

Review Article

Improving objective assessment in disorders of consciousness: An option for classification technology?

Denis Larrivee*

International Association Catholic Bioethicists, Toronto, Canada

Abstract

For patients suffering disorders of consciousness (DOC), conceptual ambiguities in the definition of consciousness affect diagnostic conclusions on patient status. Current imaging technologies offer the prospect of objective identification of consciousness correlates that can enhance diagnostic accuracy. Classification technologies coupled with parallel developments in artificial intelligence learning add a further dimension to interpretive conclusions by revealing structured information content and distributed causal influences.

Introduction

Consciousness poses an enigma, for which the philosophical resolution remains unsettled [1]. Colloquially, its understanding invokes a subjective awareness of attentionally evoked, externally sensed or internally recollected events. Such awareness must be exercised by a someone, thereby introducing difficulties in how to relate the underlying neural correlates that may contribute to consciousness to a global property that is exercised subjectively. The philosophical need for consciousness to be a subjective property of one individual, for example, has yielded models of consciousness access, rather than consciousness proper, that can be objectively explained in terms of global faculties of the brain and nervous system [2]. For patients suffering disorders of consciousness, these ambiguities impact medical care [3], influencing the choice of diagnostic methodology, prognostic possibilities, and the success of the therapy actually implemented.

Medically, consciousness has been linked to two constructs, arousal, also termed wakefulness, and awareness [4]. Arousal is generally considered to be synonymous with alertness and is conceived in terms of a level of patient responsiveness to a stimulus; clinically, it is assessed by eye opening. Awareness is qualitatively distinguished from arousal and relates to a patient's ability to respond with comprehension to the information content of a presented stimulus. In a clinical setting this distinction is manifest in a patient's ability to engage in non-reflexive motor behavior, such as eye movement in response to a verbal command. Distinctions between the constructs have also been shown to extend to their neural sources, which are located in separate brain regions. Arousal is associated with activities localized to the brainstem; awareness, by contrast, has been linked to activities originating in the fronto-parietal network and associative cortices, as well as in extrinsic connectivities with the thalamus. Despite their distinct anatomical origins, there is a positive relationship between the two, with more aware states generally correlated with more alert ones.

On the basis of the absence or presence of these constructs clinical entities for consciousness disorders have been distinguished [1]. These include coma, unaware wakefulness syndrome (UWS), minimally conscious syndrome (MCS), and Locked in Syndrome (LIS). Coma

is an acute state of unresponsiveness where there is no arousal. Signs of arousal are seen in UWS patients and appear to coincide with the restoration of the brain stem reticular system, and the absence of activity in the neocortex and thalamus. Minimally conscious patients, by contrast, display significant brain activity in these latter regions and limited awareness, such as the following of a verbal command. Patients diagnosed as LIS do not conform to the DOC categorization profile, due to extensive damage to the cortico-spinal and cortico-bulbar pathways that results in the loss of control over all voluntary muscles excepting those involving small eyelid movement.

The association of these constructs with distinct brain regions, notwithstanding, the prismatic and philosophically multidisciplinary character of consciousness has been an important factor in weighting diagnostic determinations of these disorders toward behavioral scales, today's gold standard, that need to be assessed by trained medical staff. Several such scales have been developed to assess the DOC patient. Of those in current use, the American Congress of Rehabilitation Special Interest Group has designated the Coma Recovery scale as having the strongest content validity, based on the Aspen criteria [5]. Revised in 2004 the CRS scale is composed of 23 items arranged hierarchically from reflexive to cognitively mediated responses that address patient sensory, usually visual and auditory, motor, communication, and arousal abilities. Alternate scales in common use, such as the Sensory Modality and Rehabilitation Technique scale (SMART) employed in the United Kingdom, the Western Neurosensory Stimulation Profile, and the Wessex Head Injury Matrix, likewise employ items intended to assess, with varied emphases, sensory, communication, and arousal abilities of the patient [6].

