

# Modes and Models in Disorders of Consciousness Science

M. BOLY<sup>1</sup>, A.K. SETH<sup>2</sup>

<sup>1</sup> Coma Science Group, Cyclotron Research Centre and Neurology Department, University of Liège, Belgium; <sup>2</sup> Sackler Centre for Consciousness, University of Sussex, Brighton, UK

## ABSTRACT

*The clinical assessment of non-communicative brain damaged patients is extremely difficult and there is a need for paraclinical diagnostic markers of the level of consciousness. In the last few years, progress within neuroimaging has led to a growing body of studies investigating vegetative state and minimally conscious state patients, which can be classified in two main approaches. Active neuroimaging paradigms search for a response to command without requiring a motor response. Passive neuroimaging paradigms investigate spontaneous brain activity and brain responses to external stimuli and aim at identifying neural correlates of consciousness. Other passive paradigms eschew neuroimaging in favour of behavioural markers which reliably distinguish conscious and unconscious conditions in healthy controls. In order to furnish accurate diagnostic criteria, a mechanistic explanation of how the brain gives rise to consciousness seems desirable. Mechanistic and theoretical approaches could also ultimately lead to a unification of passive and active paradigms in a coherent diagnostic approach. In this paper, we survey current passive and active paradigms available for diagnosis of residual consciousness in vegetative state and minimally conscious patients. We then review the current main theories of consciousness and see how they can apply in this context. Finally, we discuss some avenues for future research in this domain.*

## Key words

*Theories of consciousness • Vegetative state • Neuroimaging*

## Introduction

The clinical diagnosis of patients with disorders of consciousness is extremely difficult, leading to a high rate of misdiagnosis (Schnakers et al., 2009). There are several reasons for these difficulties. First, the results of patients' clinical assessment are highly dependent on the way clinicians search for behaviour at the bedside, and in particular on the details of the coma scale used. Among the many coma scales presently available, only a few explicitly incorporate diagnostic criteria differentiating vegetative state (VS) from minimally conscious state (MCS) in behavioural terms (Seel et al., 2010). Even if clinicians use appropriate scales, some behaviours, such

as blinking to visual threat (Vanhaudenhuyse et al., 2008) or grimacing to pain (Zeman, 1997), remain equivocal and are difficult to interpret as being either purely reflex, or as signs of residual conscious perception of external stimuli. The intrinsic problem of behavioural evaluation of the level of consciousness in VS and MCS is that the clinician is required to infer the presence or absence of consciousness based on behaviour, in the absence of a 'gold standard' verbal report (Boly et al., 2009a). Furthermore, behavioural responsiveness can be biased by a number of additional confounds including the quality of the patients' language comprehension, willingness to collaborate, availability of motor control, and so on (Boly et al., 2007).

Taken together, the above issues illustrate the challenges attending the behavioural diagnosis of patients with disorders of consciousness, and explain the increased interest of clinical community in the use of neuroimaging techniques to complement bedside behavioural diagnosis (Rafii and Brewer, 2010). In the present article, we will review neuroimaging studies that have been performed in patients in VS and MCS, focusing on the distinction between so-called ‘passive’ and ‘active’ paradigms. After briefly discussing non-neuroimaging passive approaches, we will discuss how mechanistic and theoretical perspectives on the fundamental problem of relating brain activity and consciousness could help unify passive and active approaches in a coherent diagnostic procedure. We will review the main theoretical frameworks currently available in this context, concluding with some potential future avenues of research.

### Modes: passive versus active neuroimaging paradigms

In the last few years, advances in neuroimaging techniques have led to the development of several novel paradigms aiming at assessing the potential for residual cognitive functions in severely brain damaged patients. Putting aside the variability in the neuroimaging techniques and the type of stimuli employed, these paradigms can be classified into two complementary categories. *Passive paradigms* are used to investigate the patients’ brain activity at rest or during administration of external stimuli, without requiring their collaboration. *Active paradigms* are used to obtain a response to command from the patients assessed by differential brain responses, by-passing motor output.

For good reasons, active paradigms have attracted considerable interest in the last few years. If they are properly designed and controlled, they can provide direct evidence for the presence of residual consciousness in individual brain damaged patients. An early example employed functional MRI (fMRI) assessment of the performance of mental imagery tasks, in order to obtain a response to command (Boly et al., 2007). In this arrangement, patients are instructed to repetitively alternate 30 seconds of motor (tennis playing) imagery or spatial naviga-

tion mental imagery with 30 seconds of rest. fMRI data analysis then aims at detecting task-specific motor or spatial navigation neural activation during the periods in which the patient was instructed to perform the task, as compared to periods of rest. This paradigm has proven able to identify the presence of awareness in some patients previously clinically diagnosed as being in a vegetative state (Owen et al., 2006). Furthermore, it allowed for the first time interactive real-time communication using fMRI brain signals in a single brain-damaged patient (Monti et al., 2010). Additional paradigms have been designed, using event-related potentials, which are more flexible and are potentially usable at the patients’ bedside (Schnakers et al., 2008; Bekinschtein et al., 2009).

Despite their definite clinical relevance and usefulness, active paradigms have a number of limitations. Recent cohort studies have shown that only a minority of patients are able to positively respond to this approach – of the order of 10% (Monti et al., 2010). In the 90% of patients that do not respond to command in a way detectable by neuroimaging methods, one cannot say anything definite about the presence or absence of residual cognition or consciousness (Boly et al., 2007). Negative findings have to be interpreted very cautiously for several reasons. Firstly, at a very general level, consciousness is not an all-or-none phenomenon but should rather be considered as a continuum (Majerus et al., 2005; Seth et al., 2008) finessing any diagnosis of residual consciousness. Secondly, some supra-tentorial brain lesions could impair the patients’ ability to perform a selected task, for example by causing apraxia or neglect (Boly et al., 2007). Brain injury could also lead to a certain amount of reorganization and plasticity, potentially resulting in the recruitment of other areas during the performance of a given cognitive task (Demertzi et al., 2010). These factors could lead to negative results in active neuroimaging paradigms, independently of the patient’s awareness or vigilance levels. In addition, maintaining mental imagery tasks for periods of 30 s is attention demanding, again emphasizing the value of a positive result. However, patients with severely altered level of vigilance may only be able to supply transient responses to the presentation of instructions. A third point is that a defining feature of voluntary actions is that one can choose whether or not to

execute them (Passingham, 1995). The success of any active paradigm is therefore dependent on the desire or willingness (if the patient is aware) of the patient to respond (Majerus 2005). Finally, although one might think that a positive behavioural response in an active paradigm should imply a positive neuroimaging response, this does not always seem to be the case. A recent study investigating MCS patients (Bardin et al., 2011) showed an absence of fMRI command following in patients who were able to show some behavioural responsiveness. Considering these limitations together with the observed frequency of false negatives reported across fMRI investigations of healthy volunteers (McGonigle et al., 2000), the conclusion is very clear: although positive results using active neuroimaging paradigms can provide direct evidence for residual consciousness, negative results can never exclude the possibility that the patient retains awareness of self or environment.

