**Beyond the neural correlates of consciousness: using brain stimulation to elucidate causal mechanisms underlying conscious states and contents**

**Corinne A. Bareham\*, Matt Oxner\*, Tim Gastrell, and David Carmel**

**School of Psychology, Victoria University of Wellington**

\* These authors contributed equally

Corresponding author: David Carmel, [david.carmel@vuw.ac.nz](mailto:david.carmel@vuw.ac.nz)

Keywords:

Consciousness, States of Consciousness, Perceptual Awareness, TMS, Vegetative State, Unresponsive Wakefulness, Minimally Conscious State, Bistable Perception

**Abstract**

We are conscious beings: Somehow, the activity of our brains and nervous systems gives rise to states in which we have subjective experiences; and when we are in such states, we are aware of specific content. Researchers are only beginning to develop an understanding of the neural mechanisms underlying these phenomena. Much of the research in the last few decades has focussed on discerning the neural correlates of consciousness, using neuroimaging methods. Correlates, however, are not causes: to draw inferences about the processes that generate consciousness, rather than accompanying it, one must manipulate brain activity and examine the effects this has on conscious states and contents. One way to do so is to use brain stimulation techniques, such as transcranial magnetic stimulation (TMS). In this review, we survey the consciousness literature with a special emphasis on TMS studies. We begin by examining what is known about the neural substrates of *states* of consciousness – the kinds of brain activity that determine whether a person is awake, asleep, or suffering from a disorder of consciousness. We then delve into the *contents* of consciousness, by examining the literature on perceptual awareness. Throughout, we highlight current controversies and promising avenues for further research.

**Introduction**

We are all intimately familiar with being conscious: We have subjective experiences of ourselves and the world around us, and every day we go through major transitions between states of consciousness such as sleep and wakefulness. But we are only beginning to develop a scientific understanding of the neural mechanisms that give rise to these different aspects of consciousness (Block et al., 2014; Michel et al., 2019). What patterns of brain activity determine whether a person is awake or asleep, in a coma or vegetative state? And when one is awake (and conscious), what are the mechanisms that determine the subjective experiences evoked by stimuli that reach awareness?

In addressing these questions, researchers have largely followed a scientific agenda formulated in the 1990s, when the pursuit of answers was framed as a search for the *neural correlates of consciousness* (Crick & Koch, 1990, 1995; Tononi & Edelman, 1998; Rees et al., 2002). Indeed, correlational neuroimaging methods – most notably, functional magnetic resonance imaging (fMRI) and electroencephalography (EEG) – have led to great progress in elucidating the neural activity that accompanies different conscious experiences or characterizes states of consciousness. Correlation, however, is insufficient for determining the causal mechanisms underlying the relevant processes. To draw causal inferences about neural functions, one must *manipulate* neural activity – for example, using brain stimulation – and measure the effects this has on relevant measures (Walsh & Cowie, 2000; Pascual-Leone et al., 2000). Here, we review research that has examined the effects of brain stimulation on measures of consciousness and associated neural activity. We outline the preliminary insights this work provides into the *neural causes of consciousness*, as well as new questions these insights raise and promising avenues for further work.

Our review focusses on research using transcranial magnetic stimulation (TMS), the most common method for non-invasive brain stimulation1. TMS operates by passing a strong, brief electric current through a conductive coil that is held against a person’s head. The coil is isolated, so the person does not receive an electric shock; however, an electric current is always accompanied by a magnetic field – and the brief magnetic pulse induces corresponding electric activity in the part of the brain closest to the coil, disrupting coordinated activity in the neural tissue (Hallett, 2007; Parkin et al., 2015).

Within well-established operating parameters, TMS poses no substantial risks and its effects are brief and transient (Rossi et al., 2009). It is a rather crude tool – it affects millions of neurons simultaneously, in a cortical region measuring a few centimetres in diameter and depth, depending on stimulation amplitude. However, it enables testing of hypotheses about the involvement of specific brain regions in particular functions. Furthermore, the neural response to a TMS pulse spreads from the stimulated locus to other regions, enabling measurement of connectivity patterns (Massimini et al., 2005). And finally, the temporal precision of brief TMS pulses allows researchers to pinpoint the timing, relative to other events, of a region’s involvement in a function (Pascual-Leone et al., 2000).

TMS can be used “online” to interfere with ongoing task performance or to create spreading activation in the brain. Alternatively, it can be used “offline”, before a function is required, to influence a region’s subsequent excitability. Most offline stimulation is inhibitory, although some (less well-established) stimulation protocols appear to facilitate activity (Huang et al., 2005; Rothkegel et al., 2010). (For a detailed description of TMS protocols and experimental design considerations, see Sandrini et al., 2011 and Parkin et al., 2015.)

This review covers two largely-independent branches of research that are currently at the centre of consciousness science. In ***Part 1***, we survey studies on the neural mechanisms that determine an individual’s *state of consciousness*. This research has shed light on the patterns of activity that underlie different states – wakefulness, sleep, anaesthesia and the states known collectively as disorders of consciousness. It has also elucidated the kinds of neural information processing that can happen in each state and during transitions between them. In ***Part 2***, we delve into research on the mechanisms that shape *perceptual awareness* – the momentary content of our consciousness when we are in a conscious state. This research addresses the processes that determine which aspects of our environment and ourselves we become aware of, and what mechanisms shape our experience (or interpretation) of stimuli. In each part, we review relevant neuroimaging findings that have provided important correlational data on the brain activity associated with consciousness. We then explore how brain stimulation experiments are taking researchers beyond neural correlates to illuminate the neural *causes* of consciousness.

**Part 1. States of consciousness: neural mechanisms and information processing**

A person’s state of consciousness can be described in terms of two main dimensions (Laureys, 2005). The first is wakefulness – our level of physiological *arousal*. The second is the presence of awareness – the ability to have conscious *content*, including thoughts, feelings, and subjective perception of ourselves and our surroundings (note that this dimension does not describe the *specific* contents of awareness – we address the processes that determine *what* we are aware of in Part 2; here, it refers to the extent to which one is capable of being aware of anything at all). Wakefulness and awareness interact, producing our overall state of consciousness: When we are fully awake, we have both high wakefulness and high awareness; every night, as we drift off to sleep, both wakefulness and awareness gradually decline; under anaesthesia, both wakefulness and awareness are reduced even further.

It may seem counterintuitive to divide consciousness into these separate dimensions: Aren’t we aware of a large variety of stimuli when we are awake, and unaware of them when we are not awake? The distinction between wakefulness and awareness is useful, though, because the two do not always go hand-in-hand; in some states of consciousness, one may be high while the other is low. During sleep, for example, there is a distinction between rapid eye movement (REM) and non-REM (nREM) sleep. In nREM sleep, which comprises light and deep sleep stages, EEG activity is clearly distinguishable from that seen during wakefulness. (Deep nREM sleep is also known as “slow-wave sleep” because it shows high-amplitude slow waves (0.5-4.5Hz) that are not seen in other states.) Importantly, nREM sleep is where wakefulness and awareness match: it is characterized by a reduction in both as sleep becomes deeper. On the other hand, REM sleep (during which people make rapid bursts of eye movements) is a deep-sleep state in which the brain is highly active, and its EEG activity resembles that seen during wakefulness. During REM sleep, we often have vivid dreams – in other words, our wakefulness is low (we are asleep), but awareness is high (we experience feelings, have thoughts, and perceive images and sounds).

The opposite combination – reduced awareness with high wakefulness – is seen in some clinical states caused by brain trauma, collectively known as disorders of consciousness (DOCs). The *vegetative state* is a DOC in which patients have sleep-wake cycles and spend long periods with their eyes open, but do not demonstrate any evidence of awareness or meaningful behaviour (Giacino, Kalmar, & Whyte, 2004; tellingly, this disorder is also commonly known as *unresponsive wakefulness syndrome*, Laureys et al., 2010). Some vegetative patients recover sufficiently to be recategorized as being in a *minimally conscious state*, where there is some – often inconsistent – evidence of awareness, such as responding appropriately to an instruction to follow a moving object with the eyes (Laureys et al., 2001).

To understand the neural mechanisms that determine conscious states, researchers have explored neural information processing in response to environmental stimuli, as well as patterns of brain activity in different states and during the transitions between them. We describe these lines of research in the next two sections, before turning to recent contributions from brain stimulation studies.

*Information processing in different states of consciousness*

In the healthy brain, reduced wakefulness usually entails reduced awareness of sensory stimuli: falling asleep impairs perception and the ability to interact with the environment (Ogilvie, 2001). Neural and behavioural measures, however, indicate that this change is gradual, with some preserved (though attenuated) responsiveness to environmental stimuli during the early stages of sleep (for a review, see Goupil & Bekinschtein, 2012).

Recent findings have demonstrated sustained cognitive processing of external stimuli without full wakefulness. People can continue to make button-press responses in sound discrimination tasks during early, light sleep (Bareham, Bekinschtein, Scott, & Manly, 2015; Bareham, Manly, Pustovaya, Scott, & Bekinschtein, 2014). Moreover, well into unresponsive nREM sleep, EEG measures continue to show responses to sounds that are salient (Chennu & Bekinschtein, 2012) or novel (Atienza, Cantero, & Escera, 2001; Chennu & Bekinschtein, 2012; Cote, 2002). In speech processing tasks, even when people stop responding as they fall asleep, EEG markers of appropriate motor-preparation for behavioural responses are still present (Kouider, Andrillon, Barbosa, Goupil, & Bekinschtein, 2014). There is even evidence that novel sensory associations can be learned while asleep: pairing pleasant and unpleasant odours with distinct tones during sleep alters subsequent sniff responses to those tones (enhancing or reducing them for the pleasant or unpleasant association, respectively) during both sleep and wakefulness (Arzi et al., 2012).

Intriguingly, during the transition from wakefulness to early sleep, awareness for the left side of the world fades before the right. For example, Bareham et al. (2014) asked participants to report which side of space auditory tones came from, and to continue doing so as they fell asleep. When they were drowsy, participants were more likely to misperceive tones on the left as coming from the right than vice versa. This mis-localization implies that drowsiness is associated with a fragmentation of awareness (indicated here by reduced ability to bind a sound with its location) although the reason for the spatial asymmetry remains unclear (Bareham et al., 2014; Chennu & Bekinschtein, 2012).

Fragmentation of awareness is also seen with sedation. Increasing blood concentrations of propofol, an anaesthetic, reduced accuracy and prolonged reaction times in a behavioural sound discrimination task where participants reported which of two sounds was played (Chennu, O’Connor, Adapa, Menon, & Bekinschtein, 2016). The reduction in accuracy (i.e., not just a reduction in the overall number of responses, but a lower proportion of correct responses out of those that were made) suggests that reduced wakefulness leads to a decline in the integration of stimulus features.