Correspondence to: Denis Larrivee International Association Catholic Bioethicists, Toronto, Canada, E-mail: sallar1@aol.com

Key words: consciousness disorders, neuroimaging, classification technology, neural decoding, artificial intelligence, unsupervised machine learning

Received: November 30, 2017; **Accepted:** December 22, 2017; **Published:** December 26, 2017

Despite widespread reliance on behavioral assessment, nearly 2 in 5 patients are misdiagnosed with respect to the status of consciousness impairment [7]. The salience of this statistic for diagnostic and therapeutic purposes is evident. Patients misdiagnosed as having more severe disorders of consciousness, for instance, are likely not to be offered treatments that have been shown to assist rehabilitation in patients with less severe disorders, like MCS. These realities have led to the search for 'objective' assessments that can supplement, or even supplant, behavioral diagnosis, several of which now see clinical use.

Among the most employed, 'objective' approaches for detecting DOC are the neuroimaging and electrophysiological technologies [8]; these offer to the medical professional the opportunity to study physical features of cognitive processing apart from the behavioral scales, with the latter's inevitable element of subjectivity. Clinical technologies in current use include positron emission tomography (PET), functional magnetic resonance imaging and magnetic resonance imaging (fMRI/MRI), electroencephalography (EEG), and transcranial magnetic stimulation coupled with electroencephalography (TMS/EEG). Emphasizing passive patient participation, these technologies have been used to detect the presence or absence of brain activity in brain regions that are associated with arousal or awareness. While fMRI has generally replaced PET analyses, its use of the blood oxygen level dependent signal (BOLD) as a parameter revealing brain activity is confounded by the considerably slower response times - on the order of several magnitudes - of the BOLD signal relative to the temporal dynamics of the neural activity actually taking place. The electrophysiological technologies, by contrast, are capable of assessing neural activity on physiological time scales, but lack the spatial resolution of the imaging procedures.

The passive paradigm used in these technologies and the indirect inference of consciousness made on the basis of brain activity alone remain caveats to their diagnostic value, which otherwise have greatly improved diagnostic accuracy. Whether observed activity is associated with consciously perceived representations, thus, remains an unknown. Recent developments in imaging technologies, which to date are used largely in the research setting, propose to assist in deciphering mental imagery and patient awareness [9]. These technologies are designed to access information content contained within brain activity, and so offer an advance over the sort of static assessment made with the imaging and electrophysiological technologies used in the customary, clinical setting. This review will consider some current knowledge about these information content, 'decoding' devices, their prospects for clinical use, purported advantages that may assist DOC diagnosis, and prospective developments in related diagnostic technologies.

Current approaches to content based classification technology

It is with the intention of managing the semantic content to be inferred from brain activity, as opposed to making inferences about brain state activation, that qualitatively new approaches to elucidate what brain states actually mean have been undertaken [10-12]. While these new approaches are to date restricted to the research setting, and have generally yet to be exploited as viable medical devices or device/approaches, they nonetheless afford the prospectus for revealing distinguishing features of brain activity patterns that can be correlated with objective features of the world. Although this is not quite the same as representational imagery of the sort needed for intersubjective communication, it does signify an advance over existing imaging techniques in exposing the structure of the information content that the brain may actually be using.

In the 'decoding' approaches that have evolved to date the central technical concern is that of 'classification', that is, mapping a brain state as imaged in its activity pattern with an externalized feature, object, or event. The oldest of these, mass univariate analysis, is based on a general linear model in which sequential brain regions are monitored for specific mental activity at a specific brain location [13]. How the brain represents content remains an unknown. However, it is presumed that mental contents are distributed over populations of cells [14]; hence, there is the presupposition of an underlying neural connectivity architecture uniting them. Classification technology is used thus to measure the covariance between multiple single units, which serves as a determinative as well as a diagnostic feature that is relevant to how select images are encoded. In fMRI imaging, for example, the presentation of a single object will activate long regions of the occipital cortex originating at multiple sites; thus, monitoring covariance is a technical objective thought to link the neural activity patterns to a structured representational content.