Passive paradigms have a longer history than their active counterparts in the neuroimaging assessment of patients with disorders of consciousness. They allow a global assessment of the patients' residual brain function, without requiring the patients' explicit collaboration. For this reason, they can provide information on brain activity and the potential for consciousness in each and every patient studied. They also allow, at least in principle, that the neural mechanisms underlying consciousness *per se* may not be identical to those mechanisms supporting explicit report, verbal or otherwise (Block, 2007, Hulme 2009, Lamme, 2010).

Passive neuroimaging paradigms, as used in the current literature, encompass a large variety of approaches which aim at identifying differences in brain function between MCS and VS. Early resting state PET studies identified a global dysfunction of a bilateral fronto-parietal thalamo-cortical network in patients in a VS as compared to controls (Laureys et al., 1999), later found to be relatively more preserved in patients in MCS (Laureys et al., 2004). PET studies investigating responses to external stimuli found a pattern of localized response and functional disconnection in patients in a VS as compared to controls (Laureys et al., 2000; 2002), while MCS patients showed a near-to-normal response (Boly et al., 2004; 2008a). Resting state fMRI studies identified impaired thalamo-cortical con-

nectivity in the default network (Boly et al., 2009b; Vanhaudenhuyse et al., 2010) and other networks (Zhou et al., 2010) in coma and VS as compared to controls. As compared to VS patients, MCS patients showed stronger precuneus involvement in the default network (Vanhaudenhuyse et al., 2010). Collectively, these findings show preserved brain function in MCS as compared to VS, in line with recent reports of better preserved cerebral structural integrity in this patient population (Fernandez-Espejo et al., 2010; 2011).

Currently, however, passive paradigms do not allow a differentiation of VS from MCS at the individual level (though see Fernandez-Espejo et al., 2011). Most studies comparing MCS to VS patients have been performed at the group level, and did not systematically investigate the inter-subject variability in the brain activity criteria observed. Furthermore, fMRI studies have shown that it is possible to elicit activation of associative cortices in response to external stimulation in individual VS patients (Di et al., 2007; Rodriguez Moreno et al., 2010), complicating the interpretation of passively acquired fMRI data. Ultimately, as discussed below, further research on neural correlates of consciousness (NCCs) is needed to bring passive neuroimaging paradigms closer to diagnostic value.

An alternative passive approach is to use behavioural or autonomic markers which reliably differentiate conscious and unconscious conditions in healthy controls. For example, several studies indicate that consciousness is needed in order to learn 'trace conditioning' contingencies, in which there is delay separating the end of the CS (conditioned stimulus) from the start of the US (unconditioned stimulus), but is not needed to learn 'delay conditioning' contingencies (in which the CS and US overlap) (Clark and Squire, 1998; Manns et al., 2000). Bekinschtein and colleagues exploited this observation to examine whether trace conditioning could be learned under anaesthesia and by patients in VS and MCS (Bekinschtein et al., 2009a). Using anticipatory electromyographical responses as an indicator of learning during eyeblink conditioning, they found that the degree of learning was a good indicator of recovery in patients; however, the method failed to provide a clean separation of conscious and (anaesthetically) unconscious control subjects; some conscious controls failed to learn and one unconscious con-

trol showed marginal learning while under general anaesthesia. In another example, Scott et al. (Scott et al., 2010) showed, using healthy controls, that a simple paradigm involving exposure to predictable *versus* unpredictable sound sequences clearly differentiated conscious from unconscious conditions (they used inattention as a proxy for unconsciousness), assessed via skin conductance responses. Their procedure, termed the ‘learned aversive contingency’ (LAC) procedure, is currently being tested in patients. An advantage to these approaches is that they can be assessed by very simple methods (e.g., galvanic skin response) in contrast to requiring an MRI scanner. These paradigms have however to be better validated at the individual level before one can assess of their clinical utility.

What validates the use of a passive paradigm? One answer is that a passive paradigm is valid if, in healthy controls, it unambiguously differentiates conscious from unconscious conditions (Bekinschtein et al., 2009a; Scott et al., 2010). However, while this is a useful practical heuristic it overlooks the obvious fact that brain damaged patients have abnormal brains, so that what may be a clear differentiation in healthy controls may not be so clear in patients. Ultimately, in the absence of a consensus regarding the neural mechanisms underlying consciousness, passive paradigms cannot lead to a direct diagnosis of the presence or absence of cognition in individual patients. Research on passive paradigms and NCCs thus go hand-in-hand (Seth et al., 2008; Tononi and Koch, 2008), advancing both the global study of brain function in disorders of consciousness, and the investigation of the links between brain activity and consciousness in the normal human brain. Such an approach facilitates a transition from ‘exploratory’ to ‘explanatory’ NCCs; where the latter refers to brain processes that do not merely *correlate with* but actually *account for* fundamental phenomenological properties of conscious experience (Tononi and Koch, 2008; Seth, 2009). This mechanistic understanding would not only allow more confidence in the diagnosis, but also the development of new therapeutic strategies in individual brain damaged patients.