Although the defining characteristic of DOCs is a deficit in awareness during wakefulness, fMRI evidence suggests that some behaviourally-unresponsive vegetative patients nonetheless show the same neural activity as healthy controls when instructed to perform a mental imagery task (Owen et al., 2006). Furthermore, a subset of such neurally-responsive patients can even answer yes/no questions through such imagery-related brain activity, by arbitrarily associating each answer with a particular type of imagery (Fernández-Espejo & Owen, 2013; Monti et al., 2010). It is currently unclear what distinguishes patients who show such neural responsiveness from those who do not, and to what extent such responsiveness can be taken to indicate partial or full awareness.

Our understanding of states of consciousness is informed by the ways they differ (or don’t) in responsiveness to environmental stimuli. Responsiveness, however, is likely to depend on the brain’s overall state – its level and patterns of activity. The neural activity underlying different states has been characterised in a wealth of neuroimaging studies, which we turn to next.

*Neural activity patterns in different states of consciousness*

During sleep, anaesthesia and DOCs, overall brain metabolism is lower than in wakefulness (Boveroux et al., 2008). However, wakefulness does not seem to be just a matter of overall activity level, but is associated with specific patterns of activity (Baars et al., 2003; Klimova, 2014). Such patterns are detected through measures of *functional connectivity*, which use various statistical techniques to assess temporal correlations between neural events in different brain regions, capturing probable networks of interacting regions.

Using simultaneously-recorded data from fMRI and EEG, researchers have shown that activity in subcortical regions known to regulate arousal (e.g., the thalamus) covaries with activity levels in the default mode network, a set of frontal and parietal cortical regions that is typically active during wakeful rest. The functional connectivity between thalamus and cortex, as well as between different cortical regions, breaks down during nREM sleep (Sadaghiani et al., 2010; Sämann et al., 2011; Spoormaker et al., 2010) and sedation (Stamatakis, Adapa, Absalom, & Menon, 2010). In DOCs, functional connectivity is typically impaired (Demertzi et al., 2015; Fernandez-Espejo et al., 2012; Vanhaudenhuyse et al., 2010), and the physical connectivity between subcortical and cortical structures is often damaged (Fernández-Espejo et al., 2011; Newcombe et al., 2010). These findings suggest that projections from subcortical and brainstem structures coordinate the activity of cortical networks that support wakefulness (Bodien, Chatelle, & Edlow, 2017; Hobson & Pace-Schott, 2002; Spoormaker et al., 2010), and highlight the importance of the integrity of connections between regions, as they allow exchange of information across the brain.

The growing understanding of the role of information exchange is in line with various theoretical frameworks that view communication and integration of information between brain regions as an essential requirement for consciousness (Dehaene, Changeux, Naccache, Sackur, & Sergent, 2006; Tononi, 2004). Several studies have applied mathematical measures of information sharing to neuroimaging data, and used these measures to successfully discriminate between states of consciousness (Chennu et al., 2014; King et al., 2013; Sitt et al., 2014), as well as to characterise transitions between states during sedation (Chennu et al., 2017; Chennu et al., 2016) and recovery from brain injury (Bareham et al., 2018). Such measures are thus extremely useful for understanding the associations between neural activity patterns and states of consciousness; however, they still rely on correlations between states of consciousness and spontaneously-arising activity patterns. We now turn, therefore, to brain stimulation research that has examined the patterns generated by direct manipulations of neural activity in different states of consciousness.

*Using TMS to probe brain activity and connectivity in different states of consciousness*

As noted above, causal inference requires manipulation. A growing number of studies has investigated how manipulating brain activity – by perturbing it with TMS – affects EEG measures of brain activity in different states of consciousness. TMS can indirectly influence activity beyond the stimulated region; these non-local effects provide evidence for *effective connectivity*, an indication of causal interactions in the brain2 (**Figure 1A**).

During wakefulness, a TMS pulse causes cortical activity that propagates across multiple brain regions in a complex pattern. During nREM sleep, however, a similar pulse evokes a response that has a larger amplitude, but is much more brief and spatially-restricted (Massimini et al., 2005). In contrast, during REM sleep TMS-evoked activity is similar (albeit shorter-lasting) to that induced during wakefulness (Massimini et al., 2010; **Figure 1B**).

Patterns of TMS-induced spreading activation have been investigated both in DOCs and in people under sedation. Rosanova et al. (2012) examined DOC patients, and observed brief, spatially-restricted responses – similar to those seen during nREM sleep – when TMS was applied to awake vegetative patients. The pattern of activity in awake minimally-conscious patients, on the other hand, was similar to that seen in healthy wakefulness, suggesting that the vegetative and minimally conscious states differ substantively in effective connectivity. Furthermore, longitudinal observations in the same study showed that improvements in vegetative patients’ conditions were predicted by increases in the complexity and duration of evoked EEG responses to TMS. These increases could precede any behavioural change, a finding that not only has important implications for prognosis in DOCs, but also suggests a possible causal connection between conscious state and effective connectivity: the physical connections between brain regions (which are present in both the vegetative and minimally conscious states) may be less important in determining conscious state than the extent to which such connections are active (as revealed by TMS-induced activity patterns). These ideas are also supported by data from sedation: Ferrarelli et al. (2010), found brief, restricted TMS-evoked responses similar to those seen during sleep in people who were rendered unconscious by the anaesthetic midazolam. The same study also showed that significant current scattering (a measure of the propagation of evoked cortical currents) could reliably distinguish awake and sedated states.

Most studies that compared TMS-evoked EEG responses between different states of consciousness have investigated spatial activity *patterns*, but a few have examined response *amplitudes* at the stimulation target. As noted above, response amplitudes are greater during nREM sleep than in wakefulness (Massimini et al., 2005). Huber et al. (2013) found that the response amplitude increases with prolonged wakefulness (the duration of wakefulness serves as a proxy measure for tiredness or sleep-pressure). Noreika et al. (2019) found that the amplitude correlates negatively, during wakefulness, with momentary EEG-defined measures of alertness. In other words, cortical excitability (indicated by the EEG amplitude of TMS-evoked responses) increases when alertness is reduced, even during wakefulness. The functional significance of this association remains unclear, but may reflect a reduction in the inhibitory activity required for coordinated neural functioning during wakefulness.

Oscillatory activity is another neural measure that TMS affects in a manner that depends on state of consciousness (Massimini et al., 2007). A TMS pulse can trigger a distinct negative deflection in the EEG signal, whose occasional spontaneous occurrence is typical of nREM sleep. Such deflections are known as K-complexes and are indicative of so-called ‘off-periods’ (transient suppression of the high-frequency oscillations that reflect neural processing). The amplitude of TMS-evoked K-complexes increases with sleep depth, and this, in turn, has been shown (Massimini et al., 2007) to lead to an increase in slow-wave activity, which is believed to be involved in memory consolidation and restorative processes during sleep. Awake vegetative patients show TMS-evoked off-periods that are similar to those seen in healthy nREM sleep (Rosanova et al., 2018), suggesting that despite being awake, these patients’ brains process external stimulation in a similar way to the healthy *sleeping* brain. Furthermore, in healthy people, the TMS-evoked response in nREM sleep is reduced when stimulation is applied during dreams (Nieminen et al., 2016; dreaming does occur during nREM sleep, though less commonly than during REM sleep).

Overall, these findings demonstrate that cortical excitability and patterns of effective connectivity are *state-dependent*. When wakefulness and awareness are reduced, local cortical excitability increases, suggesting a reduction in coordinated inhibitory control processes, and the complexity and spatial extent of effective connectivity decrease, indicating reduced communication between brain regions. Stimulation-induced oscillatory activity is also modulated by the state of consciousness, and typically indicates an increase in ‘off-states’ – a reduction in high-frequency neuronal processing – in states where awareness (the presence of conscious content) is low.

Critically, TMS investigations of effective connectivity are a controlled manipulation, and thus allow causal conclusions about the patterns of brain activity that directly follow stimulation. It is worth noting, though, that when collecting measures of effective connectivity in different states of consciousness, the states themselves (with the exception of anaesthesia) are *not* manipulated in a controlled manner but arise either naturally (sleep) or via brain trauma (DOCs). The conclusions drawn from such investigations may therefore be better characterised as concerning how the causal effects of TMS differ between (or in other words, correlate with) different states of consciousness. Characterizing the causal effects of neural manipulations in different states of consciousness can inform both theoretical perspectives and clinical practice in important ways.

This relationship between effective connectivity and states of consciousness is probably best captured by the recent development of a measure known as the perturbational complexity index, or PCI (Casali et al., 2013). This measure quantifies the overall complexity, over several hundred milliseconds and across all EEG electrodes, of the neural response to perturbation by a TMS pulse. A full explanation of the algorithm used to calculate this index is beyond the scope of the present review (Readers who are interested can find relevant technical details and sources in a footnote3); in general terms, however, the PCI is calculated by converting the spatiotemporal pattern of activity into a 2-dimensional representation, with all information sources (EEG electrodes) on one axis, and time on the other axis. Activation is represented in a binary fashion – every point on the axis system is given a value of either one (activity exceeds a statistical significance threshold at that time and location) or zero (no significant activity). The resulting matrix is then compressed in an analogous manner to compressing computer files into Zip-files. Just as a more complex file yields a larger compressed zip-file, more complex activity patterns yield a higher PCI values. **Figure 1C** provides an intuitive schematic illustration of this procedure and its relation to states of consciousness: The EEG response of a healthy, awake brain to a TMS pulse yields a high PCI value, because it involves activity in a large number of cortical regions (high *integration* of information across different parts of the brain), and develops temporally with a unique spatial pattern at any given point in time (high *differentiation* between different times). In contrast, reduced consciousness may be associated with either a brief and localised response (low integration) or a widespread, long-lasting, but unchanging response across the brain (low differentiation), both of which can be compressed more easily – due to containing less data in the former case and higher redundancy in the latter – and thus yield lower PCI values.

Importantly, the PCI shows graded differences between healthy wakefulness, stages of sleep, and sedation, and is low in DOCs (Casali et al., 2013). Furthermore, it allows accurate classification, distinguishing not only wakefulness vs sleep or anaesthesia in the healthy brain, but different severities of DOC – the PCI values of minimally-conscious patients, though lower than those of healthy people, are consistently higher than those of vegetative patients (Casarotto et al., 2016). The PCI therefore provides a reliable neural measure of state of consciousness, does not require behavioural data, and can be collected at patients’ bedside. A more recent measure, the Fractal Dimension Index, uses the same data but involves a different calculation of effective connectivity; in healthy participants, use of this measure has yielded almost-perfect within-subject discrimination between conscious and unconscious states (de Miras et al., 2019).

