CHAPTER 2

How can we know if patients in coma, vegetative state or minimally conscious state are conscious?

Morten Overgaard*

CNRU, Hammel Neurorehabilitation and Research Unit, Aarhus University Hospital, Hammel, Denmark

Abstract: This paper examines the claim that patients in coma, vegetative state and minimally conscious state may in fact be conscious. The topic is of great importance for a number of reasons — not least ethical. As soon as we know a given creature has any experiences at all, our ethical attitude towards it changes completely. A number of recent experiments looking for signs of intact or partially intact cognitive processing in the various stages of decreased level of consciousness are reviewed. Whether or not vegetative or coma patients are in fact conscious is an empirical issue that we yet do not know how to resolve. However, the simple fact that this is an unresolved empirical issue implies that the standard behavioural assessment is not sufficient to decide what it is like to be these patients. In other words, different and more sophisticated methods are necessary. From a theoretical position, the paper moves on to discuss differences in validity between reports (e.g. verbal) and signals (e.g. brain activations) in the study of consciousness, and whether results from experiments on the contents of consciousness may be of any use in the study of levels of consciousness. Finally, an integrated approach is suggested, which does not separate research in level and content as clearly as in current practice, and which may show a path to improved paradigms to determine whether patients in coma or vegetative state are conscious.

Keywords: coma; vegetative state; minimally conscious state; consciousness; experience; neural correlates of consciousness

Introduction

This paper will consider the seemingly controversial hypothesis that patients in coma, vegetative or minimally conscious state (MCS) may in fact have conscious experiences. It is a typical opinion in current neuroscience that the absence of reports or clear neurophysiological markers of consciousness in these patient groups place the burden of proof lies with the claim that there is any conscious experience left in coma or the vegetative state (VS) (Giacino and Smart, 2007). However, since we have no certain neurophysiological or behavioural markers for the absence of consciousness either, one could — at least for the sake of the argument — take on the opposite stance without violating any logical imperatives; that is there is no claim necessitated by the reason that when a given individual cannot behave in certain ways (or behave at all), then that individual can have no subjective experiences.

The question is of great importance for a number of reasons. For instance, our ethical
considerations are specifically directed at conscious beings. That is, we have no ethical problems cutting wood or kicking a football as we are convinced that these objects have no experience of pain. As soon as we know a given creature has any experiences at all, our ethical attitude towards it changes completely.

The patients

Three distinct “stages” of decreased consciousness have been described — coma, the VS and the MCS. The distinction between the stages is based on behavioural criteria. VS patients are generally thought to differ from comatose patients as coma patients can be aroused, yet they are believed to be equally unconscious (Schiff, 2005). MCS patients, however, are believed to have “some” or “fluctuating” consciousness. Other patients with severe brain injury who, however, are not in MCS, are typically believed to be “more conscious”, yet in some cases “less conscious” than healthy people. Consciousness is thus considered gradual and not necessarily stable — and measurable by different aspects of overt behaviour.

Coma is generally believed to be a state of constant, continuous unconsciousness in which the eyes remain closed and the patient cannot be aroused. The eyes remain closed, there seems to be a total absence of voluntary behaviour or any kind of purposeful motor activity or expressive language ability, and no sleep/wake cycles can be identified. There seems to be a total absence of voluntary behaviour or any kind of purposeful motor activity or expressive language ability. The comatose state almost always resolves within 2–4 weeks, leading either to the patient’s death or an improved level of consciousness.

The appearance of spontaneous eye opening marks the onset of VS. In VS, eyes are open but there is no evidence of sustained or reproducible purposeful behaviour, responses to sensory stimuli and no evidence of language comprehension. The term persistent vegetative state (PVS) refers to an ongoing VS lasting at least 1 month from the time of onset. When VS persists for one year after traumatic brain injury or three months following other types of brain injury, it is generally considered highly unlikely that the individual ever recovers there, seemingly, lost consciousness. Most research on patients with a reduced level of consciousness rests on the assumption that many of these patients, and certainly all of those in coma and VS, are fully unconscious. One central example is Laureys et al. (2000, see also Laureys, 1999) where brain activity recorded from a patient in VS was contrasted with that from healthy controls and, subsequently, with his own brain activity post-recovery. Analysis of cortico-subcortical coupling showed that, in contrast to when the patient was in VS, both healthy controls and the patient on recovery had a specific pattern of cortico-thalamic activity. This is in turn used to suggest that this pattern of coupling is part of the neural correlate of consciousness — given, of course, that the VS patients are in fact fully unconscious.