To advance mechanistic understanding, a combination of theory and experiment is required. Theoretical approaches allow taking some distance from the data and building a conceptual description

of how phenomenal experience could be generated as a consequence of different patterns of neural activity. With regard to experiment, restricting investigation to only one mode of altered consciousness can induce bias in the search for truly explanatory NCCs (Boly et al., 2009a) because patterns of brain activity can differ substantially among modes of unconsciousness (Brown et al., 2010). For example, global brain metabolism and functional brain connectivity seem to be markedly decreased in conditions such as coma, anaesthesia, or sleep, while they are paradoxically increased in some other conditions such as generalized tonico-clonic seizures and some forms of temporal lobe epilepsy (Arthuis et al., 2009). Similarly, EEG high frequency activity (e.g. in the beta and gamma bands) decreases during non rapid eye movement (REM) sleep (Nishida et al., 2005; Corsi-Cabrera et al., 2006) but can increase during anesthesia-induced loss of consciousness (LOC) (Murphy et al., 2011). Theoretical approaches to explanatory NCCs should therefore encompass a broad view of the different states of altered states of consciousness (Boly et al., 2008b), in order to potentially identify common underlying mechanisms. Combined with further experiments comparing brain function in conscious and unconscious states, the specification of explanatory NCCs could motivate a broader use of passive paradigms as diagnostic tools in patients with disorders of consciousness (Table I).

In the next section, we briefly review some conceptual issues in the identification of explanatory NCCs in the human brain, in the context of several prominent current theories of consciousness. We conclude by describing further research avenues potentially relevant to the study of brain function in brain damaged patients, and on correlates of consciousness in a larger, theoretically informed context.

## Models: Explanatory NCCs and theories of consciousness

Two general approaches to the study of NCCs can be distinguished: (i) those that contrast consciously perceived stimuli against those perceived only subliminally, usually in healthy volunteers (i.e., conscious *content*), and (ii) those that contrast different global states of consciousness (i.e., conscious *level*;



Table I. - Active versus passive paradigms: pros and cons.	
Active paradigms	Passive paradigms
If positive responses, directly usable information on the conscious state of the patient	Not directly usable for single-subject diagnosis at present.
If negative response, no conclusion possible	Requires parallel research on NCC and theories of consciousness to support interpretation.
Requires patient's collaboration	Does not require patient's collaboration
Negative response in ~90% patients	Information on the patient's brain function obtained in every case

examples include normal waking consciousness, sleep, anaesthesia, VS, MCS, and so on) (Seth et al., 2008; Hohwy, 2009). Example approaches to studying content NCCs include masking, binocular rivalry, continuous flash suppression paradigms, or studies of motion-induced blindness, change blindness or inattention blindness (for reviews see (Seth et al., 2008; Tononi and Koch, 2008)). Studies using these paradigms are almost always modality-specific, meaning they focus on auditory or visual modalities, or other specific aspects of cognition such as emotion, memory or language embedded in a specific sensory modality. Such studies have delivered many important insights into content NCCs, for example identifying general mechanisms such as the involvement of fronto-parietal cortices (Dehaene et al., 2006; Pollen, 2011), or the importance of long-latency ERP components (Del Cul, et al., 2007) and recurrent feedback processes (Lamme, 2006; 2010). An important issue raised by these studies, discussed further below, is whether these NCCs have to do with consciousness *per se*, or to do with the reportability of conscious experiences.

Disorders such as VS and MCS seem to involve not so much disturbances of specific content NCCs, but rather a more pervasive and non-modality-specific impairment of consciousness. While zero conscious level does imply a complete absence of conscious contents, inferences from conscious contents to conscious level are harder to make. For example, it seems possible to be highly conscious of some stimuli while being completely unconscious of others (Overgaard and Overgaard, 2010). In addition, there is abundant and accumulating evidence that many processes tied to specific conscious contents, or even dimensions of conscious contents (e.g., experiences of agency, volition, or subjective reality) can be dissociated from consciousness *per se*

(Tononi and Koch, 2008; Boly, et al., 2009a). These lines of evidence are summarized in Table II. Given these dissociations, it makes sense to search for indicators of conscious level that pertain to *all* conscious scenes, as far as possible independent from their specific contents. In other words, one could see a normal level of consciousness as a state making all sorts of conscious contents potentially available (repertoire), while not all cognitive functions or perception modalities are present to consciousness at a given time. Experimentally, this more global approach has been employed when investigating brain function in altered consciousness states such as coma, VS, MCS, epilepsy, sleep or anaesthesia. In these experiments, researchers investigate the difference between the presence and absence of generic ability to have conscious experience. This approach can seem at a first glance slightly simplistic, however comparing different altered states of consciousness and identifying common mechanisms renders it complex and challenging. Though content-based and level-based NCC investigations are complementary and certainly both useful, they do ask radically different questions. Generic theories of consciousness should however in principle address both of these topics in conjunction.

Block (2009) provides a useful differentiation of current theoretical frameworks of consciousness into three categories: *biological theories*, *higher-order thought theories*, and *global workspace/information integration theories*. (We note that the latter two categories are also 'biological' inasmuch as the corresponding theories do recognize that consciousness is instantiated by brains, however their key concepts are functional.) Biological theories postulate that consciousness is some sort of biological state of the brain. Quite what biological state or states count as conscious is of course not yet understood.

Table II. - Level *versus* content: current evidence for the fact that consciousness *per se* can be dissociated from content-specific sensory or cognitive processes in the human brain. Most references cite studies showing the continued presence of consciousness in the absence of a specific process; those marked with a \* show that a particular process which is usually conscious can also occur in the absence of consciousness.

Processes that can be dissociated from consciousness	Practical examples	References
Sensori-motor interaction with environment	Dream consciousness Locked-in syndrome	(Nir and Tononi, 2010) (Laureys et al., 2005)
Visual perception	Cortical blindness	(Miller 1982)
Tactile sense - Sense of body	Loss of proprioception, deafferentation	(Gallagher and Cole, 1995; Blanke et al., 2002)
Language	Aphasic patients - Wada test	(Trener and Loring, 1995)
Sense of space	Balint syndrome, neglect	(Phan et al., 2000; Vallar, 2007)
First person perspective from within the body	Out of body experiences	(Blanke et al., 2002; Lenggenhager et al., 2007)
Episodic memory	Globally Amnesic patients	(Bartsch and Deuschl, 2010)
Declarative memory*	Subliminal declarative memory processes	(Reder et al., 2009; Henke 2010)
Working memory*	Implicit working memory processes	(Hassin et al., 2009)
Sense of self – introspection or reflection	Loss of self awareness while watching a deeply absorbing movie	(Goldberg et al., 2006)
Sense of subjective reality of world and/or of self	Depersonalization and derealization; Cotard's delusion	(Phillips et al., 2001; Pearn and Gardner-Thorpe, 2002; Sierra, 2009)
Attention (1)	Consciousness without attention – perception of peripheral visual field	(Koch and Tsuchiya, 2007; Van Boxtel et al., 2010)
Attention (2)*	Attention without consciousness – possibility of cerebral treatment of attended subliminal stimuli	(Koch and Tsuchiya, 2007; Van Boxtel et al., 2010)
Volition (1)	Consciousness without volition: akinetic mutism	(Marin and Wilkosz, 2005)
Volition (2)*	Intention without consciousness	(Libet, Gleason et al., 1983; Soon et al., 2008)

