forms of feedback support consciousness—and are not just a correlate of it? And how do they operate, at the microcircuit level?

Answering these questions will make it possible to eventually test the mechanistic role of feedback connections in generating consciousness. Is proper feedback sufficient to generate consciousness, or only necessary, provided that other conditions are met? If proper feedback were sufficient to generate consciousness in information processing systems, any biological or artificial device would have the potential for being conscious (although so far only brains may have achieved the right form of feedback). Although certain configurations of feedback projections are most likely essential to achieve functions normally associated with consciousness—Chalmer's easy problems, but see also Dehaene et al. (2017)—we concur with Tononi and Koch (2015) that merely reproducing a function does not imply possessing consciousness (Tononi and Koch, 2015; Tononi et al., 2016). Thus, while a computer simulation with a proper, yet virtual, form of feedback would likely be able to match conscious systems from a functional point of view, we surmise that a proper physical substrate is a pre-requisite to achieve real consciousness, i.e., experience—see Tononi and Koch (2015) for an in-depth discussion on the topic. Rather, we support the hypothesis that proper feedback projections are necessary to achieve consciousness in computational systems (neuron-based, artificial, etc.), but only if these possess an appropriate physical substrate (Tononi and Koch, 2015; Tononi et al., 2016). Nevertheless, these fascinating questions can only be achieved after a deeper understanding of the role of feedback projections in conscious processing. This will finally allow, we believe, to unveil the mystery which the mechanisms of consciousness still represent.

AUTHOR CONTRIBUTIONS

TS, CB and UO wrote the manuscript.

FUNDING

This study was supported by the European Union's Horizon 2020 Framework Programme for Research and Innovation under the Specific Grant Agreement No. 785907 (Human Brain Project SGA2) to CB and UO, and by the FLAG-ERA JTC 2015 project CANON [co-financed by the Netherlands Organization for Scientific Research (Nederlandse Organisatie voor Wetenschappelijk Onderzoek, NWO)] to CB and UO.

REFERENCES

Adesnik, H., and Naka, A. (2018). Cracking the function of layers in the sensory cortex. *Neuron* 100, 1028–1043. doi: 10.1016/j.neuron.2018.10.032

Akeju, O., and Brown, E. N. (2017). Neural oscillations demonstrate that general anesthesia and sedative states are neurophysiologically distinct from sleep. *Curr. Opin. Neurobiol.* 44, 178–185. doi: 10.1016/j.conb.2017.04.011

Anderson, L. A., and Malmierca, M. S. (2013). The effect of auditory cortex deactivation on stimulus-specific adaptation in the inferior colliculus of the rat. *Eur. J. Neurosci.* 37, 52–62. doi: 10.1111/ejn.12018

Aru, J., Bachmann, T., Singer, W., and Melloni, L. (2012). Distilling the neural correlates of consciousness. *Neurosci. Biobehav. Rev.* 36, 737–746. doi: 10.1016/j.neubiorev.2011.12.003

Atienza, M., Cantero, J. L., and Dominguez-Marin, E. (2002). Mismatch negativity (MMN): an objective measure of sensory memory and long-lasting memories during sleep. *Int. J. Psychophysiol.* 46, 215–225. doi: 10.1016/s0167-8760(02)00113-7

Auksztulewicz, R., and Friston, K. (2016). Repetition suppression and its contextual determinants in predictive coding. *Cortex* 80, 125–140. doi: 10.1016/j.cortex.2015.11.024

Baars, B. J. (2002). The conscious access hypothesis: origins and recent evidence. *Trends Cogn. Sci.* 6, 47–52. doi: 10.1016/s1364-6613(00)01819-2

Baars, B. J. (2005). Global workspace theory of consciousness: toward a cognitive neuroscience of human experience. *Prog. Brain Res.* 150, 45–53. doi: 10.1016/s0079-6123(05)50004-9

Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., and Friston, K. J. (2012). Canonical microcircuits for predictive coding. *Neuron* 76, 695–711. doi: 10.1016/j.neuron.2012.10.038

Bastos, A. M., Vezoli, J., Bosman, C. A., Schoffelen, J.-M., Oostenveld, R., Dowdall, J. R., et al. (2015). Visual areas exert feedforward and feedback influences through distinct frequency channels. *Neuron* 85, 390–401. doi: 10.1016/j.neuron.2014.12.018