In addition to investigating impaired effective connectivity in DOCs, researchers have also examined impairments to structural connectivity in these patients. Brain damage can affect different forms of connectivity, but the interactions between connectivity types in DOCs are only partially understood. To assess structural connectivity, several studies have used diffusion-weighted imaging (DWI), a type of MRI scan that captures the direction and rate of the movement of water molecules. Water diffuses more rapidly in directions aligned with internal structures, such as the white matter axonal tracts that connect brain regions. This directionality of this diffusion is quantified in a measure known as fractional anisotropy, or FA, where higher values indicate that more diffusion occurs along a single axis of movement. High FA at a particular location is thus taken to indicate the presence of an intact white matter tract. DWI has been shown to differentiate between vegetative and minimally conscious patients (Fernandez-Espejo, 2011), with the FA of white matter tracts shown to correlate with residual levels of awareness in DOC patients (Fernandez-Espejo, 2012). Moreover, Bodart et al. (2018) used both MRI and TMS-EEG measures to elucidate the links between structural and effective connectivity in healthy and DOC brains, and found that PCI was highly correlated with FA in both groups.

**Part 2. Contents of consciousness: processes underlying perceptual awareness**

Awareness is one of the essential components of consciousness: when we are conscious, we have subjective experiences of ourselves and the world around us. In Part 1, we discussed the mechanisms that give rise to the states in which conscious content can exist *at all*. But when we are in a conscious state, what processes determine the particular contents of our awareness at any given time?

The majority of research on this issue has focussed on understanding the mechanisms that determine *perceptual awareness*, the momentary subjective experiences associated with perception. Although knowledge of perceptual processes (the transduction and hierarchical processing of sensory stimulus features) informs such investigations in important ways, it is not synonymous with an understanding of awareness. Many of the steps involved in processing the features of sensory stimuli occur without awareness (Rees et al., 2002), and may therefore be necessary conditions for the stimulus reaching awareness, but not constitutive parts of the subjective experience itself.

An appreciation of the difference between perception and perceptual awareness can be gained from considering the many ways that the latter has been investigated. These include the search for factors that distinguish what perceptual information can be processed without awareness (e.g., Nasrallah et al., 2009) from perceptual processing that requires awareness (e.g., Rabagliati et al., 2018), and how processing of similar perceptual information with and without awareness may differ (e.g., Kim & Blake, 2005; Raio et al., 2012; Tooley et al., 2017); the relationship between awareness and cognitive functions such as attention (Bahrami et al., 2007, 2008a, 2008b; Tsuchiya & Koch, 2007) and memory (Soto et al., 2005, 2011); the factors that determine awareness in bistable perception, the fluctuating subjective experience evoked by stimuli that can be perceived in more than one way (Brascamp et al., 2018); and perceptual metacognition, the relationship between objective ability to perform a task and subjective ratings of stimulus visibility or confidence (Fleming et al., 2010; Fleming & Lau, 2014; Lau & Passingham, 2006).

A growing body of research has provided a wealth of behavioural, physiological and neural data on all the above. Reviewing all of these approaches is beyond the scope of any single article; here, we will focus on the last two, bistable perception and metacognition, where promising TMS work has already augmented imaging findings and opened new avenues for research.

A central question that has been a continuous topic of debate is whether awareness arises early or late in the perceptual processing stream (Boly et al., 2017). This question provides a recurring theme throughout our review: We start by surveying effects of TMS on awareness, found when stimulation was applied to both early visual cortex and higher-level regions, then continue to work on bistable perception that has attempted to address its underlying mechanisms – and particularly whether the competition between different conscious percepts is resolved early or late. We wrap up with a discussion of high-level metacognition, and evidence that its neural substrates are not found in early perceptual areas but in the frontal lobe.

*Visual awareness and TMS effects in early visual cortex and higher-level regions*

A central question in concerning perceptual awareness is whether its contents are determined by activity in early sensory cortical regions (e.g., primary visual cortex) or whether those regions feed information into higher-level regions that use this information to construct our conscious internal model of the world (Crick & Koch, 1990, 1995). Recent brain stimulation work has contributed to debates on this issue, by investigating how manipulating neural activity in early visual (occipital) and higher-level (parietal) cortical regions alters the contents of perceptual awareness.

In one of the first investigations of TMS effects on visual processing, Amassian et al. (1989) applied single pulses to the occipital pole, the early cortical region that processes information from the foveal (central) region of the visual field4. TMS pulses delivered 80-100ms after stimulus onset (but not at other latencies) suppressed detection of briefly-presented letters. Beyond clarifying the time lag between visual inputs and their processing in early visual cortex, this finding suggests that early visual cortex is necessary (though not necessarily sufficient) for visual perception: Cutting off the stream of information at this early processing stage prevents conscious detection.

Such suppression abolishes reported stimulus awareness, but objective discrimination performance – e.g., reporting which of two possible stimuli was presented – may remain above-chance (Allen et al. 2014; Koenig & Ro, 2019). This has been termed “TMS-induced blindsight”, implying that the preserved visual performance may arise through similar mechanisms as in blindsight patients, who have damage to early visual cortex but are able to perform visual tasks despite reporting no visual experience (Weiskrantz, 1996). The neural basis of this preserved performance remains controversial (Koenig & Ro, 2019): it may be due to spared residual activity in visual cortex that could be sufficient for stimulus discrimination but not for awareness, but may also be due to alternative visual pathways that bypass early visual cortex (e.g., via the midbrain) and enable stimulus discrimination but do not give rise to conscious experience.

Stimulation of visual cortex is known to trigger phosphenes, brief visual sensations that do not correspond to any visual input (Marg & Rudiak, 1994). Above a certain threshold strength, single TMS pulses to early visual cortex evoke stationery phosphenes, whereas pulses to the part of retinotopic cortex that processes motion, area V5 (which is located laterally to early visual cortex), evoke moving phosphenes. Moving phosphenes are abolished, though, when a supra-threshold V5 pulse is followed 20-40ms later by a sub-threshold pulse to early visual cortex (Pascual-Leone and Walsh, 2001), and a moving phosphene is evoked when a sub-threshold pulse to V5 is followed by a supra-threshold pulse to early visual cortex (Silvanto et al., 2005). These manipulations of the relative timing of paired TMS pulses reveal that conscious experiences may depend on feedback projections to early visual cortex.

So far, we have described effects of TMS to retinotopic cortex on visual perception and awareness. TMS has also been used, however, to demonstrate the involvement of higher-level brain regions in measures of visual awareness. Specifically, TMS to the right intraparietal sulcus has been shown to impair detection of changes between sequentially-presented pairs of visual images (i.e., an increase in change blindness; Beck, Muggleton, Walsh, & Lavie, 2006). Stimulation of the same area also increased Troxler fading – the gradual disappearance of peripheral visual stimuli during continuous viewing (Kanai, Muggleton, & Walsh, 2008). These TMS effects are consistent with either disruption of representations that are held in this parietal region, or with disruption of feedback connections that sustain low-level representations in early visual cortex.

Although findings like those described above contribute to understanding the pathway that visual information must go through in order to reach awareness, they do not clarify what step in the process generates the conscious experience itself (e.g., are we aware of activity in primary visual cortex, or is it just a necessary way-station for information we become aware of via later feedforward/feedback processing steps? Crick & Koch, 1995), nor which processes determine what the specific conscious percept will be. We now turn to investigations of bistable perception, which have attempted to address these questions.

*Bistable perception: questions arising from behavioural and neuroimaging findings*

It has been proposed that to understand perceptual awareness, we need insight into the mechanisms that determine our conscious interpretation (what is experienced and reported) of stimuli that can be perceived in more than one way (Rees et al, 2002). A central phenomenon investigated in this context is *bistable perception*.

Bistable perception is the conscious experience associated with a class of stimuli that can be perceived, with equal validity, in different ways. The most well-known example of a bistable stimulus is the Necker cube, which can be perceived in one of two mutually-exclusive ways (**Figure 2A**). Such percepts are referred to as bistable because our perceptual awareness switches between the two interpretations periodically – one cannot perceive both at the same time. There are many other examples of such stimuli. One that is commonly used in research is the ambiguous structure-from-motion (SFM) sphere, which can be perceived as rotating in either of two possible directions (**Figure 2B**). Another widely-investigated phenomenon is binocular rivalry, in which bistability arises not from stimulus ambiguity but from competition between inputs: When each eye is shown a sufficiently-different image at corresponding retinal locations, the two images are not fused but instead compete for dominance, with awareness alternating between the two (Carmel, Arcaro et al., 2010; **Figure 2C**). Bistable stimuli can also be found in audition (e.g., Denham et al., 2018) and even in olfaction (Zhou & Chen, 2009), although the majority of research on this topic has used visual stimuli.

Consciousness researchers are interested in bistable perception because it dissociates external stimulation from perceptual awareness: Our experience keeps changing even though the external stimulus does not; the switch from one conscious percept to the other is therefore internal. Specifically, bistable perception demonstrates competition for awareness between different neural representations of the same stimulus. But at what processing level do such alternations in awareness occur? Are they due to events in early visual cortex or at higher-level regions? Understanding how this works may provide a window into how the brain constructs the subjective experience associated with sensory input.

Dominance in binocular rivalry is influenced by low-level physical stimulus features, such as contrast, which are processed in early visual cortex (Brascamp et al., 2015), but also by high-level features such as context and semantic content, which are processed outside early visual cortex (for reviews, see Klink et al., 2012; Wolf et al., 2011). Furthermore, rivalry is not necessarily between the eyes, but can occur between patchwork patterns that are grouped *across* the eyes to create coherent conscious percepts (Kovacs et al., 1996), and between patterns that alternate rapidly between eyes (Logothetis et al., 1996). Such findings suggest that the competition occurs at a neural locus that follows binocular integration and thus is further in the processing stream than the monocular representations of neurons in primary visual cortex.

Neural data have also painted a confusing picture: single-neuron recordings in monkeys have shown that many stimulus-selective neurons throughout the brain continue to respond to their preferred stimulus during both dominance and suppression in binocular rivalry. What changes across regions is the proportion of neurons whose spiking activity rises and falls in time with the dominant percept – this proportion is very low in primary visual cortex and becomes gradually larger when examining higher levels in the visual processing stream, suggesting that dominance arises in those late processing stages (Logothetis & Schall, 1989; Leopold & Logothetis, 1996, 1999).