Though adding a qualitatively significant dimension to imaging assessment, classification methodologies introduce additional analytical requirements that are less straightforward than static fMRI imaging; approaches to assessing information content have thus been considered from several perspectives. Because the covariant activity patterning is, a priori, an unknown with regard to a specific representation, it is necessary to map the activity distribution to a specific content first. In multivariate decoding brain patterns are assessed on the basis of previously determined correlations between a respective image and the activity pattern [15]. For this purpose it is necessary to 'train' the classifier to label, that is, to associate a brain pattern with a particular image. Training data sets are therefore needed to identify unique patterns of activity so that the test data may correctly identify observed patterns with a specific image. This process is complex due to the presence of intrinsic stochastic and potentially other sources of variation in activity, that therefore require the statistical processing of patterning, often with Gaussian type filters, to assess statistically significant differences or similarities [16]. For classification, processing layers additional complexity into the analysis, since both the test and training imagery are susceptible to stochastic variation. For example in a frequently used classification strategy the data set is halved and the halves designated as training and test sets respectively [17]. Signals from the training set are then compared with their respective counterparts from the test set and the highest correlation between the two is calculated and then designated as the criterial response.

In the multivariate classification approaches, further, the necessity of learning to classify each image representation limits assessment to identifiable images alone, leaving many potential representations unclassifiable. Attempts to overcome this limitation have, accordingly, attempted to expand the range of accessible imagery through an approach known as model-based classification, where classification models are used to 'predict' patterns that may then be associated with an image that is not part of the training data [18]. Applying the model to novel patterns it is possible to extract common informational features within the pattern that can be associated with discrete object elements. This approach can potentially reduce the size of the training set needed, and, conversely, extend the extrapolation to a much broader range of imagery. Such an approach has been used, for example, to greatly expand relationships among word data sets [19].

Another important consideration in classification assessment is that of the time variant nature of representational brain states. Since the brain is continually active image representations are, consequently,

highly variable. Time variant representation affects classification in two ways, in the fluctuation of dynamical brain flow, hence, in the spatial and temporal transitioning of representations; and in the likelihood of causal relations between representational events. Significant parameters that need to be determined in the observed brain activity also include onset and duration, and the boundaries that distinguish temporally independent and successive events. On the other hand the detection of the time evolution of patterning is likely to expose temporal relations among such features that show, for example, phase locking [20] that can reveal synchronized and globally distributed events. Indeed, such correlations have constituted the conceptual basis for current approaches that infer causal relations amongst distinct activity patterns [21], and which require a distributed analysis of the entire brain [22].

Assessing the DOC patient

For the DOC patient the diagnostic value of content based devices rests in their ability to detect a covert awareness. A presupposition that is adopted is that the detection of a covariant pattern in response to an object presentation reveals a conscious awareness. Such a premise is complicated in multiple ways. The need to map a brain pattern to a given image, unlike the observation of passive activity in a given brain region, makes one or more of several additional assumptions. The classification must presume that the patient can respond to a given image, a presupposition likely to be satisfied only in less severe instances of DOC where patients are capable of reacting to an image; or, presuppose that the neural correlates affected by DOC do not affect the neural representation, an unknown; or, be a feature or biomarker common to large groups of individuals. This last, for example, was taken up in a classification study that sought to discriminate healthy individuals from patients suffering unresponsive wakefulness syndrome [23].

Additionally, it must be presumed that the observed activity pattern is one for which the patient has an awareness; that an object is not, for example, detected but unavailable for consciousness, or that the absence of the pattern, conversely, does not mean that the individual is unaware. To overcome such objections experimenters have taken the tack of having the patient demonstrate an ability to actively respond to the presence of an image on the presupposition that this would require the patient's awareness of the image to do so. In one such 'active' study a limited number of patients were capable of specifically modulating motor brain activity as assessed by BOLD responses in otherwise ambiguous behavioral diagnoses [24]. This was interpreted to indicate that the technology correctly assessed an uncertain behavioral diagnosis in a small cohort of patients. Other studies have attempted to elicit activity in the premotor cortical area, that is, activity interpreted as signals preparatory for overt motor actions, hence, they identified an activity more closely associated with conscious and willful thought processes. While the studies were numerically small, the ability to detect a specific activity in a correct region of the brain associated with willful action suggests that the classification technology possesses a potential for detecting a covert awareness, a feature that may potentially be exploited for development.