Bekinschtein, T. A., Dehaene, S., Rohaut, B., Tadel, F., Cohen, L., and Naccache, L. (2009). Neural signature of the conscious processing of auditory regularities. *Proc. Natl. Acad. Sci. U S A* 106, 1672–1677. doi: 10.1073/pnas.0809667106

Boly, M., Garrido, M. I., Gosseries, O., Bruno, M.-A., Boveroux, P., Schnakers, C., et al. (2011). Preserved feedforward but impaired top-down processes in the vegetative state. *Science* 332, 858–862. doi: 10.1126/science.1202043

Boly, M., Massimini, M., Tsuchiya, N., Postle, B. R., Koch, C., and Tononi, G. (2017). Are the neural correlates of consciousness in the front or in the back of the cerebral cortex? Clinical and neuroimaging evidence. *J. Neurosci.* 37, 9603–9613. doi: 10.1523/JNEUROSCI.3218-16.2017

Bosman, C. A., and Aboitiz, F. (2015). Functional constraints in the evolution of brain circuits. *Front. Neurosci.* 9:303. doi: 10.3389/fnins.2015.00303

Bosman, C. A., Lansink, C. S., and Pennartz, C. M. A. (2014). Functions of gamma-band synchronization in cognition: from single circuits to functional diversity across cortical and subcortical systems. *Eur. J. Neurosci.* 39, 1982–1999. doi: 10.1111/ejn.12606

Buffalo, E. A., Fries, P., Landman, R., Buschman, T. J., and Desimone, R. (2011). Laminar differences in gamma and alpha coherence in the ventral stream. *Proc. Natl. Acad. Sci. U S A* 108, 11262–11267. doi: 10.1073/pnas.1011284108

Buschman, T. J., and Kastner, S. (2015). From behavior to neural dynamics: an integrated theory of attention. *Neuron* 88, 127–144. doi: 10.1016/j.neuron.2015.09.017

Casali, A. G., Gosseries, O., Rosanova, M., Boly, M., Sarasso, S., Casali, K. R., et al. (2013). A theoretically based index of consciousness independent of sensory processing and behavior. *Sci. Transl. Med.* 5:198ra105. doi: 10.1126/scitranslmed.3006294

Casarotto, S., Comanducci, A., Rosanova, M., Sarasso, S., Fecchio, M., Napolitani, M., et al. (2016). Stratification of unresponsive patients by an independently validated index of brain complexity. *Ann. Neurol.* 80, 718–729. doi: 10.1002/ana.24779

Chalmers, D. J. (1995). Facing up to the problem of consciousness. *J. Conscious. Stud.* 2, 200–219.

Chao, Z. C., Takaura, K., Wang, L., Fujii, N., and Dehaene, S. (2018). Large-scale cortical networks for hierarchical prediction and prediction error in the primate brain. *Neuron* 100, 1252.e3–1266.e3. doi: 10.1016/j.neuron.2018.10.004

Chen, I.-W., Helmchen, F., and Lütcke, H. (2015). Specific early and late oddball-evoked responses in excitatory and inhibitory neurons of mouse auditory cortex. *J. Neurosci.* 35, 12560–12573. doi: 10.1523/JNEUROSCI.2240-15.2015

Chennu, S., and Bekinschtein, T. A. (2012). Arousal modulates auditory attention and awareness: insights from sleep, sedation and disorders of consciousness. *Front. Psychol.* 3:65. doi: 10.3389/fpsyg.2012.00065

Chennu, S., Finoia, P., Kamau, E., Allanson, J., Williams, G. B., Monti, M. M., et al. (2014). Spectral signatures of reorganised brain networks in disorders of consciousness. *PLoS Comput. Biol.* 10:e1003887. doi: 10.1371/journal.pcbi.1003887

Clark, A. (2013). Whatever next? Predictive brains, situated agents and the future of cognitive science. *Behav. Brain Sci.* 36, 181–204. doi: 10.1017/s0140525x12000477

Cohen, M. A., and Dennett, D. C. (2011). Consciousness cannot be separated from function. *Trends Cogn. Sci.* 15, 358–364. doi: 10.1016/j.tics.2011.06.008

Constantinople, C. M., and Bruno, R. M. (2013). Deep cortical layers are activated directly by thalamus. *Science* 340, 1591–1594. doi: 10.1126/science.1236425

Crick, F., and Koch, C. (1990). Towards a neurobiological theory of consciousness. *Semin. Neurosci.* 2, 263–275.