On the other hand, fMRI findings in humans have shown a correlation between activity in primary visual cortex and the current dominant percept (Plonsky et al., 2000; Tong & Engel, 2001). Furthermore, activity in temporal-lobe regions that are specifically sensitive to face and house images fluctuates in a time-locked manner to the dominance of their preferred stimuli in rivalry: activity in the fusiform face area (FFA) rises when a face is dominant and falls when it is suppressed, and activity in the parahippocampal place area (PPA) rises when a house is dominant and falls when it is suppressed. This activity is similar to that observed during physically-alternating stimulus presentations, indicating the competition between stimuli is resolved at a lower level (Tong et al., 1998).

The apparent contradiction between single-neuron findings (implicating higher-level regions) and fMRI (implicating primary visual cortex) was resolved when Maier et al. (2008) used single-neuron and fMRI recordings in the same monkeys and found that fMRI indeed reflected conscious perception in primary visual cortex, whereas single neuron spiking did not. Rather than clarifying where binocular rivalry is resolved, however, this finding highlights the importance of acknowledging that these measures are sensitive to different aspects of neural activity: Single-cell spiking rates measure action potentials, and are vulnerable to sampling biases as recording is limited to a subset of the neurons in a region. Conversely, the blood-oxygen-level-dependent (BOLD) fMRI signal reflects oxygenated blood flow, a proxy measure for neuronal activity, in a spatially-defined region, and is influenced by both excitatory and inhibitory activity in cells and synapses; a large fMRI signal may arise when spiking increases, but also when there is no change in spiking (e.g., if increased excitation and inhibition cancel each other out) or even a reduction in spiking (due to increased inhibition). It remains unclear, therefore, which of these measures – spiking or BOLD – is directly associated with the processes that determine the dominant precept. The early visual activity reflected in fMRI signals may determine conscious content, which then propagates to higher-level regions where it influences the kinds of activity reflected in spiking. Conversely, the early activity observed with fMRI may reflect feedback from higher-level regions that determine awareness.

This uncertainty has contributed to current suggestions that bistability is probably a complex phenomenon, arising not from activity in a single area, but from interactions between cortical regions that may play distinct roles in the process (Blake & Logothetis, 2002; Brascamp et al., 2018). Specific regions may *contain* clusters of neurons that represent the different percepts, and these clusters’ spontaneous activity may feature inherent oscillations (via inter-cluster inhibition and intra-cluster facilitation) such that at any given time one is more active than the other. Importantly, however, it is likely that such clusters and their within-cluster processes occur in more than one location in the brain, and that different regions *modulate* *one anothers’* activity. The specific modulatory roles of different regions in bistable perception are a crucial target for investigation.

In line with this, other neuroimaging studies have focussed not on the neural correlates of the current percept, but rather on searching for activity related to the perceptual switches. Using fMRI, Lumer et al. (1998) asked observers to report their dominant percept continuously during binocular rivalry and during a subsequent replay condition (in which stimuli were physically switched at the same time points reported during rivalry). They found a right-lateralized network of frontal and parietal regions that was more active during rivalry alternations than during analogous replay alternations. Similar locations were identified when a different analysis method was applied to the same data (Lumer & Rees, 1999) and in other fMRI studies that used various kinds of bistable stimuli (Brascamp et al., 2015; Frassle et al., 2014; Kleinschmidt et al., 1998; for the neural locations, see **Figure 2D**). These studies suggested that the modulatory activity responsible for switches in bistable perception occurs in high-level regions outside early visual cortex. However, such correlational findings cannot establish that this kind of activity *causes* the switches; it may trigger them, but may also, for example, be involved in noticing that they happen (i.e., directing attention to them), or arise as an epiphenomenal by-product of processes that are entirely resolved in early visual cortex, whose outcome is then propagated to other regions. As noted above, insights into causal mechanisms can be gleaned from brain stimulation. We now turn, therefore, to studies that have investigated how administering TMS to various neural loci modulates bistable perception.

*Using TMS to elucidate causal cortical influences on bistable perception*

TMS has been used to probe the involvement of various brain regions in the dynamics of bistable perception. Pearson et al. (2007) targeted early visual cortex: Single pulses applied online, during viewing of binocular rivalry, increased the probability of a switch (or, in other words, shortened current dominance durations). Similar stimulation had no effect, however, when it was applied during rivalry between stimuli that were rapidly swapped between the eyes (a type of rivalry previously demonstrated by Logothetis et al., 1996). These findings confirm a causal role for early visual cortex in ‘conventional’ binocular rivalry – TMS may have disrupted processing of early stimulus representations, either by differentially affecting one of them (e.g., weakening the current dominant percept or boosting the suppressed percept) or altering the interaction between the neural clusters representing each percept. The absence of an effect for eye-swapping rivalry, however, supports multi-level views of rivalry: similar phenomenal outcomes (dominance alternations) may arise from neural substrates that are either distributed across different regions, or arise from similar types of interactions that occur in different regions for different bistable stimuli.

These results converge with TMS findings described earlier, showing that TMS to parietal cortex can alter other awareness-related phenomena (change detection and Troxler fading), and with the fMRI findings implicating a fronto-parietal network in the perceptual alternations of bistable stimuli. Together, these findings suggested the necessity of testing the effects that TMS to this network would have on bistability. To investigate parietal involvement in binocular rivalry, Carmel, Walsh et al. (2010) applied offline TMS with inhibitory parameters to the right superior parietal lobule (SPL), the parietal location that had shown maximal fMRI activation during rivalry switches (Lumer et al., 1998). This stimulation caused subsequent switching in binocular rivalry to become faster, *shortening* the durations of individual dominant percepts. Rather than suggesting that this parietal region triggers switches – in which case, inhibitory TMS should have increased dominance durations – this finding is therefore consistent with the idea (mentioned earlier in the context of other parietal TMS effects) that parietal activity may sustain and stabilise perception, perhaps through feedback to lower-level visual areas.

In a study that was published in the same journal issue, however, Kanai et al. (2010) also applied offline TMS to SPL, and found that doing so *prolonged* percept durations for a bistable structure-from-motion (SFM) sphere. This apparent contradiction between studies may have been due to the use of different bistable stimuli, or different TMS parameters (although both were inhibitory), or – most interestingly – the fact that the stimulated locations were not identical: Whereas Carmel, Walsh et al. (2010) had based their site selection on fMRI activation coordinates, Kanai et al. (2010) selected stimulation sites in which a structural MRI measure (grey matter volume) had correlated with bistable percept durations. Kanai et al.’s (2010) site, while still within the SPL (and within the activation ‘blob’ seen in previous fMRI studies), was 3cm posterior to that of Carmel, Walsh et al. (2010; **Figure 2D**).

To address this apparent contradiction, the authors of these two studies conducted a joint study (Kanai, Carmel et al., 2011), which used the same bistable SFM sphere and the same TMS parameters as in Kanai et al. (2010), but applied TMS to the more anterior site used by Carmel, Walsh et al. (2010). Under these conditions, TMS *shortened* percept durations, as it had previously when applied to that location (Carmel, Walsh et al., 2010). These results confirmed the existence of sub-regions with distinct functional properties within parietal cortex – TMS to one location prolongs dominance, whereas TMS to the other shortens it independently of the specific bistable stimulus and inhibitory stimulation parameters.

Kanai, Carmel et al. (2011) suggested that taken together, these findings can be accommodated within the theoretical framework known as predictive coding, which provides a compelling account of the underlying mechanism. Predictive coding views the brain as perpetually generating and maintaining models of the likely causes of sensory information (Hohwy, 2008; Clark, 2013). In perception, these models operate in a top-down fashion as predictions of incoming sensory information. These predictions are what we are aware of; any discrepancy between predictions and sensory inputs generates a prediction error signal, which is used to update the prediction. In bistable perception, according to this framework, the process is perpetual: sensory evidence for a nondominant percept (i.e. prediction error) is always present and eventually outweighs the influence of ongoing predictions representing the dominant percept, leading to a perceptual switch; the process then repeats itself. Kanai, Carmel et al. (2011) propose that the dominant percept (or current prediction) is maintained by the anterior parietal location. Inhibiting this location’s activity with TMS thus weakens the prediction, making it easier for the prediction error generated by the non-dominant percept to cause a switch. Conversely, the posterior site generates prediction error signals computed from sensory data, and inhibiting its activity thus hinders the updating of the current prediction, leading to longer periods of perceptual dominance.

Importantly, this account does not make strong claims about the locus of the percept representations themselves: These are likely to exist in several locations in the visual hierarchy. TMS may affect representations held in the stimulated parietal regions, or alter the way these regions exert top-down modulations on lower parts of the visual hierarchy, including early visual cortex. Furthermore, it is important to note that the theoretical interpretation of the findings is a post-hoc proposal: it accounts for the observed pattern of results nicely, but was not part of the studies’ original rationale; therefore, it has to be substantiated by using it to yield novel, testable predictions.

Further studies using various bistable stimuli have provided generally consistent evidence on the effects of parietal TMS. Zaretskaya et al. (2013) demonstrated that inhibitory TMS to anterior right SPL shortened percept durations for a bistable stimulus that alternates between global-gestalt and local-element percepts, although the effect was limited to the global percept. Using a different ambiguous stimulus, Vernet et al. (2015) found that applying single-pulse TMS to a similar site during the temporal gap between brief stimulus presentations increased the probability of a perceptual switch; interestingly, when TMS was applied to frontal and parietal sites together, this effect was abolished, suggesting a complex (and poorly understood) fronto-parietal interaction in determining perceptual dominance.

One major exception to the emerging pattern is a study by Zaretskaya et al. (2010), who applied TMS to anterior and posterior parietal sites bilaterally while participants viewed binocular rivalry, and found a significant effect only in the right anterior site (which was very close to the anterior parietal site stimulated by Carmel, Walsh et al., 2010, and Kanai, Carmel et al., 2011; **Figure 2D**). This effect, however, was in the opposite direction to those previous studies: TMS *prolonged* dominance durations. This contradiction may be due to methodological differences. Whereas Carmel, Walsh et al. (2010) and Kanai, Carmel et al. (2011) had used group-level coordinates from previous fMRI data, this study used individual participants’ fMRI activation to localise stimulation sites, which varied between participants (it is the *average* location that was similar to the previous studies). Additionally, this study used online TMS (applying it *while* participants were viewing rivalry and reporting their percepts), whereas the previous studies had used offline stimulation. Unlike offline TMS, which is known to cause a lasting inhibitory effect after stimulation (Sandrini et al., 2011), online TMS causes momentary disruption that increases noise in the system (Schwarzkopf et al., 2011). It remains unclear, therefore, how to interpret the contradictory findings5.