Future prospects

Given the ambiguity inherent in behavioral diagnosis an important goal of advancing classification technology is not only the improvement of communicating content, hence of pinpointing known features that are aberrant in consciousness disorders, but also how the technologies may in turn reveal previously uncharacterized features of consciousness that may be impacted by the impairments. Several areas now under active scrutiny include 1) expanding the range of active paradigms in

which the DOC patient can be enabled to assist in eliciting information about his status, 2) clarifying how content based information is causally related, distributed, or mutually influencing, and 3) appropriating parallel developments in related, assistive technologies such as predictive learning and artificial intelligence. In the first, eliciting the active participation of the patients has been to date restricted to MCS states where it is possible to obtain positive responses to command requests. Expanding elicitation batteries, however, may help to extend to UWS the ability to discriminate among smaller, more discrete variations within the syndrome category. For example, existing electro potential methods that rely on the evoked potential currently exploit multisensorial stimulation, such as combinations of auditory and tactile stimuli, which have been used to distinguish between UWS and LIS [25]. While the BOLD signal poses a physical obstacle to a similar level of temporal resolution, improved summation detection could better detect small differences between voxels that may be revealed through the concerted use of paradigm changes, electropotential technology, and neuroimaging. Clarifying how activity cells are related, second, may serve to identify not only pattern biomarkers for neural correlates of consciousness, but establish functional and performance links among spatially discrete events. Oscillation profiles, for instance, are widely suspected not only to provide global coherence, of the sort needed for consciousness [26], but also to entrain local circuit assemblies, such as those that are postulated to be the basis of memory retrieval [27]. Perhaps most significant, third, are improvements in the processing capabilities of classification through a concerted merger with advances in machine learning and artificial intelligence. Existing training on data sets now routinely use predictive, unsupervised learning, methods that significantly expand the range achieved by reinforcement learning strategies alone. Recent advances now emphasize learning paradigms that simulate the natural learning of neural networks, a strategy termed unsupervised learning [28]. This learning approach does not rely on the isolation of discrete elements within data samples as in supervised or predictive learning, but rather extracts correlations that enable the reconstruction of data samples *de novo*, further expanding the range of detectable imagery.

Taken together, classification technology offers a qualitatively new dimension in the diagnosis of the DOC patient. While temporal and signal processing complexities continue to restrict the significance of diagnostic conclusions, dynamic paradigms and expanded possibilities for interpretive accuracy through new learning architectures offer a previously unavailable window for clinicians to assess consciousness disorders.

References

1. Eapen BC, Georgekutty J, Subbarao B, Bavishi S, Cifu DX (2017) Disorders of consciousness. *Phys Med Rehabil Clin N Am* 28: 245-258. [[Crossref](#)]
2. DeHaene S (2015) Consciousness and the Brain: Deciphering How the Brain Codes our Thoughts. *J Undergrad Neurosci Educ* 12. [[Crossref](#)]
3. Di Perri C, Thibaut A, Heine L, Soddu A, Demertzi A, et al. (2014) Measuring consciousness in coma and related states. *World J Radiol* 6: 589-597. [[Crossref](#)]
4. Laureys S (2005) The neural correlate of (un)awareness: lessons from the vegetative state. *Trends Cogn Sci* 9: 556-559. [[Crossref](#)]
5. Giacino JT, Fins JJ, Laureys S, Schiff ND (2014) Disorders of consciousness after acquired brain injury: the state of the science. *Nat Rev Neurol* 10: 99-114. [[Crossref](#)]
6. Gill-Thwaites H, Munday R (2004) The sensory modality assessment and rehabilitation technique (SMART): a valid and reliable assessment for vegetative state and minimally conscious state patients. *Brain Inj* 18:1255-1269. [[Crossref](#)]
7. Schnakers C, Vanhauwenhuyse A, Giacino J, Ventura M, Boly M, et al. (2009) Diagnostic accuracy of the vegetative and minimally conscious state: clinical consensus