Crick, F., and Koch, C. (2003). A framework for consciousness. *Nat. Neurosci.* 6, 119–126. doi: 10.1038/nn0203-119

Dehaene, S., Lau, H., and Kouider, S. (2017). What is consciousness and could machines have it? *Science* 358, 486–492. doi: 10.1126/science.aan8871

Dennett, D. C. (1991). *Consciousness Explained.* Boston: Little, Brown and Co.

Dennett, D. C. (2018). Facing up to the hard question of consciousness. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 373:20170342. doi: 10.1098/rstb.2017.0342

Destexhe, A., Contreras, D., and Steriade, M. (1999). Spatiotemporal analysis of local field potentials and unit discharges in cat cerebral cortex during natural wake and sleep states. *J. Neurosci.* 19, 4595–4608. doi: 10.1523/jneurosci.19-11-04595.1999

Douglas, R. J., and Martin, K. A. C. (2004). Neuronal circuits of the neocortex. *Annu. Rev. Neurosci.* 27, 419–451. doi: 10.1146/annurev.neuro.27.070203.144152

Ercsey-Ravasz, M., Markov, N. T., Lamy, C., Van Essen, D. C., Knoblauch, K., Toroczkai, Z., et al. (2013). A predictive network model of cerebral cortical connectivity based on a distance rule. *Neuron* 80, 184–197. doi: 10.1016/j.neuron.2013.07.036

Erickson, M. A., Ruffle, A., and Gold, J. M. (2016). A meta-analysis of mismatch negativity in schizophrenia: from clinical risk to disease specificity and progression. *Biol. Psychiatry* 79, 980–987. doi: 10.1016/j.biopsych.2015.08.025

Fahrenfort, J. J., Scholte, H. S., and Lamme, V. A. F. (2007). Masking disrupts reentrant processing in human visual cortex. *J. Cogn. Neurosci.* 19, 1488–1497. doi: 10.1162/jocn.2007.19.9.1488

Fahrenfort, J. J., Scholte, H. S., and Lamme, V. A. F. (2008). The spatiotemporal profile of cortical processing leading up to visual perception. *J. Vis.* 8:12. doi: 10.1167/8.1.12

Fahrenfort, J. J., Snijders, T. M., Heinen, K., van Gaal, S., Scholte, H. S., and Lamme, V. A. F. (2012). Neuronal integration in visual cortex elevates face category tuning to conscious face perception. *Proc. Natl. Acad. Sci. U S A* 109, 21504–21509. doi: 10.1073/pnas.1207414110

Farley, B. J., Quirk, M. C., Doherty, J. J., and Christian, E. P. (2010). Stimulus-specific adaptation in auditory cortex is an NMDA-independent process distinct from the sensory novelty encoded by the mismatch negativity. *J. Neurosci.* 30, 16475–16484. doi: 10.1523/jneurosci.2793-10.2010

Felleman, D. J., and Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex* 1, 1–47. doi: 10.1093/cercor/1.1.1

Ferrarelli, F., Massimini, M., Sarasso, S., Casali, A., Riedner, B. A., Angelini, G., et al. (2010). Breakdown in cortical effective connectivity during midazolam-induced loss of consciousness. *Proc. Natl. Acad. Sci. U S A* 107, 2681–2686. doi: 10.1073/pnas.0913008107

Fiser, A., Mahringer, D., Oyibo, H. K., Petersen, A. V., Leinweber, M., and Keller, G. B. (2016). Experience-dependent spatial expectations in mouse visual cortex. *Nat. Neurosci.* 19, 1658–1664. doi: 10.1038/nn.4385

Flores, F. J., Hartnack, K. E., Fath, A. B., Kim, S.-E., Wilson, M. A., Brown, E. N., et al. (2017). Thalamocortical synchronization during induction

and emergence from propofol-induced unconsciousness. *Proc. Natl. Acad. Sci. U S A* 114, E6660–E6668. doi: 10.1073/pnas.1700148114

Fontolan, L., Morillon, B., Liegeois-Chauvel, C., and Giraud, A.-L. (2014). The contribution of frequency-specific activity to hierarchical information processing in the human auditory cortex. *Nat. Commun.* 5:4694. doi: 10.1038/ncomms5694