MCS is distinguished from VS by the presence of one or several signs of knowledge about self or the environment; for example the following of simple commands, recognizable verbal or gestural yes/no-responses (accurate or not) or movements that seem to be beyond mere reflexes. MCS typically occurs as a progression from VS, but may also be observed during the course of progressive decline in neurodegenerative diseases. Although MCS may involve reactions to emotional stimuli or reaching toward objects placed in the immediate visual field, the general assumption in neurological wards appears to be that these patients are “less conscious” than are healthy subjects. The assumption is not just they have decreased cognitive functions or, due to their impairments, are conscious of fewer things, but their consciousness itself is somehow diminished. Although one should think that such a claim should be supported by literature discussing what it is like to be in MCS (or, for that matter, in VS), this is extremely rare (see however Laureys and Boly, 2007). Instead, MCS is discussed in terms of behavioural and/or neurological signs only. Emergence from MCS is signalled by the recovery of some kind of meaningful interaction with the environment affording the assessment of higher cognitive functions.
Such criteria for diagnosing hypothesised levels of consciousness do not stand without criticism. Taylor et al. (2007) have suggested that the requirements for reliable communication and functional object use confuse central aspects of the posttraumatic amnesia syndrome (PTA) with MCS. The loss of executive control during PTA may cause communicative difficulties so that an “actual” emergence from MCS goes unnoticed. But although the clinical criteria for establishing these supposedly distinguishable levels of consciousness are quite debated (see also Giacino and Smart, 2007), the robustness of the levels themselves, curiously enough, are uncritically accepted from research papers to neurological wards. This is particularly interesting as, in the absence of a verifiable or merely consensual operationalisation of consciousness, clinical assessment currently relies on the strictly pragmatic principle that people can only be considered to be unequivocally conscious if they can report that this is indeed the case. Thus, the discrimination between VS and MCS, and, in effect, the discrimination between states of conscious and unconscious being, depends upon such communication. Quite obviously, this approach is seriously flawed and represents a central, if not the crucial, problem in the study of decreased levels of consciousness.

Signs of consciousness?

A number of recent experiments have looked for signs of intact or partially intact cognitive processing in the various stages of decreased level of consciousness in the absence of any behavioural signals or communication. This has been done by looking for neural signals, such as event-related potentials (ERPs) or patterns of functional brain activation, typically associated with conscious cognitive processing in healthy individuals.

Some ERP studies have focussed on the P300 response, a positive wave elicited 300 ms after a stimulus, which is usually seen when a subject detects a “surprising” (unpredicted) stimulus in a train of other stimuli, for example in an “oddball paradigm” (Sutton et al., 1965). One sub-component of P300 is the P3b amplitude, which seems sensitive to the importance of the stimulus to the subject — for example the subject’s own name (Perrin et al., 1999). Thus, P300 is typically associated with attentional discrimination, anticipation and emotional states. Signorino et al. (1997) used, in one experimental condition, a conventional auditory oddball paradigm and, in another, a paradigm in which the tones were coupled to emotional verbal stimuli. P300 responses were obtained in 36–38% of comatose patients in the first condition, and in 52–56% in the second. Other experiments have confirmed that emotional stimuli evoke a larger P300 than do meaningless stimuli (e.g. Lew et al., 1999), suggesting that even comatose patients process auditory stimuli to a semantic level. One study (Perrin et al., 2006) found that the patient’s own name elicited stronger P3 responses in VS patients than do other names, and, interestingly, found no significant differences between VS and MCS patients in this regard. Obviously, this is of specific interest given the common conception that the difference between these patient groups marks the difference between being conscious and unconscious.