Candidates include proposals linking consciousness to gamma band activity (Llinas et al., 1998), recurrent processing (Lamme, 2006) or long-distance connectivity or synchrony (Melloni et al., 2007). A distinguishing feature of biological theories is that they allow for conscious contents to exist without these contents necessarily being reportable; or, to put it more strongly, subjects could have phenomenally conscious states that the subject does not know about or have 'access' to (Block, 2007; Hulme et al., 2009; Lamme, 2010). In Block's terminology, biological theories allow a distinction between *phenomenal consciousness* and *access consciousness*. This distinction has important implications for diagnosis of residual consciousness. Non-reportable conscious contents (i.e. contents which are phenomenally conscious but not access conscious) are by definition beyond the reach of active paradigms, which require

subjects to make (non-behavioural) reports based on access. [An interesting middle ground may involve dissociations between conscious contents and conscious selfhood. For example, people with pain asymbolia (Aydede, 2005). A critical and quasi-historical essay on theories of pain. Pain: New essays on its nature and the methodology of its study. M. Aydede. Cambridge, MA, MIT press.) claim to have pain experiences but do not mind having these experiences, sometimes describing the pain as painful for someone else. Possibly such dissociated conscious contents could still be reportable]. Only passive paradigms could in principle detect such contents. However, as mentioned above, in order to do so one first needs to understand what biological states or processes count as conscious.

According to higher-order thought (HOT) theories, a mental state is conscious when a person is actu-

ally aware, or disposed towards being aware, of being in that state; either by directly perceiving it, or by thinking about it (Rosenthal, 2000; 2005). For example, on HOT theories the conscious experience of red consists in a first-order neural representation of red (probably in the visual system) accompanied by a perception or thought (possibly originating in frontal cortices), directed at that representation, with the content that the subject is having the experience of red. HOT theories offer an attractive epistemological simplicity, since being conscious of X is now equivalent to having a HOT potentially allowing communication/report of that fact; in other words, consciousness implies *access*. However, HOT theories arguably underestimate the richness of primary, phenomenal consciousness by asserting that components of conscious scenes are necessarily accompanied by HOTs. With respect to diagnosing residual consciousness, HOT theories suffer the opposite problem to biological theories. Whereas biological theories allow for non-reportable conscious contents but (as yet) provide insufficient criteria for their identification, HOT theories rule out such experiences by theoretical fiat. Consider the important example of pain: even if patients are not able to report it, the existence of a conscious pain sensation is a, if not *the*, crucial matter from the perspective of proper diagnosis and clinical management.

The final frameworks, global workspace and information integration theories, are in our view more relevant in the context of patients with disorders of consciousness though they are not without their own problems. Global workspace theory (GWT) was introduced by Baars (Baars, 1988) and can be viewed as a ‘theatre’ metaphor of mental functioning (Baars and Laureys, 2005). According to Baars et al. (2003), once conscious sensory content is established, it is distributed widely to a decentralized “audience” of expert networks – executive interpreters, involving parietal and prefrontal cortices. GWT has been developed in a more neural direction by Dehaene, Naccache, and Changeux (Dehaene and Naccache, 2001; Dehaene and Changeux, 2004; 2005; Dehaene et al., 2006) and by Shanahan (Shanahan, 2008). In terms of empirical data relevant to conscious content, GWT predicts that conscious perception should involve widespread brain sources, and unconscious sensory processing should be much more limited (Baars, 2005). Regarding overall conscious

level, loss of consciousness in states like coma, VS, sleep, and anaesthesia should be accompanied by decreased activity in “observing self” fronto-parietal regions (Baars et al., 2003). In addition to fronto-parietal activity, these authors suggested that the level of consciousness should be determined by the amount of spontaneous fast frequency oscillatory activity in the thalamo-cortical system (Dehaene and Changeux, 2005). A disadvantage of GWT is that, like HOT theories, it is primarily a theory of conscious access rather than of phenomenal consciousness. On GWT, phenomenal consciousness consists in the process of global broadcast, which entails access. Thus, like HOT theories, GWT may be insensitive to non-reportable conscious experiences, should they exist. GWT does however have the advantage of attempting to make empirical predictions, providing a potentially useful foothold for clinical application. Future developments of GWT in this regard could focus on: i) unpacking, functionally and mechanistically, the crucial concept of ‘broadcast’ in the genesis of conscious experience (i.e., what exactly is broadcast, and how does it happen?); and ii) leveraging the useful discussions of the role of *context* in shaping workspace contents, a rich but usually overlooked aspect of Baars’ original theory (Baars, 1988).

Information integration theories, such as Tononi’s ‘information integration theory of consciousness’ (IITC, (Tononi, 2008)) and Tononi and Edelman’s previous ‘dynamic core hypothesis’ (Tononi and Edelman, 1998) start from a different intuition. In contrast to GWT, these theories do not focus on access consciousness; they have the more globalist aim of explaining links between brain activity and the emergence of any kind of conscious contents. Information integration theories emphasize the *dynamical complexity* of conscious scenes, namely that conscious scenes are simultaneously integrated (i.e., they are experienced ‘all of a piece’) and differentiated (i.e., each conscious scene is one among a vast repertoire of possible conscious scenes). Therefore, the occurrence of any particular conscious experience generates an enormous quantity of information by ruling out a vast repertoire of alternative possibilities (differentiation) and because each scene is integrated, this is information *for* the system (i.e., not for an observer of the system, as might be the case, for example, for a digital camera).

Theories of this kind predict that conscious level will be associated with a measure of ‘dynamical complexity’ tracking coexisting differentiation and integration, applied to the relevant neural variables. If true, then such measures could be useful when applied to neural data recorded from patients in the context of passive paradigms.