Fries, P. (2015). Rhythms for cognition: communication through coherence. *Neuron* 88, 220–235. doi: 10.1016/j.neuron.2015.09.034

Friston, K. (2005). A theory of cortical responses. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 360, 815–836. doi: 10.1098/rstb.2005.1622

Friston, K. (2008). Hierarchical models in the brain. *PLoS Comput. Biol.* 4:e1000211. doi: 10.1371/journal.pcbi.1000211

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* 11, 127–138. doi: 10.1038/nrn2787

Garrido, M. I., Kilner, J. M., Kiebel, S. J., Stephan, K. E., Baldeweg, T., and Friston, K. J. (2009a). Repetition suppression and plasticity in the human brain. *Neuroimage* 48, 269–279. doi: 10.1016/j.neuroimage.2009.06.034

Garrido, M. I., Kilner, J. M., Stephan, K. E., and Friston, K. J. (2009b). The mismatch negativity: a review of underlying mechanisms. *Clin. Neurophysiol.* 120, 453–463. doi: 10.1016/j.clinph.2008.11.029

Gazzaley, A., and Nobre, A. C. (2012). Top-down modulation: bridging selective attention and working memory. *Trends Cogn. Sci.* 16, 129–135. doi: 10.1016/j.tics.2011.11.014

Gilbert, C. D., and Li, W. (2013). Top-down influences on visual processing. *Nat. Rev. Neurosci.* 14, 350–363. doi: 10.1038/nrn3476

Grill-Spector, K., Henson, R., and Martin, A. (2006). Repetition and the brain: neural models of stimulus-specific effects. *Trends Cogn. Sci.* 10, 14–23. doi: 10.1016/j.tics.2005.11.006

Gugino, L. D., Chabot, R. J., Prichep, L. S., John, E. R., Formanek, V., and Aglio, L. S. (2001). Quantitative EEG changes associated with loss and return of consciousness in healthy adult volunteers anaesthetized with propofol or sevoflurane. *Br. J. Anaesth.* 87, 421–428. doi: 10.1093/bja/87.3.421

Hall, S. D., Yamawaki, N., Fisher, A. E., Clauss, R. P., Woodhall, G. L., and Stanford, I. M. (2010). GABA(A) alpha-1 subunit mediated desynchronization of elevated low frequency oscillations alleviates specific dysfunction in stroke--a case report. *Clin. Neurophysiol.* 121, 549–555. doi: 10.1016/j.clinph.2009.11.084

Hamm, J. P., and Yuste, R. (2016). Somatostatin interneurons control a key component of mismatch negativity in mouse visual cortex. *Cell Rep.* 16, 597–604. doi: 10.1016/j.celrep.2016.06.037

Hamm, J. P., Shymkiv, Y., Han, S., Yang, W., and Yuste, R. (2018). Cortical subnetworks encode context of a visual stimulus. *bioRxiv* [Preprint]. doi: 10.1101/452219

Harms, L., Michie, P. T., and Näätänen, R. (2016). Criteria for determining whether mismatch responses exist in animal models: focus on rodents. *Biol. Psychol.* 116, 28–35. doi: 10.1016/j.biopsycho.2015.07.006

Harris, K. D., and Shepherd, G. M. G. (2015). The neocortical circuit: themes and variations. *Nat. Neurosci.* 18, 170–181. doi: 10.1038/nn.3917

Hubel, D. H., and Wiesel, T. N. (1977). Ferrier lecture. Functional architecture of macaque monkey visual cortex. *Proc. R. Soc. Lond. B Biol. Sci.* 198, 1–59. doi: 10.1098/rspb.1977.0085

King, J.-R., Bekinschtein, T., and Dehaene, S. (2011). Comment on "Preserved feedforward but impaired top-down processes in the vegetative state." *Science* 334:1203. doi: 10.1126/science.1210012

Koch, C., Massimini, M., Boly, M., and Tononi, G. (2016). Neural correlates of consciousness: progress and problems. *Nat. Rev. Neurosci.* 17, 307–321. doi: 10.1038/nrn.2016.22

Koelsch, S., Heinke, W., Sammler, D., and Olthoff, D. (2006). Auditory processing during deep propofol sedation and recovery from unconsciousness. *Clin. Neurophysiol.* 117, 1746–1759. doi: 10.1016/j.clinph.2006.05.009