In sum, at least six studies have demonstrated effects of right parietal TMS on the dynamics of bistable perception, with mostly-consistent results. Two other studies provide indirect evidence of this region’s causal role in bistable perception: Megumi, Kanai, & Rees (2015) applied dynamic causal modelling to fMRI data, and found that a model corresponding to the predictive-coding-based parietal fractionation proposed by Kanai, Carmel et al. (2011) provided a good account for activation patterns. Schauer et al. (2016) applied TMS to previously-used posterior (Kanai et al., 2010) and anterior (Carmel et al., 2010; Kanai, Carmel et al., 2011) parietal sites while participants completed attentional tasks known to be associated with parietal activity and found no effects, suggesting that the effects of parietal TMS on bistable perception are not due to modulations of attention but rather to a causal role these regions play in determining conscious perception.

*The plot thickens: no-response and no-awareness paradigms in bistable perception*

In studies of bistable perception, participants typically report their current percept continuously or indicate percept switches. This involves a host of processes other than the conscious perception itself – executive control to maintain monitoring of the stimulus in line with task instructions, decision processes about when to make a report, planning and executing the motor response – all of which also (presumably) happen in the brain. To what extent do the neural activity patterns observed in previous studies represent these processes as well, and is it possible to isolate only the processes associated with awareness, removing potential task-related confounds while keeping the bistable stimulus constant, to avoid introducing stimulus-related confounds?

To address this, recent studies have introduced “no-report” paradigms, which use objective measures instead of active subjective reporting to establish when switches are happening (for a review, see Tsuchiya et al., 2015). Frassle et al. (2014), for example, used the microsaccades that involuntarily track a moving stimulus (optokinetic nystagmus, or OKN): They capitalised on the previous finding (Naber et al., 2011) that in binocular rivalry between gratings that move in opposite directions, the OKN’s direction corresponds reliably to participants’ dominance reports. In this study’s fMRI recordings, participants viewed the same stimuli and either reported their percepts or not (in the report condition, the authors verified that the OKN indeed tracked reports; in the no-report condition, dominance switches were estimated from participants’ OKN). Switch-related parietal activity was observed both with and without reports; but the previously-reported frontal activity was only seen when reports were made, suggesting that frontal activity is related to report generation rather than to perceptual awareness.

Even without report, watching bistable stimuli involves cognitive activity related to the switches (e.g., noticing, contemplating; Block, 2019). Brascamp et al. (2015) went a step further, introducing a “no awareness” design: They showed participants a different array of randomly-moving dots in each eye. Various stimulus manipulations enabled assessment of dominance, both with and without awareness. Participants were aware of perceptual switches when differently-coloured dots were presented to each eye, and they reported changes in the dominant colour. Awareness of switches was prevented in a condition where same-colour dots were presented to both eyes. In this condition, participants could not tell that they were viewing two different displays, but the perceived dot density was the same as for a single display, indicating that one display was dominant at any given time. Dominance was assessed by having participants report the direction of brief coherent motion intervals, where different directions were presented to each eye. fMRI recordings showed consistent switch-related activity in early visual cortex. Activity in both frontal and parietal cortex, however, was present when participants were aware of the switches, but negligible when they were not. Brascamp et al. (2015) concluded that perceptual switches do not necessarily require activity in higher-level regions outside the early visual system, and that selection for dominance in rivalry can occur without high-level involvement.

While these are reasonable conclusions, it must be noted that in terms of the search for mechanisms of consciousness, this study may have thrown the awareness baby out with the task-confound bathwater: It is definitely possible that competition between neural representation may take place outside of awareness, and it is plausible that such non-conscious competitive interactions would take place early in the visual processing stream. As noted earlier, activity in early visual areas may indicate stimulus processing, but it remains unknown whether the processes involved in bringing a percept into awareness rely on such activity being propagated further up the processing stream in a feedforward manner, on feedback from other regions, or both. Rivalry between *conscious* representations (and by extension, any visual awareness) may require activity beyond early visual cortex, even without report. A crucial next step, which has not been carried out to date, would therefore be to use TMS to investigate the effects of frontal and parietal stimulation on switching in no-report paradigms. If effects under such conditions are similar to those observed previously, this would serve as strong evidence of higher-level cortical involvement independent of report.

Another potentially-fruitful avenue may be using TMS to assess effective connectivity during bistable perception, in a similar manner to that used in studies of states of consciousness and other cognitive functions. Ruff et al. (2008), for example, found that parietal TMS modulated neural activity (measured with fMRI) in distant lower-level visual areas. A similar approach could be applied to test whether the causal influence of parietal areas regulates rivalry competition in the early visual cortex.

*Frontal activity and metacognition*

Finally, it is worth noting a different approach to assessing perceptual awareness that has been associated with frontal activity: Metacognition, or thinking about thinking, is operationalised in perception as the relationship between objective task performance and subjective ratings of confidence or stimulus visibility. Many researchers consider metacognitive measures as a meaningful proxy for conscious perception – when perceptual content reaches awareness, confidence and performance should covary (Persaud, McLeod, & Cowey, 2007; Fleming & Dolan, 2014). According to this view, brain regions associated with monitoring perceptual decisions and producing confidence judgments may also be those that produce subjective awareness.

Fleming et al. (2010) used a simple perceptual decision task and manipulated its difficulty so that objective performance was similar for all participants. They also collected a confidence rating on each trial, and found that metacognitive accuracy (how well participants were able to assess their own task performance) and an MRI-derived structural measure (grey matter volume) correlated in a frontal region (specifically, bilateral anterior prefrontal cortex). Using a different perceptual task, but again keeping objective performance constant, Lau & Passingham (2006) found that fMRI activity in a similar frontal area covaried with reported stimulus visibility.

To assess the causal significance of such findings, Rounis et al. (2010) investigated the effect of inhibitory TMS to the above frontal region on subsequent perceptual decisions and metacognitive judgments. They found that perceptual accuracy was not impaired by TMS, but metacognitive accuracy was. In a close replication of this study, however, Bor et al. (2017) failed to find evidence for a TMS-induced impairment to either task performance of metacognition. The mixed evidence means the jury is still out on whether prefrontal cortex is causally involved in judgments of metacognition.

**Concluding remarks**

The physical basis of consciousness remains one of the great mysteries confronting science. The most fundamental question concerns how the activity of a physical system like the brain can give rise to subjective experience. Different versions of this question have been asked for many centuries; the most recent is its branding as ‘the hard problem of consciousness’ by the philosopher David Chalmers (Chalmers, 1995). It remains a philosophical problem – current science does not have the tools to address it.

In the current review, we have addressed a few of the so-called ‘easy’ problems – easy not because they are easy to solve (they are not), but because we can recognize an answer when we find one, and have some idea of how to go about looking for it. These problems concern the neural functions and activities that give rise to consciousness-related phenomena. We don’t know why they do so, but we can make progress by characterising them. Indeed, researchers have made some progress on some of these problems: we are developing an understanding of the kinds of brain activity that underlie different states of consciousness, and of the neural systems whose activity leads to subjective conscious experiences.

A substantial portion of our current understanding has been obtained with the use of neuroimaging methods, in the search for the neural correlates of consciousness. As this review has hopefully clarified, such neural evidence can be meaningfully augmented by brain stimulation methods, which enable the experimental manipulation of brain activity and can yield conclusions about causal mechanisms. Our review has also highlighted that each new answer leads to further questions. The field is still in its infancy, and although the challenge of understanding consciousness is formidable, the progress made in just a few decades should be a cause for optimism.

**Notes**

1) Other forms of brain stimulation have also been used, but fall outside the scope of the present review: deep brain stimulation, which involves invasive implanting of electrodes in the brain, has shown promise in clinical settings, where stimulating subcortical structures has increased arousal in unresponsive patients (e.g., Schiff et al., 2007). Other research has used transcranial electric stimulation, a non-invasive method that applies a weak current through electrodes attached to the head; however, the efficacy and reliability of this method have been called into question (Horvath et al., 2015; Parkin et al., 2015; Vöröslakos et al., 2018).

2) Effective connectivity can also be estimated through modelling approaches applied to neuroimaging data. Dynamic causal modelling (DCM), for example, is a technique in which a biologically plausible mathematical model is tested for its ability to explain observed neural signals, by postulating changes in the activity of defined neuronal regions and their effects on activity in other regions (Friston, 2009). Such methods still rely on inferences from correlational data; here, we focus on findings from TMS studies, which provide data allowing direct inference about causal interactions in the brain.

3) Operationally, PCI is defined as the normalized algorithmic complexity of the spatiotemporal pattern of activation triggered by a TMS pulse (Casali et al., 2013). The computation of PCI involves performing source modelling and nonparametric statistics to create a binary matrix of significant neural sources of the activation caused by the TMS (Casali et al., 2010). This matrix is then compressed to determine its information content using the Lempel-Ziv complexity index (Lempel & Ziv, 1976). This complexity index is then normalised by the source entropy of the binary matrix of significant sources, which is a measure of the randomness of the spatiotemporal pattern of activity (Kaspar & Schuster, 1987). The PCI is therefore high when the Lempel-Ziv index of complexity is high and the source entropy (randomness) is low.

4) Many studies that have stimulated the occipital pole suggest the observed effects were due to stimulation of V1, primary visual cortex. However, the retinotopic representations of the fovea in primary, secondary, and tertiary visual cortex (areas V1, V2, and V3) all converge onto the occipital pole. This part of retinotopic cortex is thus referred to as the foveal confluence, and stimulating it cannot be assumed to affect V1 alone. Here, we will therefore simply refer to such stimulation as targeting early visual cortex.

5) We note that in a recent – as yet unpublished – study in our lab, we directly compared the effects of offline and online TMS to right anterior SPL on binocular rivalry. In two experiments, we replicated the previously-reported effect of offline stimulation (shortening of dominance durations); however, we did not find any effect of online stimulation. As noted, these results have not been through peer-review, and should thus be taken with caution.

**Acknowledgments**

DC is supported by the European Research Council (Co-I, ERC Advanced Grant XSPECT – DLV-692739). CAB is supported by a Neurological Foundation of New Zealand Repatriation Fellowship (1839 RF). We would like to acknowledge Dr David Lloyd for his assistance with figure preparation.

**References**

Alkire MT, Hudetz AG, Tononi G. 2008. Consciousness and anesthesia. Science. 322(5903):876--880.

Allen CPG, Sumner P, Chambers CD. 2014. The timing and neuroanatomy of conscious vision as revealed by TMS-induced blindsight. Journal of Cognitive Neuroscience. 26:1507-1518.