Other ERP studies have focused on the N400 potential. The N400 seems less related to focused attention and more related to verbal stimuli discordant to preceding verbal stimuli (Vanhaudenhuyse et al., 2008). Schoenle and Witzke (2004) presented different patient groups with semantically congruent and incongruent sentences while recording ERP and found an N400 response to incongruent words in 12% of the vegetative population, 77% in a population named “near-VS” and in 90% of a population with “severe brain damage” (probably MCS).

A number of brain imaging studies in VS patients have, in addition, shown that areas of the brain increase their metabolic activity in response to sensory stimuli — for example the auditory processing areas of such patients might be activated in response to hearing a familiar voice such as their name (Perrin et al., 2006). Owen et al. (2006) used fMRI to study visual imagery in a patient fulfilling all the behavioural criteria for a diagnosis of VS. This 23-year-old
woman sustained a severe brain injury in a traffic accident. After an initial comatose state, she opened her eyes and demonstrated sleep–wake cycles. However, even during the waking periods, she was unresponsive to stimuli and did not manifest spontaneous intentional behaviour. In an experiment, the patient was asked to perform two mental imagery tasks — either to imagine visiting the rooms in her home or to imagine that she was playing tennis. Patterns of brain activation observed using fMRI during each task were indistinguishable from those recorded from a group of conscious control subjects.

It seems impossible to explain these results without accepting that this patient retained the ability to comprehend verbal instructions, to remember them from the time they were given (before scanning began) to the appropriate time during the scan itself, and to act on those instructions, thereby willfully producing specific mental, or at least neural, states. It may be tempting to dismiss this as a simple case of error in the behavioural assessment of her as vegetative, but examination of the exhaustive case report reveals this as unlikely. Indeed, at testing, the patient exhibited no evidence of sustained or reproducible purposeful behaviours consistent with the criteria defining the MCS. The diagnosis of VS was thus entirely appropriate, given current criteria.

One way to oppose the view that this patient and, as a logical consequence, perhaps all other patients diagnosed as vegetative are in fact conscious, would be to argue that the neural activations only represent unconscious cognitive processes involved in the mental task. The fact that the healthy subjects had vivid experiences of their visual imageries would accordingly rely on other brain processes as those observed to be shared with the vegetative patient.

Whether or not this patient, or all patients in VS, is in fact conscious is an empirical issue that we yet do not know how to resolve. However, the simple fact that this is an unresolved empirical issue implies that the standard behavioural assessment is not sufficient to decide what it is like to be in VS. In other words, different and more sophisticated methods may be necessary.

Conscious states and conscious levels

There seem to be persuasive arguments indicating that patients with impaired consciousness are not merely able to passively receive external stimuli, but that they are able to perform distinctly different kinds of cognitive processing. Current debates about conscious and unconscious cognitive processing are centred on studies of conscious content rather than levels of consciousness. Even though this distinction is widespread, both in definitions and in actual research, it may not be fruitful as discussed in section below. For this reason, I will briefly summarise relevant discussions from the content approach to aid the ongoing debate about levels.

Studies of conscious content seek to identify those specific factors that make a subject conscious of something rather than something else (e.g. the taste of coffee or the visual impression of a tree). Typically, this is done by comparing brain states in conditions where a specific conscious content is present to conditions where it is absent. Studies of levels of consciousness also look for enabling factors (using the terminology of Koch, 2004) making it possible to be conscious at all. Here, differences between different states such as being awake, being in dreamless sleep or in a coma are typically compared.