Kwon, S. E., Yang, H., Minamisawa, G., and O'Connor, D. H. (2016). Sensory and decision-related activity propagate in a cortical feedback loop during touch perception. *Nat. Neurosci.* 19, 1243–1249. doi: 10.1038/nn.4356

Lamme, V. A. F. (2006). Towards a true neural stance on consciousness. *Trends Cogn. Sci.* 10, 494–501. doi: 10.1016/j.tics.2006.09.001

Lamme, V. A. F. (2018). Challenges for theories of consciousness: seeing or knowing, the missing ingredient and how to deal with panpsychism. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 373:20170344. doi: 10.1098/rstb.2017.0344

Lamme, V. A., and Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends Neurosci.* 23, 571–579. doi: 10.1016/s0166-2236(00)01657-x

LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature* 521, 436–444. doi: 10.1038/nature14539

Levine, J. (1983). Materialism and qualia: the explanatory gap. *Pac. Philos. Q.* 64, 354–361. doi: 10.1111/j.1468-0114.1983.tb00207.x

Lewis, C. M., Bosman, C. A., and Fries, P. (2015). Recording of brain activity across spatial scales. *Curr. Opin. Neurobiol.* 32, 68–77. doi: 10.1016/j.conb.2014.12.007

Manita, S., Suzuki, T., Homma, C., Matsumoto, T., Odagawa, M., Yamada, K., et al. (2015). A top-down cortical circuit for accurate sensory perception. *Neuron* 86, 1304–1316. doi: 10.1016/j.neuron.2015.05.006

Markov, N. T., Ercsey-Ravasz, M. M., Ribeiro Gomes, A. R., Lamy, C., Magrou, L., Vezoli, J., et al. (2014a). A weighted and directed interareal connectivity matrix for macaque cerebral cortex. *Cereb. Cortex* 24, 17–36. doi: 10.1093/cercor/bhs270

Markov, N. T., Vezoli, J., Chameau, P., Falchier, A., Quilodran, R., Huissoud, C., et al. (2014b). Anatomy of hierarchy: feedforward and feedback pathways in macaque visual cortex. *J. Comp. Neurol.* 522, 225–259. doi: 10.1002/cne.23458

Markov, N. T., Ercsey-Ravasz, M., Lamy, C., Ribeiro Gomes, A. R., Magrou, L., Misery, P., et al. (2013a). The role of long-range connections on the specificity of the macaque interareal cortical network. *Proc. Natl. Acad. Sci. U S A* 110, 5187–5192. doi: 10.1073/pnas.1218972110

Markov, N. T., Ercsey-Ravasz, M., Van Essen, D. C., Knoblauch, K., Toroczkai, Z., and Kennedy, H. (2013b). Cortical high-density counterstream architectures. *Science* 342:1238406. doi: 10.1126/science.1238406

Massimini, M., Ferrarelli, F., Huber, R., Esser, S. K., Singh, H., and Tononi, G. (2005). Breakdown of cortical effective connectivity during sleep. *Science* 309, 2228–2232. doi: 10.1126/science.1117256

Massimini, M., Ferrarelli, F., Sarasso, S., and Tononi, G. (2012). Cortical mechanisms of loss of consciousness: insight from TMS/EEG studies. *Arch. Ital. Biol.* 150, 44–55. doi: 10.4449/aib.v150i2.1361

May, P. J. C., and Tiitinen, H. (2010). Mismatch negativity (MMN), the deviance-elicited auditory deflection, explained. *Psychophysiology* 47, 66–122. doi: 10.1111/j.1469-8986.2009.00856.x

Michalareas, G., Vezoli, J., van Pelt, S., Schoffelen, J.-M., Kennedy, H., and Fries, P. (2016). Alpha-beta and gamma rhythms subserve feedback and feedforward influences among human visual cortical areas. *Neuron* 89, 384–397. doi: 10.1016/j.neuron.2015.12.018

Mikulan, E., Hesse, E., Sedeño, L., Bekinschtein, T., Sigman, M., García, M. D. C., et al. (2018). Intracranial high-γ connectivity distinguishes wakefulness from sleep. *Neuroimage* 169, 265–277. doi: 10.1016/j.neuroimage.2017.12.015