A first way to check these predictions empirically was provided by the combination of transcranial magnetic stimulation (TMS) with recordings using high-density EEG. An early study (Massimini et al., 2005) showed that TMS pulses administered during slow wave sleep evoked activity only local to the TMS focus; in contrast, TMS applied during conscious wakefulness led to a spatiotemporally complex pattern of activity with several distinct components over time, consistent with a rich underlying effective and functional connectivity. This paradigm has recently been validated in other states of altered consciousness, conditions including anaesthesia (Ferrarelli et al., 2010) and REM sleep (Massimini et al., 2010). In each case, putatively unconscious conditions led to evoked responses similar to those observed during slow wave sleep; by contrast, the REM sleep condition was very similar to normal conscious wakefulness. Perhaps the most promising aspect of this approach is that it leverages theoretical work linking consciousness with high levels of dynamical complexity in the underlying neuronal networks. This provides the hope to better differentiate conscious from unconscious patients at the individual level (Massimini et al., 2009). TMS-EEG studies in VS and MCS patients are currently ongoing.

The use of neuroimaging methods as a diagnostic tools would benefit from the ability to summarize results in a simple score (number), as a very generic measure. In the context of TMS-EEG, the use of the response entropy or algorithmic complexity has been proposed (Massimini et al., 2009) and is currently being evaluated. Several other theoretically-grounded candidate measures also exist for characterising dynamical complexity in neural data. The earliest was ‘neural complexity’ (Tononi and Edelman, 1998) which is determined by the average mutual information among subsets of all possible sizes for all bipartitions of a system. This measure has been criticized however for not reflecting causal interactions (mutual information is a symmetric measure of general statistical dependence). Addressing this,

Tononi introduced the measure ‘integrated information’ ( $\Phi$ , Tononi, 2004)) which measures the information generated when a system transitions from one state to the next (differentiation), to the extent that this information is generated by the whole system and not by the parts considered independently (integration). This measure forms the cornerstone of Tononi’s IITC according to which consciousness is considered as a ‘capacity’ or ‘potential’ (Tononi, 2005).  $\Phi$  however is to date difficult or infeasible to calculate for nontrivial systems [though see a recent adaptation (Barrett and Seth, 2011) which is computable from general time series data and conceives of integrated information as a ‘process’ rather than as a ‘capacity/potential’]. A third measure, ‘causal density’, reflects the overall density of causal interactions among elements (or bipartitions) of system (Seth et al., 2006; 2011) using a statistical measure of causality based on precedence and relative predictability. Causal density entails fewer assumptions on the data and, in contrast to Tononi’s  $\Phi$ , reflects a process rather than a capacity. It bears emphasizing that these measures, while all operationalizing dynamical complexity, do so in different ways which have substantial implications for how consciousness is conceptualized (e.g., as a capacity, or as a process) (Seth et al., 2011).

In contrast to (Block’s) biological theories, GWT and information integration theories are essentially functionalist; it is conceivable that machines or other systems could incorporate global workspace architectures and/or generate high levels of dynamic complexity. Unlike GWT, information integration theories are theories of phenomenal as well as access consciousness; content will be conscious if it is subtended by neural dynamics with sufficiently high dynamical complexity; access will ensue if these processes incorporate neural mechanisms associated with broadcast and reportability. Also, information integration theories and GWT make different predictions regarding empirical data. For instance, on information integration theories widespread activity has to be complex in order to furnish both integration and differentiation, a requirement that is not explicit in GWT (Balduzzi and Tononi, 2008). Evidence on this point is clear: under some circumstances widespread activity can actually signify absence of consciousness, for example during hypersynchronous states associated with epileptic absence seizures



Table III. - Predictions of each theory and remaining issues. HOT: Higher Order Thought; LOC: loss of consciousness.		
Theories	Predictions	Remaining issues
Biological	Depending on the theory: 1) Fast gamma activity; 2) Widespread activation/connectivity; 3) Recurrent processing	Verify that the biological mechanisms hold in each and every conscious state, as compared to unconscious states.
HOT	Higher order activity should be involved in each and every conscious perception	Accounts for access consciousness but no account for phenomenal consciousness. More precise predictions to be generated - not biologically rooted yet. What is the neural correlate of a HOT?
Global workspace	Generally widespread activity and fast background activity (Baars); activation of higher order fronto-parietal cortices (Dehaene and colleagues)	More precise predictions to be generated; theory needs to be refined to account for widespread activity during LOC. Also, emphasis is on access rather than phenomenal consciousness.
Information integration theories	Consciousness should be absent if: bistable dynamics; hypo/hyper activity and/or connectivity	More precise predictions to be generated - not sufficiently biologically rooted and proposed measures are hard to calculate in practice.

(Arthuis et al., 2009). While the loss of consciousness in this condition has a natural interpretation in terms of information integration theories (loss of differentiation), GWT theories would need to postulate additional mechanisms (e.g., that hypersynchronous activity may somehow ‘block access’ to the global workspace) to potentially account for loss of consciousness (Boly et al., 2009a). In general, information integration theories successfully retrodict that patterns of brain activity such as very low neuronal firing/connectivity, hyperactivity/hypersynchrony, and bistable dynamics are especially unfavourable for the brain to generate meaningful levels of consciousness (Tononi, 2008; Barrett and Seth, 2011). These theories are also highly consistent with the TMS-EEG evidence discussed earlier (see above) inasmuch as spatiotemporally complex responses to TMS pulses suggest an underlying functional connectivity with high, or potentially high, integrated information or causal density. An explicit quantitative connection between TMS-EEG data and these quantities is still however lacking, though research on this topic is presently ongoing (Table III).