Amassian VE, Cracco RQ, Maccabee PJ, Cracco JB, Rudell A, Eberle L. 1989. Suppression of visual perception by magnetic coil stimulation of human occipital cortex. Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section. 74:458-462.

Arzi A, Shedlesky L, Ben-Shaul M, Nasser K, Oksenberg A, Hairston IS, Sobel N. 2012. Humans can learn new information during sleep. Nature neuroscience. 15(10):1460.

Atienza M, Cantero J, Escera C. 2001. Auditory information processing during human sleep as revealed by event-related brain potentials. Clinical Neurophysiology. 112(11):2031--2045.

Baars BJ, Ramsøy TZ, Laureys S. 2003. Brain, conscious experience and the observing self. Trends in neurosciences. 26(12):671-675.

Bahrami B, Carmel D, Walsh V, Rees G, Lavie N. 2008a. Unconscious orientation processing depends on perceptual load. Journal of Vision, 8(3): 12-12.

Bahrami B, Carmel D, Walsh V, Rees G, Lavie N. 2008b. Spatial attention can modulate unconscious orientation processing. Perception, 37(10): 1520-1528.

Bahrami B, Lavie N, Rees G. 2007. Attentional load modulates responses of human primary visual cortex to invisible stimuli. Current Biology, 17(6): 509-513.

Bareham C, Allanson J, Roberts N, Hutchinson PJ, Pickard JD, Menon DK, Chennu S. 2018. Longitudinal bedside assessments of brain networks in disorders of consciousness: case reports from the field. Frontiers in neurology. 9:676.

Bareham CA, Bekinschtein TA, Scott SK, Manly T. 2015. Does left-handedness confer resistance to spatial bias? Sci Rep. 5:9162.

Bareham CA, Manly T, Pustovaya OV, Scott SK, Bekinschtein TA. 2014. Losing the left side of the world: rightward shift in human spatial attention with sleep onset. Sci Rep. 4:5092.

Beck DM, Muggleton N, Walsh V, Lavie N. 2006. Right parietal cortex plays a critical role in change blindness. Cerebral Cortex. 16(5):712-717.

Blake R, Logothetis NK. 2002. Visual competition. Nature Reviews Neuroscience. 3(1):13-21.

Block N, Carmel D, Fleming SM, Kentridge RW, Koch C, Lamme VA, Lau H, Rosenthal D. 2014. Consciousness science: real progress and lingering misconceptions. Trends in cognitive sciences. 18(11):556-557.

Block, N., 2019. What is wrong with the no-report paradigm and how to fix it. *Trends in cognitive sciences*, *23*(12), pp.1003-1013.

Bodart O, Amico E, Gómez F, Casali AG, Wannez S, Heine L, Thibaut A, Annen J, Boly M, Casarotto S. 2018. Global structural integrity and effective connectivity in patients with disorders of consciousness. Brain stimulation. 11(2):358-365.

Bodart O, Gosseries O, Wannez S, Thibaut A, Annen J, Boly M, Rosanova M, Casali AG, Casarotto S, Tononi G et al. 2017. Measures of metabolism and complexity in the brain of patients with disorders of consciousness. Neuroimage Clin. 14:354-362.

Functional networks in disorders of consciousness. Seminars in neurology; 2017: Thieme Medical Publishers.

Boly M, Massimini M, Tsuchiya N, Postle BR, Koch C, Tononi G. 2017. Are the neural correlates of consciousness in the front or in the back of the cerebral cortex? Clinical and neuroimaging evidence. Journal of Neuroscience, 37(40): 9603-9613.

Bor D, Schwartzman DJ, Barrett AB, Seth AK. 2017. Theta-burst transcranial magnetic stimulation to the prefrontal or parietal cortex does not impair metacognitive visual awareness. PLoS ONE. 12.

Boveroux P, Bonhomme V, Boly M, Vanhaudenhuyse A, Maquet P, Laureys S. 2008. Brain function in physiologically, pharmacologically, and pathologically altered states of consciousness. International anesthesiology clinics. 46(3):131-146.

Brascamp J, Blake R, Knapen T. 2015. Negligible fronto-parietal BOLD activity accompanying unreportable switches in bistable perception. Nature Neuroscience. 18:1672-1678.

Brascamp J, Klink P, Levelt WJ. 2015. The ‘laws’ of binocular rivalry: 50 years of Levelt’s propositions. Vision research. 109:20-37.

Brascamp J, Sterzer P, Blake R, Knapen T. 2018. Multistable Perception and the Role of the Frontoparietal Cortex in Perceptual Inference. Annual review of psychology. 69:77-103.

Brown RJ, Norcia AM. 1997. A method for investigating binocular rivalry in real-time with the steady-state VEP. Vision Research. 37:2401-2408.

Carmel D, Arcaro M, Kastner S, Hasson U. 2010. How to create and use binocular rivalry. JoVE (Journal of Visualized Experiments).(45):e2030.

Carmel D, Walsh V, Lavie N, Rees G. 2010. Right parietal TMS shortens dominance durations in binocular rivalry. Current Biology. 20:R799-R800.

Casali AG, Gosseries O, Rosanova M, Boly M, Sarasso S, Casali KR, Casarotto S, Bruno MA, Laureys S, Tononi G et al. 2013. A theoretically based index of consciousness independent of sensory processing and behavior. Science translational medicine. 5(198):198ra105. eng.

Casarotto S, Comanducci A, Rosanova M, Sarasso S, Fecchio M, Napolitani M, Pigorini A, G. Casali A, Trimarchi PD, Boly M et al. 2016. Stratification of unresponsive patients by an independently validated index of brain complexity. Annals of Neurology.

Chalmers, DJ. 1995. Facing up to the problem of consciousness. Journal of consciousness studies, 2(3), 200-219.

Chennu S, Annen J, Wannez S, Thibaut A, Chatelle C, Cassol H, Martens G, Schnakers C, Gosseries O, Menon D et al. 2017. Brain networks predict metabolism, diagnosis and prognosis at the bedside in disorders of consciousness. Brain. 140(8):2120-2132.

Chennu S, Bekinschtein TA. 2012. Arousal modulates auditory attention and awareness: insights from sleep, sedation and disorders of consciousness. Frontiers in Psychology. 3(00065).

Chennu S, Finoia P, Kamau E, Allanson J, Williams GB, Monti MM, Noreika V, Arnatkeviciute A, Canales-Johnson A, Olivares F et al. 2014. Spectral signatures of reorganised brain networks in disorders of consciousness. PLOS Computational Biology. 10(10):e1003887.

Chennu S, O’Connor S, Adapa R, Menon DK, Bekinschtein TA. 2016. Brain Connectivity Dissociates Responsiveness from Drug Exposure during Propofol-Induced Transitions of Consciousness. PLOS Computational Biology. 12(1):e1004669.

Clark A. 2013. Whatever next? Predictive brains, situated agents, and the future of cognitive science. Behavioral and Brain Sciences. 36:181-204.

Colombo MA, Napolitani M, Boly M, Gosseries O, Casarotto S, Rosanova M, Brichant J-F, Boveroux P, Rex S, Laureys S. 2019. The spectral exponent of the resting EEG indexes the presence of consciousness during unresponsiveness induced by propofol, xenon, and ketamine. NeuroImage. 189:631-644.

Cote KA. 2002. Probing awareness during sleep with the auditory odd-ball paradigm. International Journal of Psychophysiology. 46(3):227 - 241.

Crick F, Koch C. 1990. Towards a neurobiological theory of consciousness. Seminars in the neurosciences. 2:263-275.

Crick F, Koch C. 1995. Are we aware of neural activity in primary visual cortex? Nature. 375(6527):121-123.

de Miras JR, Soler F, Iglesias-Parro S, Ibáñez-Molina A, Casali A, Laureys S, Massimini M, Esteban F, Navas J, Langa J. 2019. Fractal dimension analysis of states of consciousness and unconsciousness using transcranial magnetic stimulation. Computer methods and programs in biomedicine. 175:129-137.

Dehaene S, Changeux J-P, Naccache L, Sackur J, Sergent C. 2006. Conscious, preconscious, and subliminal processing: a testable taxonomy. Trends in Cognitive Sciences. 10(5):204-211.

Del Cul A, Baillet S, Dehaene S. 2007. Brain Dynamics Underlying the Nonlinear Threshold for Access to Consciousness. PLoS Biology. 5(10):e260.

Demertzi A, Antonopoulos G, Heine L, Voss HU, Crone JS, de Los Angeles C, Bahri MA, Di Perri C, Vanhaudenhuyse A, Charland-Verville V et al. 2015. Intrinsic functional connectivity differentiates minimally conscious from unresponsive patients. Brain. 138(9):2619-2631.

Denham SL, Farkas D, Van Ee R, Taranu M, Kocsis Z, Wimmer M, Carmel D, Winkler I. 2018. Similar but separate systems underlie perceptual bistability in vision and audition. Scientific reports. 8(1):1-10.

Fernández-Espejo D, Bekinschtein T, Monti MM, Pickard JD, Junque C, Coleman MR, Owen AM. 2011. Diffusion weighted imaging distinguishes the vegetative state from the minimally conscious state. NeuroImage. 54(1):103 - 112.

Fernandez-Espejo D, Soddu A, Cruse D, Palacios EM, Junque C, Vanhaudenhuyse A, Rivas E, Newcombe V, Menon DK, Pickard JD et al. 2012. A role for the default mode network in the bases of disorders of consciousness. Ann Neurol. 72(3):335-343.

Ferrarelli F, Massimini M, Sarasso S, Casali A, Riedner BA, Angelini G, Tononi G, Pearce RA. 2010. Breakdown in cortical effective connectivity during midazolam-induced loss of consciousness. Proc Natl Acad Sci U S A. 107(6):2681-2686.

Fleming SM, Dolan RJ. 2014. The neural basis of metacognitive ability. The Cognitive Neuroscience of Metacognition. Berlin, Heidelberg: Springer Berlin Heidelberg. p. 245-265.

Fleming SM, Lau HC. 2014. How to measure metacognition. Frontiers in human neuroscience, 8: 443.

Fleming, S.M., Weil, R.S., Nagy, Z., Dolan, R.J. and Rees, G., 2010. Relating introspective accuracy to individual differences in brain structure. *Science*, *329*(5998), pp.1541-1543.

Frässle S, Sommer J, Jansen A, Naber M, Einhäuser W. 2014. Binocular rivalry: Frontal activity relates to introspection and action but not to perception. Journal of Neuroscience. 34:1738-1747.

Giacino JT, Kalmar K, Whyte J. 2004. The JFK Coma Recovery Scale-Revised: Measurement characteristics and diagnostic utility1. Archives of physical medicine and rehabilitation. 85(12):2020-2029.