Miller, E. K., and Desimone, R. (1994). Parallel neuronal mechanisms for short-term memory. *Science* 263, 520–522. doi: 10.1126/science.8290960

Näätänen, R., Paavilainen, P., Rinne, T., and Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clin. Neurophysiol.* 118, 2544–2590. doi: 10.1016/j.clinph.2007.04.026

Näätänen, R., Tervaniemi, M., Sussman, E., Paavilainen, P., and Winkler, I. (2001). "Primitive intelligence" in the auditory cortex. *Trends Neurosci.* 24, 283–288. doi: 10.1016/s0166-2236(00)01790-2

Näätänen, R., Todd, J., and Schall, U. (2016). Mismatch negativity (MMN) as biomarker predicting psychosis in clinically at-risk individuals. *Biol. Psychol.* 116, 36–40. doi: 10.1016/j.biopsycho.2015.10.010

Naccache, L. (2018). Why and how access consciousness can account for phenomenal consciousness. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 373:20170357. doi: 10.1098/rstb.2017.0357

Natan, R. G., Briguglio, J. J., Mwilambwe-Tshilobo, L., Jones, S. I., Aizenberg, M., Goldberg, E. M., et al. (2015). Complementary control of sensory adaptation by two types of cortical interneurons. *Elife* 4:e09868. doi: 10.7554/elife.09868

Nelken, I., and Ulanovsky, N. (2007). Mismatch negativity and stimulus-specific adaptation in animal models. *J. Psychophysiol.* 21, 214–223. doi: 10.1027/0269-8803.21.34.214

Nourski, K. V., Steinschneider, M., Rhone, A. E., Kawasaki, H., Howard, M. A., and Banks, M. I. (2018). Auditory predictive coding across awareness states under anesthesia: an intracranial electrophysiology study. *J. Neurosci.* 38, 8441–8452. doi: 10.1523/JNEUROSCI.0967-18.2018

Odegaard, B., Knight, R. T., and Lau, H. (2017). Should a few null findings falsify prefrontal theories of conscious perception? *J. Neurosci.* 37, 9593–9602. doi: 10.1523/jneurosci.3217-16.2017

Oizumi, M., Albantakis, L., and Tononi, G. (2014). From the phenomenology to the mechanisms of consciousness: integrated information theory 3.0. *PLoS Comput. Biol.* 10:e1003588. doi: 10.1371/journal.pcbi.1003588

Olcese, U. (2018). Non-REM sleep and the neural correlates of consciousness: more than meets the eye. *Arch. Ital. Biol.* 156, 137–148. doi: 10.12871/00039829201835

Olcese, U., Bos, J. J., Vinck, M., and Pennartz, C. M. A. (2018a). Functional determinants of enhanced and depressed interareal information flow in nonrapid eye movement sleep between neuronal ensembles in rat cortex and hippocampus. *Sleep* 41, 1–18. doi: 10.1093/sleep/zsy167

Olcese, U., Oude Lohuis, M. N., and Pennartz, C. M. A. (2018b). Sensory processing across conscious and nonconscious brain states: from single neurons to distributed networks for inferential representation. *Front. Syst. Neurosci.* 12:49. doi: 10.3389/fnsys.2018.00049

Olcese, U., Bos, J. J., Vinck, M., Lankelma, J. V., van Mourik-Donga, L. B., Schlumm, F., et al. (2016). Spike-based functional connectivity in cerebral cortex and hippocampus: loss of global connectivity is coupled to preservation of local connectivity during non-REM sleep. *J. Neurosci.* 36, 7676–7692. doi: 10.1523/JNEUROSCI.4201-15.2016

Parras, G. G., Nieto-Diego, J., Carbajal, G. V., Valdés-Baizabal, C., Escera, C., and Malmierca, M. S. (2017). Neurons along the auditory pathway exhibit a hierarchical organization of prediction error. *Nat. Commun.* 8:2148. doi: 10.1038/s41467-017-02038-6

Pennartz, C. M. A. (2009). Identification and integration of sensory modalities: neural basis and relation to consciousness. *Conscious. Cogn.* 18, 718–739. doi: 10.1016/j.concog.2009.03.003

Pennartz, C. M. A. (2015). *The Brain's Representational Power*. The MIT Press Available online at: https://mitpress.mit.edu/books/brains-representational-power. Accessed November 24, 2017.