## Conclusion and Perspectives

While active paradigms have considerable clinical and ethical relevance, work on these paradigms has to be paired with further research on passive paradigms and neural correlates of consciousness. Combining

the results of different neuroimaging modalities, as well as the use of computational modelling, could also furnish new insights into how biological mechanisms are translated into neural activity patterns potentially underlying conscious contents. Although model-based and relying on a number of underlying assumptions, computational modelling is currently very well placed to take advantage of advances in neuroscience and neuroanatomy, as well as of new information technology developments such as the adaptation of graphics processing units (GPUs) to implement desktop parallel computing capable of simulating large scale realistic brain models (Izhikevich and Edelman, 2008; Nageswaran et al., 2009; Ching et al., 2010). Data-informed approaches such as Dynamic Causal Modelling (Friston et al., 2003; David et al., 2006) and Granger causality (Granger, 1969; Seth, 2010), which attempt to infer the underlying neural mechanisms generating brain signals in a given data set, could provide a bridge between modelling and experimental evidence in order to gain further insights into the links between brain signals, neuronal activity and consciousness (Boly et al., 2011). Multi-centric studies on larger groups of patients (Monti et al., 2010), as well as their testing in other conditions such as anaesthesia and sleep (Boly et al., 2008b), are also warranted in order to better evaluate the diagnostic and prognostic value of the different paradigms in use. Finally, a better understanding of lesion patterns and associated clinical and prognostic states could improve mechanistic

understanding of the link between consciousness and the brain (Fernandez-Espejo et al., 2010). Ultimately, however, we stand in need of progress on the different theories of consciousness in order to render them more precise, such that they furnish predictions that can be tested on actual brain data in a way that allows their refinement, confirmation, or rejection. Ideally, theoretical advances will furnish explanatory correlates of consciousness that have sufficient generality and explanatory power that they can support the confident use of passive paradigms as diagnostic tools in patients with disorders of consciousness. Such theoretically grounded and experimentally validated passive paradigms are greatly to be desired since they circumvent the limitations of active paradigms in relying on patient comprehension and cooperation.

## References

- Arthuis M., Valton L., et al. Impaired consciousness during temporal lobe seizures is related to increased long-distance cortical-subcortical synchronization. *Brain*, **132** (Pt 8): 2091-2101, 2009.
- Aydede M. A critical and quasi-historical essay on theories of pain. In: *Pain: New essays on its nature and the methodology of its study*. Cambridge, MA, MIT Press, 2005.
- Baars B.J. *A Cognitive Theory of Consciousness*. Cambridge, Cambridge University Press. 1988.
- Baars B.J., Ramsoy T.Z., et al. Brain, conscious experience and the observing self. *Trends Neurosci.*, **26**: 671-675, 2003.
- Baars B.J. Global workspace theory of consciousness: toward a cognitive neuroscience of human experience. *Prog. Brain. Res.*, **150**: 45-53, 2005.
- Baars B.J. and Laureys S. One, not two, neural correlates of consciousness. *Trends Cogn. Sci.*, **9**: 269 (Author reply: 270), 2005.
- Balduzzi D. and Tononi G. Integrated information in discrete dynamical systems: motivation and theoretical framework. *PLoS Comput. Biol.*, **4**: e1000091, 2008.
- Bardin J., Fins J., et al. Dissociations between behavioral and fMRI-based evaluations of cognitive function after brain injury. *Brain*, **134** (Pt 3): 769-782, 2011.
- Barrett A.B. and Seth A.K. Practical measures of integrated information for time-series data. *PLoS Comput. Biol.*, **7**: e1001052, 2011.
- Bartsch T. and Deuschl G. Transient global amnesia: functional anatomy and clinical implications. *Lancet Neurol.*, **9**: 205-214, 2010.
- Bekinschtein T.A., Shalom D.E., et al. Classical conditioning in the vegetative and minimally conscious state. *Nat. Neurosci.*, **12**: 1343-1349, 2009a.
- Bekinschtein T.A., Dehaene S., et al. Neural signature of the conscious processing of auditory regularities. *Proc. Natl. Acad. Sci. U S A*, **106**: 1672-1677, 2009b.
- Blanke O., Ortigue S., et al. Stimulating illusory own-body perceptions. *Nature*, **419**: 269-270, 2002.
- Block N. Consciousness, accessibility, and the mesh between psychology and neuroscience. *Behav. Brain. Sci.*, **30**: 481-499; discussion 499-548, 2007.
- Block N. Comparing the Major Theories of Consciousness. In: Gazzaniga M.S. (Ed.). *The Cognitive Neuroscience*. Cambridge MA, MIT Press. IV: 11, 2009.
- Boly M., Faymonville M.E., et al. Auditory processing in severely brain injured patients: differences between the minimally conscious state and the persistent vegetative state. *Arch. Neurol.*, **61**: 233-238, 2004.
- Boly M., Coleman M.R., et al. When thoughts become action: an fMRI paradigm to study volitional brain activity in non-communicative brain injured patients. *Neuroimage*, **36**: 979-992, 2007.
- Boly M., Faymonville M.E., et al. Perception of pain in the minimally conscious state with PET activation: an observational study. *Lancet Neurol.*, **7**: 1013-1020, 2008a.
- Boly M., Phillips C., et al. Intrinsic brain activity in altered states of consciousness: how conscious is the default mode of brain function? *Ann. N Y Acad. Sci.*, **1129**: 119-129, 2008b.
- Boly M., Massimini M., et al. Theoretical approaches to the diagnosis of altered states of consciousness. *Prog. Brain Res.*, **177**: 383-398, 2009a.
- Boly M., Tshibanda L., et al. Functional connectivity in the default network during resting state is preserved in a vegetative but not in a brain dead patient. *Hum. Brain. Mapp.*, **30**: 2393-2400, 2009b.
- Boly M., Garrido M.I., et al. Preserved feedforward but impaired top-down processes in the vegetative state. *Science*, **332**: 858-862, 2011.
- Brown E.N., Lydic R., et al. General anesthesia, sleep, and coma. *N. Engl. J. Med.*, **363**: 2638-2650, 2010.
- Ching S., Cimenser A., et al. Thalamocortical model for a propofol-induced (alpha)-rhythm associated