Goupil L, Bekinschtein TA. 2012. Cognitive processing during the transition to sleep. Arch Ital Biol. 150(2-3):140-154.

Hallett M. 2007. Transcranial magnetic stimulation: a primer. Neuron. 55(2):187-199.

Hobson JA, Pace-Schott EF. 2002. The cognitive neuroscience of sleep: neuronal systems, consciousness and learning. Nature Reviews Neuroscience. 3(9):679-693.

Hohwy J, Roepstorff A, Friston K. 2008. Predictive coding explains binocular rivalry: An epistemological review. Cognition. 108:687-701.

Horvath JC, Forte JD, Carter O. 2015. Quantitative review finds no evidence of cognitive effects in healthy populations from single-session transcranial direct current stimulation (tDCS). Brain stimulation. 8(3):535-550.

Huang, Y.Z., Edwards, M.J., Rounis, E., Bhatia, K.P. and Rothwell, J.C., 2005. Theta burst stimulation of the human motor cortex. *Neuron*, *45*(2), pp.201-206.

Huber R, Esser SK, Ferrarelli F, Massimini M, Peterson MJ, Tononi G. 2007. TMS-induced cortical potentiation during wakefulness locally increases slow wave activity during sleep. PloS one. 2(3).

Huber R, Mäki H, Rosanova M, Casarotto S, Canali P, Casali AG, Tononi G, Massimini M. 2013. Human cortical excitability increases with time awake. Cerebral cortex. 23(2):1-7.

Kanai R, Bahrami B, Rees G. 2010. Human parietal cortex structure predicts individual differences in perceptual rivalry. Current Biology. 20:1626-1630.

Kanai R, Carmel D, Bahrami B, Rees G. 2011. Structural and functional fractionation of right superior parietal cortex in bistable perception. Current Biology. 21:R106-R107.

Kanai R, Muggleton NG, Walsh V. 2008. TMS over the intraparietal sulcus induces perceptual fading. Journal of Neurophysiology. 100:3343-3350.

Kim CY, Blake R. 2005. Psychophysical magic: rendering the visible ‘invisible’. Trends in cognitive sciences, 9(8): 381-388.

King JR, Sitt JD, Faugeras F, Rohaut B, El Karoui I, Cohen L, Naccache L, Dehaene S. 2013. Information sharing in the brain indexes consciousness in noncommunicative patients. Curr Biol. 23(19):1914-1919.

Kleinschmidt A, Büchel C, Zeki S, Frackowiak RSJ. 1998. Human brain activity during spontaneously reversing perception of ambiguous figures. Proceedings of the Royal Society B: Biological Sciences. 265:2427-2433.

Klimova M. 2014. What is lost during dreamless sleep: the relationship between neural connectivity patterns and consciousness. Journal of European Psychology Students. 5(3).

Klink P, van Wezel RJA, van Ee R. 2012. United we sense, divided we fail: context-driven perception of ambiguous visual stimuli. Philosophical Transactions of the Royal Society B: Biological Sciences. 367(1591):932-941.

Knapen T, Brascamp J, Pearson J, van Ee R, Blake R. 2011. The role of frontal and parietal brain areas in bistable perception. Journal of Neuroscience. 31:10293-10301.

Koch C, Tsuchiya N. 2007. Attention and consciousness: two distinct brain processes. Trends in cognitive sciences, 11(1): 16-22.

Koenig L, Ro T. 2019. Dissociations of conscious and unconscious perception in TMS-induced blindsight. Neuropsychologia. 128:215-222.

Koivisto M, Lähteenmäki M, Kaasinen V, Parkkola R, Railo H. 2014. Overlapping activity periods in early visual cortex and posterior intraparietal area in conscious visual shape perception: A TMS study. NeuroImage. 84:765-774.

Kouider S, Andrillon T, Barbosa LS, Goupil L, Bekinschtein TA. 2014. Inducing task-relevant responses to speech in the sleeping brain. Curr Biol. 24(18):2208-2214.

Kovács I, Papathomas TV, Yang M, Fehér Á. 1996. When the brain changes its mind: Interocular grouping during binocular rivalry. Proceedings of the National Academy of Sciences of the United States of America. 93:15508-15511.

Lau, H.C. and Passingham, R.E., 2006. Relative blindsight in normal observers and the neural correlate of visual consciousness. *Proceedings of the National Academy of Sciences*, *103*(49), pp.18763-18768.

Laureys S. 2005. The neural correlate of (un)awareness: lessons from the vegetative state. Trends Cogn Sci. 9(12):556-559.

Laureys S, Berré J, Goldman S. 2001. Cerebral function in coma, vegetative state, minimally conscious state, locked-in syndrome, and brain death. Yearbook of Intensive Care and Emergency Medicine 2001. Springer; p. 386-396.

Lee M, Baird B, Gosseries O, Nieminen JO, Boly M, Postle BR, Tononi G, Lee S-W. 2019. Connectivity differences between consciousness and unconsciousness in non-rapid eye movement sleep: a TMS–EEG study. Scientific reports. 9(1):1-9.

Leopold DA, Logothetis NK. 1996. Activity changes in early visual cortex reflect monkeys' percepts during binocular rivalry. Nature. 379(6565):549-553.

Leopold DA, Logothetis NK. 1999. Multistable phenomena: Changing views in perception. Trends in Cognitive Sciences. Elsevier Current Trends. p. 254-264.

Logothetis NK, Leopold DA, Sheinberg DL. 1996. What is rivalling during binocular rivalry? Nature. 380(6575):621-624.

Logothetis NK, Schall JD. 1989. Neuronal correlates of subjective visual perception. Science. 245(4919):761-763.

Lumer ED, Friston KJ, Rees G. 1998. Neural correlates of perceptual rivalry in the human brain. Science. 280:1930-1934.

Lumer ED, Rees G. 1999. Covariation of activity in visual and prefrontal cortex associated with subjective visual perception. Proceedings of the National Academy of Sciences of the United States of America. 96:1669-1673.

Maier A, Wilke M, Aura C, Zhu C, Frank QY, Leopold DA. 2008. Divergence of fMRI and neural signals in V1 during perceptual suppression in the awake monkey. Nature neuroscience. 11(10):1193-1200.

Marg E, Rudiak D. 1994. Phosphenes induced by magnetic stimulation over the occipital brain: Description and probable site of stimulation. Optometry and Vision Science. 71:301-311.

Massimini M, Ferrarelli F, Esser SK, Riedner BA, Huber R, Murphy M, Peterson MJ, Tononi G. 2007. Triggering sleep slow waves by transcranial magnetic stimulation. Proceedings of the National Academy of Sciences. 104(20):8496-8501.

Massimini M, Ferrarelli F, Huber R, Esser SK, Singh H, Tononi G. 2005. Breakdown of cortical effective connectivity during sleep. Science. 309(5744):2228-2232.

Massimini M, Ferrarelli F, Murphy M, Huber R, Riedner B, Casarotto S, Tononi G. 2010. Cortical reactivity and effective connectivity during REM sleep in humans. Cognitive neuroscience. 1(3):176-183.

Megumi F, Bahrami B, Kanai R, Rees G. 2015. Brain activity dynamics in human parietal regions during spontaneous switches in bistable perception. NeuroImage. 107:190-197.

Michel M, Beck D, Block N, Blumenfeld H, Brown R, Carmel D, Carrasco M, Chirimuuta M, Chun M, Cleeremans A. 2019. Opportunities and challenges for a maturing science of consciousness. Nature human behaviour. 3(2):104-107.

Miller SM, Liu GB, Ngo TT, Hooper G, Riek S, Carson RG, Pettigrew JD. 2000. Interhemispheric switching mediates perceptual rivalry. Current Biology. 10:383-392.

Naber M, Frässle S, Einhäuser, W. 2011. Perceptual rivalry: reflexes reveal the gradual nature of visual awareness. *PLoS One*, *6*(6): p.e20910.

Nasrallah M, Carmel D, Lavie N. 2009. Murder, she wrote: Enhanced sensitivity to negative word valence. Emotion, 9(5): 609.

Newcombe VF, Williams GB, Scoffings D, Cross J, Carpenter TA, Pickard JD, Menon DK. 2010. Aetiological differences in neuroanatomy of the vegetative state: insights from diffusion tensor imaging and functional implications [Research Support, Non-U.S. Gov't]. J Neurol Neurosurg Psychiatry. 81(5):552-561.

Nieminen JO, Gosseries O, Massimini M, Saad E, Sheldon AD, Boly M, Siclari F, Postle BR, Tononi G. 2016. Consciousness and cortical responsiveness: a within-state study during non-rapid eye movement sleep. Scientific reports. 6(1):1-10.

Noreika V, Kamke MR, Canales-Johnson A, Chennu S, Bekinschtein TA, Mattingley JB. 2019. Alertness fluctuations during task performance modulate cortical evoked responses to transcranial magnetic stimulation. bioRxiv.155754.

Ogilvie RD. 2001. The process of falling asleep. Sleep medicine reviews. 5(3):247-270.

Parkin BL, Ekhtiari H, Walsh VF. 2015. Non-invasive human brain stimulation in cognitive neuroscience: a primer. Neuron. 87(5):932-945.

Pascual-Leone A, Walsh V. 2001. Fast backprojections from the motion to the primary visual area necessary for visual awareness. Science. 292:510-512.

Pascual-Leone A, Walsh V, Rothwell J. 2000. Transcranial magnetic stimulation in cognitive neuroscience–virtual lesion, chronometry, and functional connectivity. Current opinion in neurobiology. 10(2):232-237.

Pearson J, Brascamp J. 2008. Sensory memory for ambiguous vision. Trends in Cognitive Sciences. 12:334-341.

Pearson J, Tadin D, Blake R. 2007. The effects of transcranial magnetic stimulation on visual rivalry. Journal of Vision. 7:2.

Polonsky A, Blake R, Braun J, Heeger DJ. 2000. Neuronal activity in human primary visual cortex correlates with perception during binocular rivalry. Nature neuroscience. 3(11):1153-1159.

Rabagliati H, Robertson A, Carmel D. 2018. The importance of awareness for understanding language. Journal of Experimental Psychology: General, 147(2): 190.

Raio CM, Carmel D, Carrasco M, Phelps EA. 2012. Nonconscious fear is quickly acquired but swiftly forgotten. Current Biology, 22(12): R477-R479.

Rees G, Kreiman G, Koch C. 2002. Neural correlates of consciousness in humans. Nature Reviews Neuroscience. 3:261-270.

Rosanova M, Casali A, Bellina V, Resta F, Mariotti M, Massimini M. 2009. Natural frequencies of human corticothalamic circuits. J Neurosci. 29(24):7679-7685.