Pesaran, B., Vinck, M., Einevoll, G. T., Sirota, A., Fries, P., Siegel, M., et al. (2018). Investigating large-scale brain dynamics using field potential recordings: analysis and interpretation. *Nat. Neurosci.* 21, 903–919. doi: 10.1038/s41593-018-0171-8

Rao, R. P., and Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat. Neurosci.* 2, 79–87. doi: 10.1038/4580

Richter, C. G., Thompson, W. H., Bosman, C. A., and Fries, P. (2017). Top-down beta enhances bottom-up gamma. *J. Neurosci.* 37, 6698–6711. doi: 10.1523/jneurosci.3771-16.2017

Riesenhuber, M., and Poggio, T. (1999). Are cortical models really bound by the "binding problem"? *Neuron* 24, 87–93, 111–125. doi: 10.1016/s0896-6273(00)80824-7

Rohenkohl, G., Bosman, C. A., and Fries, P. (2018). Gamma synchronization between V1 and V4 improves behavioral performance. *Neuron* 100, 953.e3–963.e3. doi: 10.1016/j.neuron.2018.09.019

Sergent, C., and Dehaene, S. (2004). Neural processes underlying conscious perception: experimental findings and a global neuronal workspace framework. *J. Physiol. Paris* 98, 374–384. doi: 10.1016/j.jphysparis.2005.09.006

Sergent, C., Baillet, S., and Dehaene, S. (2005). Timing of the brain events underlying access to consciousness during the attentional blink. *Nat. Neurosci.* 8, 1391–1400. doi: 10.1038/nn1549

Shepherd, G. M. (2011). The microcircuit concept applied to cortical evolution: from three-layer to six-layer cortex. *Front. Neuroanat.* 5:30. doi: 10.3389/fnana.2011.00030

Siclari, F., and Tononi, G. (2017). Local aspects of sleep and wakefulness. *Curr. Opin. Neurobiol.* 44, 222–227. doi: 10.1016/j.conb.2017.05.008

Siclari, F., Baird, B., Perogamvros, L., Bernardi, G., LaRocque, J. J., Riedner, B., et al. (2017). The neural correlates of dreaming. *Nat. Neurosci.* 20, 872–878. doi: 10.1038/nn.4545

Stefanics, G., Kremláček, J., and Czigler, I. (2014). Visual mismatch negativity: a predictive coding view. *Front. Hum. Neurosci.* 8:666. doi: 10.3389/fnhum.2014.00666

Storm, J. F., Boly, M., Casali, A. G., Massimini, M., Olcese, U., Pennartz, C. M. A., et al. (2017). Consciousness regained: disentangling mechanisms, brain systems and behavioral responses. *J. Neurosci.* 37, 10882–10893. doi: 10.1523/JNEUROSCI.1838-17.2017

Strauss, M., Sitt, J. D., King, J.-R., Elbaz, M., Azizi, L., Buiatti, M., et al. (2015). Disruption of hierarchical predictive coding during sleep. *Proc. Natl. Acad. Sci. U S A* 112, E1353–E1362. doi: 10.1073/pnas.1501026112

Sutton, S., Braren, M., Zubin, J., and John, E. R. (1965). Evoked-potential correlates of stimulus uncertainty. *Science* 150, 1187–1188. doi: 10.1126/science.150.3700.1187

Taaseh, N., Yaron, A., and Nelken, I. (2011). Stimulus-specific adaptation and deviance detection in the rat auditory cortex. *PloS One* 6:e23369. doi: 10.1371/journal.pone.0023369

Tiitinen, H., May, P., Reinikainen, K., and Näätänen, R. (1994). Attentive novelty detection in humans is governed by pre-attentive sensory memory. *Nature* 372, 90–92. doi: 10.1038/372090a0

Tononi, G. (2004). An information integration theory of consciousness. *BMC Neurosci.* 5:42. doi: 10.1186/1471-2202-5-42

Tononi, G., and Edelman, G. M. (1998a). Consciousness and complexity. *Science* 282, 1846–1851. doi: 10.1126/science.282.5395.1846

Tononi, G., and Edelman, G. M. (1998b). Consciousness and the integration of information in the brain. *Adv. Neurol.* 77, 245–279.