- with loss of consciousness. *Proc. Natl. Acad. Sci. U S A*, **107**: 22665-22670, 2010.
- Clark R.E. and Squire L.R. Classical conditioning and brain systems: the role of awareness. *Science*, **280**: 77-81, 1998.
- Corsi-Cabrera M., Munoz-Torres Z., et al. Power and coherent oscillations distinguish REM sleep, stage 1 and wakefulness. *Int. J. Psychophysiol.*, **60**: 59-66, 2006.
- David O., Kiebel S.J., et al. Dynamic causal modeling of evoked responses in EEG and MEG. *Neuroimage*, **30**: 1255-1272, 2006.
- Dehaene S. and Naccache L. Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition*, **79**: 1-37, 2001.
- Dehaene S. and Changeux J.P. Neural Mechanisms for Access to Consciousness. In: Gazzaniga M.S. (Ed.). *The cognitive neurosciences*, 3rd edition. New York, Norton: 1145-1157, 2004.
- Dehaene S. and Changeux J.P. Ongoing spontaneous activity controls access to consciousness: a neuronal model for inattention blindness. *PLoS Biol*, **3**: e141, 2005.
- Dehaene S., Changeux J.P., et al. Conscious, pre-conscious, and subliminal processing: a testable taxonomy. *Trends Cogn. Sci.*, **10**: 204-211, 2006.
- Del Cul A., Baillet S., et al. Brain dynamics underlying the nonlinear threshold for access to consciousness. *PLoS Biol.*, **5**: e260, 2007.
- Demertzi A., Schnakers C., et al. Neural plasticity lessons from disorders of consciousness. *Front. Psychol.*, **1**: 245, 2010.
- Di H.B., Yu S.M., et al. Cerebral response to patient's own name in the vegetative and minimally conscious states. *Neurology*, **68**: 895-899, 2007.
- Fernandez-Espejo D., Junque C., et al. Reductions of thalamic volume and regional shape changes in the vegetative and the minimally conscious States. *J. Neurotrauma*, **27**: 1187-1193, 2010.
- Fernandez-Espejo D., Bekinschtein T., et al. Diffusion weighted imaging distinguishes the vegetative state from the minimally conscious state. *Neuroimage*, **54**: 103-112, 2011.
- Ferrarelli F., Massimini M., et al. Breakdown in cortical effective connectivity during midazolam-induced loss of consciousness. *Proc. Natl. Acad. Sci. U S A*, **107**: 2681-2686, 2010.
- Friston K.J., Harrison L., et al. Dynamic causal modelling. *Neuroimage*, **19**: 1273-1302, 2003.
- Gallagher S. and Cole J. Body schema and body image in a deafferented subject. *J. Mind Behav.*, **16**: 369-390, 1995.
- Goldberg II, Harel M., et al. When the brain loses its self: prefrontal inactivation during sensorimotor processing. *Neuron.*, **50**: 329-339, 2006.
- Granger C.W. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, **37**: 424-438, 1969.
- Hassin R.R., Bargh J.A., et al. Implicit working memory. *Conscious. Cogn.*, **18**: 665-678, 2009.
- Henke K. A model for memory systems based on processing modes rather than consciousness. *Nat. Rev. Neurosci.*, **11**: 523-532, 2010.
- Hohwy J. The neural correlates of consciousness: new experimental approaches needed? *Conscious. Cogn.*, **18**: 428-438, 2009.
- Hulme O.J., Friston K.F., et al. Neural correlates of stimulus reportability. *J. Cogn. Neurosci.*, **21**: 1602-1610, 2009.
- Izhikevich E.M. and Edelman G.M. Large-scale model of mammalian thalamocortical systems. *Proc. Natl. Acad. Sci. U S A*, **105**: 3593-3598, 2008.
- Koch C. and Tsuchiya N. Attention and consciousness: two distinct brain processes. *Trends Cogn. Sci.*, **11**: 16-22, 2007.
- Lamme V.A. Towards a true neural stance on consciousness. *Trends Cogn. Sci.*, **10**: 494-501, 2006.
- Lamme V.A. How neuroscience will change our views on consciousness. *Cogn. Neurosci.*, **1**: 204-220, 2010.
- Laureys S., Lemaire C., et al. Cerebral metabolism during vegetative state and after recovery to consciousness. *J. Neurol. Neurosurg. Psychiatry*, **67**: 121, 1999.
- Laureys S., Faymonville M.E. et al. Auditory processing in the vegetative state. *Brain*, **123** (Pt 8): 1589-1601, 2000.
- Laureys S., Faymonville M.E., et al. Cortical processing of noxious somatosensory stimuli in the persistent vegetative state. *Neuroimage*, **17**: 732-741, 2002.
- Laureys S., Owen A.M., et al. Brain function in coma, vegetative state, and related disorders. *Lancet Neurol.*, **3**: 537-546, 2004.
- Laureys S., Pellas F., et al. The locked-in syndrome: what is it like to be conscious but paralyzed and voiceless? *Prog. Brain Res.*, **150**: 495-511, 2005.
- Lenggenhager B., Tadi T., et al. Video ergo sum: manipulating bodily self-consciousness. *Science*, **317**: 1096-1099, 2007.
- Libet B., Gleason C.A., et al. Time of conscious intention to act in relation to onset of cerebral