Rosanova M, Fecchio M, Casarotto S, Sarasso S, Casali AG, Pigorini A, Comanducci A, Seregni F, Devalle G, Bodart O. 2018. Sleep-like bistability, loss of causality and complexity in the brain of Unresponsive Wakefulness Syndrome patients. bioRxiv.242644.

Rosanova M, Gosseries O, Casarotto S, Boly M, Casali AG, Bruno MA, Mariotti M, Boveroux P, Tononi G, Laureys S et al. 2012. Recovery of cortical effective connectivity and recovery of consciousness in vegetative patients [Research Support, Non-U.S. Gov't]. Brain. 135(Pt 4):1308-1320.

Rossi, S., Hallett, M., Rossini, P.M., Pascual-Leone, A. and Safety of TMS Consensus Group, 2009. Safety, ethical considerations, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research. *Clinical neurophysiology*, *120*(12), pp.2008-2039.

Rothkegel H, Sommer M, Paulus W. 2010. Breaks during 5 Hz rTMS are essential for facilitatory after effects. Clinical neurophysiology. 121(3):426-430.

Rounis E, Maniscalco B, Rothwell JC, Passingham RE, Lau H. 2010. Theta-burst transcranial magnetic stimulation to the prefrontal cortex impairs metacognitive visual awareness. Cognitive Neuroscience. 1:165-175.

Ruff CC, Bestmann S, Blankenburg F, Bjoertomt O, Josephs O, Weiskopf N, Deichmann R, Driver J. 2008. Distinct causal influences of parietal versus frontal areas on human visual cortex: Evidence from concurrent TMS-fMRI. Cerebral Cortex. 18:817-827.

Sadaghiani S, Scheeringa R, Lehongre K, Morillon B, Giraud A-L, Kleinschmidt A. 2010. Intrinsic connectivity networks, alpha oscillations, and tonic alertness: a simultaneous electroencephalography/functional magnetic resonance imaging study. Journal of Neuroscience. 30(30):10243-10250.

Sämann PG, Wehrle R, Hoehn D, Spoormaker VI, Peters H, Tully C, Holsboer F, Czisch M. 2011. Development of the Brain's Default Mode Network from Wakefulness to Slow Wave Sleep. Cerebral Cortex. 21(9):2082-2093.

Sandrini M, Umiltà C, Rusconi E. 2011. The use of transcranial magnetic stimulation in cognitive neuroscience: a new synthesis of methodological issues. Neuroscience & Biobehavioral Reviews. 35(3):516-536.

Schauer G, Kanai R, Brascamp JW. 2016. Parietal theta burst TMS: Functional fractionation observed during bistable perception not evident in attention tasks. Consciousness and Cognition. 40:105-115.

Schiff, N.D., Giacino, J.T., Kalmar, K., Victor, J.D., Baker, K., Gerber, M., Fritz, B., Eisenberg, B., O’connor, J., Kobylarz, E.J. and Farris, S., 2007. Behavioural improvements with thalamic stimulation after severe traumatic brain injury. *Nature*, *448*(7153), pp.600-603.

Schwarzkopf DS, Silvanto J, Rees G. 2011. Stochastic resonance effects reveal the neural mechanisms of transcranial magnetic stimulation. Journal of Neuroscience. 31(9):3143-3147.

Silvanto J, Cowey A, Lavie N, Walsh V. 2005. Striate cortex (V1) activity gates awareness of motion. Nature neuroscience. 8(2):143-144.

Sitt JD, King JR, El Karoui I, Rohaut B, Faugeras F, Gramfort A, Cohen L, Sigman M, Dehaene S, Naccache L. 2014. Large scale screening of neural signatures of consciousness in patients in a vegetative or minimally conscious state. Brain. 137(8):2258-2270.

Soto D, Heinke D, Humphreys GW, Blanco MJ. 2005. Early, involuntary top-down guidance of attention from working memory. Journal of Experimental Psychology: Human Perception and Performance, 31(2): 248.

Soto D, Mäntylä T, Silvanto J. 2011. Working memory without consciousness. Current Biology, 21(22): R912-R913.

Spoormaker VI, Schröter MS, Gleiser PM, Andrade KC, Dresler M, Wehrle R, Sämann PG, Czisch M. 2010. Development of a Large-Scale Functional Brain Network during Human Non-Rapid Eye Movement Sleep. The Journal of Neuroscience. 30(34):11379-11387.

Stamatakis EA, Adapa RM, Absalom AR, Menon DK. 2010. Changes in resting neural connectivity during propofol sedation [Research Support, Non-U.S. Gov't]. PLoS ONE. 5(12):e14224. eng.

Thielscher A, Reichenbach A, Uurbil K, Uluda K. 2010. The cortical site of visual suppression by transcranial magnetic stimulation. Cerebral Cortex. 20:328-338.

Tong F, Engel SA. 2001. Interocular rivalry revealed in the human cortical blind-spot representation. Nature. 411(6834):195-199.

Tong F, Nakayama K, Vaughan JT, Kanwisher N. 1998. Binocular rivalry and visual awareness in human extrastriate cortex. Neuron. 21:753-759.

Tononi G. 2004. An information integration theory of consciousness. BMC Neuroscience. 5(1):42.

Tononi, G. and Edelman, G.M., 1998. Consciousness and complexity. *science*, *282*(5395), pp.1846-1851.

Tononi G, Srinivasan R, Russell DP, Edelman GM. 1998. Investigating neural correlates of conscious perception by frequency-tagged neuromagnetic responses. Proceedings of the National Academy of Sciences of the United States of America. 95:3198-3203.

Tooley MD, Carmel D, Chapman A, Grimshaw GM. 2017. Dissociating the physiological components of unconscious emotional responses. Neuroscience of Consciousness, 2017(1): nix021.

Tsuchiya N, Wilke M, Frässle S, Lamme VAF. 2015. No-Report Paradigms: Extracting the True Neural Correlates of Consciousness. Trends in Cognitive Sciences. 19:757-770.

Vanhaudenhuyse A, Noirhomme Q, Tshibanda LJ-F, Bruno M-A, Boveroux P, Schnakers C, Soddu A, Perlbarg V, Ledoux D, Brichant J-F et al. 2010. Default network connectivity reflects the level of consciousness in non-communicative brain-damaged patients. Brain. 133(1):161-171.

Vernet M, Brem A-K, Farzan F, Pascual-Leone A. 2015. Synchronous and opposite roles of the parietal and prefrontal cortices in bistable perception: A double-coil TMS–EEG study. Cortex. 64:78-88.

Vöröslakos M, Takeuchi Y, Brinyiczki K, Zombori T, Oliva A, Fernández-Ruiz A, Kozák G, Kincses ZT, Iványi B, Buzsáki G. 2018. Direct effects of transcranial electric stimulation on brain circuits in rats and humans. Nature communications. 9(1):1-17.

Walsh V, Cowey A. 2000. Transcranial magnetic stimulation and cognitive neuroscience. Nature Reviews Neuroscience. 1(1):73-80.

Weiskrantz L. 1996. Blindsight revisited. Current opinion in neurobiology. 6(2):215-220.

Wolf M, Hochstein S. 2011. High-level binocular rivalry effects. Frontiers in Human Neuroscience. 5:129.

Zaretskaya N, Anstis S, Bartels A. 2013. Parietal cortex mediates conscious perception of illusory gestalt. Journal of Neuroscience. 33:523-531.

Zaretskaya N, Thielscher A, Logothetis NK, Bartels A. 2010. Disrupting parietal function prolongs dominance durations in binocular rivalry. Current Biology. 20:2106-2111.

Zhou, W., & Chen, D. (2009). Binaral rivalry between the nostrils and in the cortex. Current Biology, 19(18), 1561–1565. https://doi.org/10.1016/j.cub.2009.07.052

**Figure captions**

**Figure 1. The effect of TMS on EEG activity in different states of consciousness. A. Effective connectivity.** A TMS pulse influences EEG measures of neural activity not only at the site of stimulation (the red ‘hot-spot’ at the centre of the TMS coil), but also in other areas that the stimulated region communicates with (arrows). **B. Different states of consciousness have different effective connectivity patterns.** The axes depict the two dimensions, wakefulness and awareness, that interact to produce conscious states. Being awake is characterised by high levels of both, and both are gradually reduced when one falls asleep or is sedated. REM sleep, where wakefulness is low, often features dreaming, in which awareness is high; in DOCs like the vegetative and minimally conscious states, wakefulness is high but awareness is low. The figure shows three illustrative TMS-induced activity patterns that occur in healthy brains: compared to wakefulness, effective connectivity is reduced during early (nREM) sleep, but not during REM sleep. **C. The perturbational complexity index (PCI).** A schematic illustration of TMS-evoked EEG activity from all electrodes for a conscious (top row) and vegetative (bottom row) brain. Left: Activity from all electrodes is shown on the same graph. Stimulation of a fully-conscious brain leads to more complex and prolonged activity. Middle: The activity of each electrode (vertical axis) over time is binarized, showing (in black) only times in which activity was greater than a pre-determined significance threshold. Right: The resulting pattern of activity is then algorithmically compressed, yielding larger amounts of information and higher values of PCI for the complex patterns typical of consciousness than for simple patterns typical of its absence.

**Figure 2. Bistable perception and brain regions involved in perceptual switches. A. The Necker cube.** This classic bistable stimulus can be perceived in two different ways: with the bottom-left side or the top-right side in front. Subjective experience switches between the two. **B. Schematic depiction of a structure-from-motion (SFM) sphere.** This ambiguous stimulus is composed of moving dots, which can be perceived as rotating in either of two directions, with periodic alternations in perceived direction. **C. Binocular rivalry.** When the top stimulus is viewed through red-green glasses, each eye sees a different image – one eye sees a face, the other a house. This is one way to present each eye with different images at the same location, creating binocular rivalry. The bottom row of images shows the resulting conscious perception. One image dominates awareness while the other is suppressed, and dominance switches periodically. If you don’t have red-green glasses, you can still experience rivalry by free-fusing images in the bottom row: cross your eyes until the face image overlaps with one of the house images next to it (the small central checkerboards and frames should help you maintain the overlap); if you have normal vision, perceptual dominance will alternate between the images. **D. frontal and parietal activation and stimulation sites in bistable perception studies.** Crosses show stimulation sites in TMS studies. Squares and circles show switch-related activity in bistable perception, under conditions in which participants reported perceptual switches (squares indicate locations based on coordinates from the original study; circles indicate approximate locations based on figures from a study that did not provide coordinates). Regions that are commonly referred to in the literature are circled.