The research strategy currently dominant in consciousness studies per se is the identification of neural correlates of consciousness (NCC). A term coined by David Chalmers (2000), the NCC for content consciousness is those minimally sufficient neural conditions for a specific (mostly representational) content. The basic methodology was set out early by Baars (1988) as a contrastive analysis between being conscious (i.e. having specific conscious content) and unconscious (i.e. having this content in an unconscious form), thus either identifying (a) equal levels of performance, accompanied by different degrees of awareness (e.g. blindsight), (b) changes in performance unaccompanied by changes in awareness (e.g. implicit learning) and (c) changes in awareness despite stimulation remaining constant (e.g. binocular rivalry). A classic example of subliminal abilities is the phenomenon of “blindsight”.

Conscious states and conscious levels
Blindsight refers to the observations that at least some patients with lesions to the primary visual cortex resulting in blindness have nevertheless preserved such visual functions such as perception of movement direction (Weiskrantz et al., 1995), target detection (Pöppel et al., 1973) and spatial summation (Leh et al., 2006) even though they report to be fully blind in that part of the visual field corresponding to the location of the injury. As such, blindsight might be considered “less interesting” than subliminal perception in healthy subjects, as the phenomenon has so far only been studied in a few patients. However, in those patients, blindsight has proven so consistent and persuasive as an example of an almost unbelievable discrepancy between subjective report and behavioural reactions (such as the ability to discriminate) that many researchers see it as the primary source of evidence for subliminal processing. In 1986, however, Weiskrantz and co-workers found evidence which argues that blindsight should be subdivided into two “types” — type 1 and type 2. Type 1 blindsight patients are characterised, as above described, that is by preserved visual functions despite verbal reports of having no visual experiences. Type 2 blindsight patients report seeing after-images or “shadows” when presented with stimuli.

Ramsøy and Overgaard (2004) developed a new approach to introspective reports of conscious and unconscious processes. Subjects were here asked to create their own categories for subjective reports during long training sessions. They were asked what they were shown and how they experienced stimuli in terms of clarity. Here, stimuli were simple visual figures (triangles, circles or squares) presented in one of three possible colours (blue, green or red). In the study, subjects conformed to a four-point scale categorised as “not seen”, “weak glimpse” (meaning “something was there but I had no idea what it was”), “almost clear image” (meaning “I think I know what was shown”) and “clear image”. When subjects tried to use more than four categories in the scale, they found it confusing and quickly abandoned the extra categories. In the experiment, after the category-generating training process, subjects found the categories easy to use, and in free reports, they explained that the categories seemed very straightforward. Ramsøy and Overgaard showed that in an experimental design where one should expect to find subliminal perception, there was in fact none. In a later study using this scale (named Perceptual Awareness Scale, or PAS), two different report methods were compared directly to investigate subliminal perception (Overgaard et al., 2006). Again, it was found that PAS fully eliminated subliminal perception, which was otherwise heavily present using binary reports. Even more problematic for the concept of unconscious perception is a recent study by Overgaard et al. (2008) presenting a blindsight patient, GR, who exhibits the subliminal capabilities associated with blindsight using a dichotomous report. However, when the patient was asked to report using PAS, there was a significant correlation between correctness and consciousness in her “blind” field, just as in her “healthy” field. Essentially, these experiments indicate that subliminal perception at least in some cases is a methodological artefact based on flawed methods to study conscious states.

As argued by Overgaard and Timmermans (in press), subliminal perception may not be a real phenomenon at all. Instead, subliminal perception may be an artefact of (a) the result of objective measures that can be reduced to other behavioural measures and the a priori assumption of congruence between sensitivity and consciousness, and (b) crude subjective measures (e.g. dichotomous or arbitrary 10-point scales) which claim to measure subjectively conscious experiences, but that presumably do not succeed. The notion of “unconscious cognitive processing” has had a turbulent history in psychology, and it is, to say the least, an open question how to interpret the current status of concepts like “unconscious” and “subliminal”.