- activity (readiness-potential). The unconscious initiation of a freely voluntary act. *Brain*, **106** (Pt 3): 623-642, 1983.
- Llinas R., Ribary U., et al. The neuronal basis for consciousness. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, **353**: 1841-1849, 1998.
- Majerus S., Gill-Thwaites H., et al. Behavioral evaluation of consciousness in severe brain damage. *Prog. Brain Res.*, **150**: 397-413, 2005.
- Manns J.R., Clark R.E., et al. Parallel acquisition of awareness and trace eyeblink classical conditioning. *Learn Mem.*, **7**: 267-272, 2000.
- Marin R.S. and Wilkosz P.A. Disorders of diminished motivation. *J. Head Trauma. Rehabil.* **20**: 377-388, 2005.
- Massimini M., Ferrarelli F., et al. Breakdown of cortical effective connectivity during sleep. *Science*, **309**: 2228-2232, 2005.
- Massimini M., Boly M., et al. A perturbational approach for evaluating the brain's capacity for consciousness. *Prog. Brain Res.*, **177**: 201-214, 2009.
- Massimini M., Ferrarelli F., et al. Cortical reactivity and effective connectivity during REM sleep in humans. *Cogn. Neurosci.*, **1**: 176-183, 2010.
- McGonigle D.J., Howseman A.M., et al. Variability in fMRI: an examination of intersession differences. *Neuroimage*, **11** (6 Pt 1): 708-734, 2000.
- Melloni, L., Molina C., et al. Synchronization of neural activity across cortical areas correlates with conscious perception. *J Neurosci* **27**(11): 2858-2865, 2007.
- Miller N.R. *Walsh and Hoyt's Clinical Neuro-Ophthalmology*. Baltimore, 1982.
- Monti M.M., Vanhaudenhuyse A., et al. Willful modulation of brain activity in disorders of consciousness. *N. Engl. J. Med.*, **362**: 579-589, 2010.
- Murphy M., Bruno M.A., et al. Propofol anesthesia and sleep: a high-density EEG study. *Sleep.*, **34**: 283-291, 2011.
- Nageswaran J.M., Dutt N., et al. A configurable simulation environment for the efficient simulation of large-scale spiking neural networks on graphics processors. *Neural Netw.*, **22**: 791-800, 2009.
- Nir Y. and Tononi G. Dreaming and the brain: from phenomenology to neurophysiology. *Trends Cogn. Sci.*, **14**: 88-100, 2010.
- Nishida M., Uchida S., et al. High frequency activities in the human orbitofrontal cortex in sleep-wake cycle. *Neurosci. Lett.*, **379**: 110-115, 2005.
- Overgaard M. and Overgaard R. Neural correlates of contents and levels of consciousness. *Front. Psychol.*, **1**: 164, 2010.
- Owen A.M., Coleman M.R., et al. Detecting awareness in the vegetative state. *Science*, **313**: 1402, 2006.
- Passingham R.E. *The Frontal Lobes and Voluntary Action*. Oxford, Oxford University Press, 1995.
- Pearn J. and Gardner-Thorpe C. Jules Cotard (1840-1889): his life and the unique syndrome which bears his name. *Neurology*, **58**: 1400-1403, 2002.
- Phan M.L., Schendel K.L., et al. Auditory and visual spatial localization deficits following bilateral parietal lobe lesions in a patient with Balint's syndrome. *J. Cogn. Neurosci.*, **12**: 583-600, 2000.
- Phillips M.L., Medford N., et al. Depersonalization disorder: thinking without feeling. *Psychiatry Res.*, **108**: 145-160, 2001.
- Pollen D.A. On the Emergence of Primary Visual Perception. *Cereb. Cortex*, 2011.
- Rafii M.S. and Brewer J.B. Functional brain imaging in the clinical assessment of consciousness. *PLoS Biol.*, **8**: e1000548, 2010.
- Reder L.M., Park H., et al. Memory systems do not divide on consciousness: Reinterpreting memory in terms of activation and binding. *Psychol. Bull.* **135**: 23-49, 2009.
- Rodriguez Moreno D., Schiff N.D., et al. A network approach to assessing cognition in disorders of consciousness. *Neurology*, **75**: 1871-1878, 2010.
- Rosenthal D.M. Consciousness, content, and meta-cognitive judgments. *Conscious. Cogn.*, **9** (2 Pt 1): 203-214, 2000.
- Rosenthal D.M. *Consciousness and mind*. Oxford, Oxford University Press, 2005.
- Schnakers C., Perrin F., et al. Voluntary brain processing in disorders of consciousness. *Neurology*, **71**: 1614-1620, 2008.
- Schnakers C., Vanhaudenhuyse A., et al. Diagnostic accuracy of the vegetative and minimally conscious state: clinical consensus versus standardized neurobehavioral assessment. *BMC Neurol.*, **9**: 35, 2009.
- Scott R.B., Minati L., et al. Detecting conscious awareness from involuntary autonomic responses. *Conscious. Cogn.*, **20**: 936-942, 2010.
- Seel R.T., Sherer M., et al. Assessment scales for disorders of consciousness: evidence-based recommendations for clinical practice and research. *Arch. Phys. Med. Rehabil.*, **91**: 1795-1813, 2010.



- Seth A.K., Izhikevich E., et al. Theories and measures of consciousness: an extended framework. *Proc. Natl. Acad. Sci. U S A*, **103**: 10799-10804, 2006.
- Seth A.K., Dienes Z., et al. Measuring consciousness: relating behavioural and neurophysiological approaches. *Trends Cogn. Sci.*, **12**: 314-321, 2008.
- Seth A.K. Explanatory correlates of consciousness: Theoretical and computational challenges. *Cogn. Comput.* **1**: 50-63, 2009.
- Seth A.K. A MATLAB toolbox for Granger causal connectivity analysis. *J. Neurosci. Methods*, **186**: 262-273, 2010.
- Seth A.K., Barrett A.B., et al. Causal density and integrated information as measures of conscious level. *Philos Transact. A Math. Phys. Eng. Sci.*, **369**: 3748-3767, 2011.
- Shanahan M. A spiking neuron model of cortical broadcast and competition. *Conscious. Cogn.*, **17**: 288-303, 2008.
- Sierra M. *Depersonalization: A new look at a neglected syndrome*. Cambridge, Cambridge University Press, 2009.
- Soon C.S., Brass M., et al. Unconscious determinants of free decisions in the human brain. *Nat. Neurosci.*, **11**: 543-545, 2008.
- Tononi G. and Edelman G.M. Consciousness and complexity. *Science*, **282**: 1846-1851, 1998.
- Tononi G. An information integration theory of consciousness. *BMC Neurosci.*, **5**: 42, 2004.
- Tononi G. Consciousness, information integration, and the brain. *Prog. Brain Res.*, **150**: 109-126, 2005.
- Tononi G. Consciousness as integrated information: a provisional manifesto. *Biol. Bull.*, **215**: 216-242, 2008.
- Tononi G. and Koch C. The neural correlates of consciousness: an update. *Ann. N Y Acad. Sci.*, **1124**: 239-261, 2008.
- Trenerry M.R. and Loring D.W. Intracarotid amobarbital procedure. The Wada test. *Neuroimaging Clin. N. Am.* **5**: 721-728, 1995.
- Vallar G. Spatial neglect, Balint-Homes' and Gerstmann's syndrome, and other spatial disorders. *CNS Spectr.* **12**: 527-536, 2007.
- Van Boxtel J.J., Tsuchiya N., et al. Consciousness and attention: on sufficiency and necessity. *Front. Psychol.*, **1**: Epub 2010 (Dec 20).
- Vanhaudenhuyse A., Giacino J., et al. Blink to visual threat does not herald consciousness in the vegetative state. *Neurology*, **71**: 1374-1375, 2008.
- Vanhaudenhuyse A., Noirhomme Q., et al. Default network connectivity reflects the level of consciousness in non-communicative brain damaged patients. *Brain*, **133** (Pt 1): 161-171, 2010.
- Zeman A. Persistent vegetative state. *Lancet*, **350**: 795-799, 1997.
- Zhou J., Liu X, et al. Specific and nonspecific thalamocortical functional connectivity in normal and vegetative states. *Conscious. Cogn.*, **20**: 257-268, 2010.