- versus standardized neurobehavioral assessment. *BMC Neurol* 9: 35. [[Crossref](#)]
8. Gosseries O, Pistoia F, Charland-Verville V, Carolei A, Sacco S, et al. (2016) The role of neuroimaging techniques in establishing diagnosis, prognosis and therapy in disorders of consciousness. *Open Neuroimag J* 10: 52-68. [[Crossref](#)]
 9. Haynes JD (2013) Decoding mental states from patterns of brain activity. In: Spencer JP, Thomas MSC, McClelland JL, editors. *Toward a Unified theory of Development: Connectionism and Dynamic Systems Theory Reconsidered*. Oxford: oxford University Press.
 10. Norman KA, Polyn SM, Detre GJ, Hawby JV (2006) Beyond mind reading: multi-voxel pattern analysis of fMRI data. *Trends in Cog Sci* 10: 424-430.
 11. Lemm S, Blankertz T, Dickhaus T, Muller KR (2011) Introduction to machine learning for brain imaging. *NeuroImage* 56: 387-399. [[Crossref](#)]
 12. Kriegeskorte N (2011) Pattern information analysis: from stimulus decoding to computational-model testing. *NeuroImage* 56: 411-421. [[Crossref](#)]
 13. Haynes JD, Rees G (2006) Decoding mental states from brain activity in humans. *Nat Rev Neurosci* 7: 523-534. [[Crossref](#)]
 14. Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, et al. (2001) Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293: 2425-2430. [[Crossref](#)]
 15. Kriegeskorte N, Goebel R, Bandettini P (2006) Information based functional brain mapping. *Proc Natl Acad Sci U S A* 103: 3863-3868. [[Crossref](#)]
 16. Pereira F, Mitchell T, Botvinick M (2009) Machine learning classifiers and fMRI: a tutorial overview. *NeuroImage* 45: 199-209. [[Crossref](#)]
 17. Misaki M, Kim Y, Bandettini PA, Kriegeskorte N (2010) Comparison of multivariate classifiers and response normalizations for pattern information fMRI. *NeuroImage* 53:103-118. [[Crossref](#)]
 18. Nevado A, Young MP, Panzeri S (2004) Functional imaging and neural information coding. *NeuroImage* 21:1083-1095. [[Crossref](#)]
 19. Mitchell TM et al (2008) Predicting human brain activity associated with the meaning of nouns. *Science* 320:1191-1195.
 20. Lachaux JP, Rodriguez E, Martinerie J, Varela FJ (1999) Measuring phase synchrony in brain signals. *Hum Brain Mapp* 8: 194-208. [[Crossref](#)]
 21. Kahan J, Urner M, Moran R, Flandin G, Marreiros A, et al. (2014) Resting state functional MRI in Parkinson's disease: the impact of deep brain stimulation on 'effective' connectivity. *Brain* 137: 1130-1144. [[Crossref](#)]
 22. Deco G, Jirsa V, Friston K (2013) The dynamical and structural basis of brain activity. In: Spencer JP, Thomas MSC, McClelland JL, editors. *Toward a Unified theory of Development: Connectionism and Dynamic Systems Theory Reconsidered*. Oxford: Oxford University Press
 23. Phillips CL, Bruno MA, Maquet P, Boly M, Noirhomme Q, et al. (2011) Relevance vector machine consciousness classifier applied to cerebral metabolism of vegetative and locked-in patients. *NeuroImage* 56:797-808. [[Crossref](#)]
 24. Gosseries O, Pistoia F, Charland-Verville V, Carolei A, Sacco S, et al. (2016) The role of neuroimaging techniques in establishing diagnosis, prognosis and therapy in disorders of consciousness. *Open Neuroimag J* 10: 52-68. [[Crossref](#)]
 25. Schnurer A, Espinosa A, Guger C (2015) P125 mindBEAGLE: a BCI for communication and assessment of consciousness for patients with disorders of consciousness. *Clinical Neurophysiol* 126: 152-153.
 26. Buzsaki G (2006) *Rhythms of the Brain*. Oxford: Oxford University Press
 27. Kaplan R, Adhikari MH, Hindriks R, Mantini D, Murayama, et al. (2016) Hippocampal sharp-wave ripples influence selective activation of the default mode network. *Curr Biol* 26: 686-691. [[Crossref](#)]
 28. Lerch-Hostalot D, Megias D (2016) Unsupervised steganalysis based on artificial training sets. *Engineer Applications Artificial Intelligence* 50: 45-59.