Returning to the question of levels of consciousness, two things are suggested by this line of argument. First, the research discussed above indicates that we currently have no certain knowledge that totally unconscious cognitive processing exists — or, at least, how common unconscious processing is. This, in and of itself, casts further doubts on the interpretation of the
Owen et al. (2006) experiment that the vegetative patient had an “unconscious version” of the same cognitive process as the healthy subjects did. Secondly, it becomes evident that current methods used to study conscious content are intimately linked to introspective reports: How we ask subjects what they experience is crucial. Although it has been argued that there can be objective measures of consciousness (Persaud et al., 2007) that do not need to involve subjective reports at all, these suggestions are all methodologically flawed (Overgaard and Timmermans, in press). Arguing, say, that some objective method lends a “more direct” insight into the contents of consciousness than does a subjective report rests upon circularity (Overgaard, 2006). That is, associating a certain objective measure such as the ability to perform correct identifications of stimuli with consciousness is only possible based on empirical evidence, that is a correlation between this performance and the relevant conscious state. Since the conscious state cannot in itself be observed from the outside, the use of a subjective report about the relevant state seems the only possible methodology. Accordingly, no other kind of response can be a more reliable indication of a given conscious state than the subjective report itself. The objective performance correlated with the subjective report, given this correlation is perfect, is thus exactly as valid a measure of consciousness as the report itself.

Reports and signals

As argued above, the study of consciousness from a methodological point of view is a study of reports. Obviously, not all experiments are designed in such a way that it is practical — or possible — to use verbal reports with the explicit content “I am now conscious of...”. In many cognitive psychology experiments, subjects are asked to press a button or give some sort of behavioural gesture to report. In other situations, we may have to suffice with bodily or behavioural signals which may be interpreted as signs of consciousness, such as increased arousal, reflexes or neural activations. To make correct use of these different types of data, however, we need a closer look at their respective validity.

There is, initially, an important distinction between reports and signals. A report is an intended communication from a conscious subject. That is, it involves a subject with metacognitive insight in their own conscious content and the intended, self-controlled giving of information about this content. A signal lacks this intention and is outside the control of the subject. A signal may be any kind of information obtained from the subject that previous research has indicated can be correlated with consciousness — typically, this will be data from technological measurement techniques such as brain scanners, EEG, eye tracking, galvanic skin response, or, more rarely, the observation of uncontrolled behaviours such as reflexes.

As already mentioned, consciousness is subjective. That is, we may have insight into the contents of our own consciousness, but no existing method lends such insight into the contents of the consciousness of other individuals. One may in fact argue that this will always be the case, in spite of any possible technological achievement, as any kind of representation of the experience of other people will always be perceived or looked at from one’s own point of view, thus missing the very essence of what it is like to be this other subject. For this reason, reports are indirect evidence of a given conscious content. Nevertheless, they get us as close as we may come. Signals are even more indirect and much more dubious. First of all, even if a perfect correlation is established between a specific signal and a report, it is not possible to test the correlation in all possible situations. Thus, it is always an open possibility that the signal in some cases may fail as an indicator of consciousness. Following this reasoning, neither signals nor reports may count as measures of consciousness, but as indicators only.

When studying non-communicating patients, we only have signals. As hopefully is made clear from the discussion above, state of the art research and debate makes this a highly difficult yet not necessarily impossible situation. Although we at the present do not have a finished research paradigm to handle this situation, some pitfalls and possibilities can be identified.
Examining the central Owen et al. study, their findings are of no less interest in the light of the report–signal distinction, keeping in mind that the patient, as well as the healthy subjects, participated in cognitive tasks demanding voluntary control and insight into the contents of one’s own consciousness. That is, in the experiment, the patient and the healthy subjects were both asked to sustain their visual imagery for approximately 30 s and to stop when requested to rest. Although this is far from any proof, it gives us reason to speculate that the patient could have reported the contents of her conscious state, were she physically able to, as insight and control are the essential features distinguishing reports from signals.

The complexity of these issues is obviously difficult to handle — even for researchers specialising in these matters. Thus, there are conflations even in the foundational issues of how to interpret neural activations obtained from the patients. In the Owen et al. study, brain signals are used to discuss whether VS patients are conscious or not. This stands in clear contrast to the approach used in Laureys et al. (2000), where cortico-thalamic activations were suggested as parts of the NCC because VS patients are not conscious: Is the answer to the first question positive, the Laureys et al. approach is invalid. Is the Laureys et al. approach somehow shown valid, the Owen et al. study can no longer be interpreted to suggest that VS patients are conscious. At best, we may argue that Owen et al. have examined a very special and unique case story with no general implications.

As we have no prima facie reasons, nor any empirical evidence, to conclude that vegetative patients are not conscious, we should however at present avoid experiments accepting this as a background assumption. For this reason, at least until we have found a way to settle the issue whether patients in arguably reduced states of consciousness are in fact conscious, the whole contrastive approach to finding neural correlates to levels of consciousness is problematic — at least insofar as patients are involved in the research. For this reason, the approach of Owen et al. seems of exceptional value.

Future directions

Jakob Hohwy (2009) has recently argued that a specific conflation exists in those research paradigms looking for neural correlates of conscious contents and those looking for neural correlates of levels of consciousness. The content approach assumes that subjects are already conscious. That is, while the research goal seems to be finding the neural correlate of consciousness per se, experiments actually look for the neural correlate of the selection of specific mental content for conscious experience, rather than it not being selected, and, further, implied that the subject or animal under investigation is already in an overall conscious state. The approach assumes there are both conscious and unconscious contents in an otherwise conscious subject, and that something (in the brain) differentiates between the selection of content for consciousness. While these are questions that research may actually answer, the approach does not inform us what may be the neural correlates of being conscious at all — as this is already assumed and thus not a relevant variable in experiments.

With the other approach, looking specifically for conscious levels without studying conscious content, the idea is to intervene on a creature’s overall conscious state in conditions where there is no conscious content at all. For instance, the philosopher John Searle (2004) insists in the primacy of consciousness. However, very few would agree with Searle that this viable reasoning is anything but theoretical speculation.

Justifiably, Hohwy suggests a combination of the two kinds of approaches, yet gives no exact suggestions how this could be done in real life. One approach that deserves special interest in this regard is the recent advances in neuroimaging attempting to decode a person’s conscious experience based only on non-invasive measurements of their brain activity. Recent work (Haynes and Rees, 2006) demonstrates that pattern-based decoding of BOLD contrast fMRI signals acquired at relatively low spatial resolution can successfully predict the perception of low-level perceptual features. For example, the orientation, direction of motion and even perceived colour of
a visual stimulus presented to an individual can be predicted by decoding spatially distributed patterns of signals from local regions of the early visual cortex. Strikingly, despite the relatively low spatial resolution of conventional fMRI, the decoding of image orientation is possible with high accuracy and even from brief measurements of primary visual cortex (V1) activity.

Perceptual fluctuations during binocular rivalry can be dynamically decoded from fMRI signals in highly specific regions of the early visual cortex. This was achieved by training a pattern classifier to distinguish between the distributed fMRI response patterns associated with the dominance of each monocular percept. The classifier was then applied to an independent test dataset to attempt dynamic prediction of any perceptual fluctuations. Dynamic prediction of the currently dominant percept during rivalry was achieved with high temporal precision.

This approach holds out the promise of achieving important improvements in patients with claimed reduced levels of consciousness. Current experiments using ERP or fMRI, as reviewed above, investigate cognitive processes that may exist consciously as well as unconsciously. Thus, one may raise the criticism that we learn nothing of the patients’ possible conscious contents with these approaches — we may in fact be studying cognitive processes that occur fully unconsciously. However, using a “decoding approach”, one may decode neural patterns specific for conscious content (e.g. in a binocular rivalry paradigm) as verbally verified by healthy subjects able to report. If a strong report–signal correlation can be found, the experiment can be applied to comatose or VS patients looking for similar activations. Although such an approach, even with a 100% match between patients and controls, cannot be said to finally prove conscious experience in coma or VS, it will utilise the reflections above to get us as far as it has here been claimed possible.

Acknowledgment

Morten Overgaard was supported by The Mind-Bridge project, co-funded by the European Commission under the Sixth Framework Programme, Contract No. 043457.

References