Tononi, G., and Koch, C. (2015). Consciousness: here, there and everywhere? *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 370:20140167. doi: 10.1098/rstb.2014.0167

Tononi, G., Boly, M., Massimini, M., and Koch, C. (2016). Integrated information theory: from consciousness to its physical substrate. *Nat. Rev. Neurosci.* 17, 450–461. doi: 10.1038/nrn.2016.44

Townsend, R. G., Solomon, S. S., Chen, S. C., Pietersen, A. N. J., Martin, P. R., Solomon, S. G., et al. (2015). Emergence of complex wave patterns in primate cerebral cortex. *J. Neurosci.* 35, 4657–4662. doi: 10.1523/JNEUROSCI.4509-14.2015

Ulanovsky, N., Las, L., and Nelken, I. (2003). Processing of low-probability sounds by cortical neurons. *Nat. Neurosci.* 6, 391–398. doi: 10.1038/nn1032

van Gaal, S., Lamme, V. A. F., Fahrenfort, J. J., and Ridderinkhof, K. R. (2011). Dissociable brain mechanisms underlying the conscious and unconscious control of behavior. *J. Cogn. Neurosci.* 23, 91–105. doi: 10.1162/jocn.2010.21431

van Kerkoerle, T., Self, M. W., Dagnino, B., Gariel-Mathis, M.-A., Poort, J., van der Togt, C., et al. (2014). Alpha and gamma oscillations characterize feedback and feedforward processing in monkey visual cortex. *Proc. Natl. Acad. Sci. U S A* 111, 14332–14341. doi: 10.1073/pnas.1402773111

van Vugt, B., Dagnino, B., Vartak, D., Safaai, H., Panzeri, S., Dehaene, S., et al. (2018). The threshold for conscious report: signal loss and response bias in visual and frontal cortex. *Science* 360, 537–542. doi: 10.1126/science.aar7186

Vijayan, S., Ching, S., Purdon, P. L., Brown, E. N., and Kopell, N. J. (2013). Thalamocortical mechanisms for the anteriorization of α rhythms during propofol-induced unconsciousness. *J. Neurosci.* 33, 11070–11075. doi: 10.1523/JNEUROSCI.5670-12.2013

von Stein, A., Chiang, C., and König, P. (2000). Top-down processing mediated by interareal synchronization. *Proc. Natl. Acad. Sci. U S A* 97, 14748–14753. doi: 10.1073/pnas.97.26.14748

Wacongne, C., Changeux, J.-P., and Dehaene, S. (2012). A neuronal model of predictive coding accounting for the mismatch negativity. *J. Neurosci.* 32, 3665–3678. doi: 10.1523/JNEUROSCI.5003-11.2012

Wacongne, C., Labyt, E., van Wassenhove, V., Bekinschtein, T., Naccache, L., and Dehaene, S. (2011). Evidence for a hierarchy of predictions and prediction errors in human cortex. *Proc. Natl. Acad. Sci. U S A* 108, 20754–20759. doi: 10.1073/pnas.1117807108

Williams, S. T., Conte, M. M., Goldfine, A. M., Noirhomme, Q., Gosseries, O., Thonnard, M., et al. (2013). Common resting brain dynamics indicate a possible mechanism underlying zolpidem response in severe brain injury. *eLife* 2:e01157. doi: 10.7554/elife.01157

Woldorff, M. G., Hillyard, S. A., Gallen, C. C., Hampson, S. R., and Bloom, F. E. (1998). Magnetoencephalographic recordings demonstrate attentional modulation of mismatch-related neural activity in human auditory cortex. *Psychophysiology* 35, 283–292. doi: 10.1017/s0048577298961601

Womelsdorf, T., Valiante, T. A., Sahin, N. T., Miller, K. J., and Tiesinga, P. (2014). Dynamic circuit motifs underlying rhythmic gain control, gating and integration. *Nat. Neurosci.* 17, 1031–1039. doi: 10.1038/nn.3764

Yarden, T. S., and Nelken, I. (2017). Stimulus-specific adaptation in a recurrent network model of primary auditory cortex. *PLoS Comput. Biol.* 13:e1005437. doi: 10.1371/journal.pcbi.1005437

Zhang, S., Xu, M., Kamigaki, T., Hoang Do, J. P., Chang, W.-C., Jenvay, S., et al. (2014). Selective attention. Long-range and local circuits for top-down modulation of visual cortex processing. *Science* 345, 660–665. doi: 10.1126/science.1254126

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Sikkens, Bosman and Olcese. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms