BRAIN-DAMAGE AND THE MORAL SIGNIFICANCE OF CONSCIOUSNESS

Abstract

Neuroimaging studies of brain-damaged patients diagnosed as in the vegetative state suggest that they might be conscious. This might seem to raise no new ethical questions given that in related disputes both sides agree that evidence for consciousness gives strong reason to preserve life. We question this assumption. We clarify the widely held but obscure principle that consciousness is morally significant. It is hard to apply this principle to difficult cases given that philosophers of mind distinguish between a range of notions of consciousness and that it is unclear which of these is assumed by the principle. We suggest that the morally relevant notion is that of phenomenal consciousness, and then use our analysis to interpret cases of brain damage. We argue that enjoyment of consciousness might actually give stronger moral reasons not to preserve a patient’s life, and indeed that these might be stronger when patients retain significant cognitive function.

Keywords

VEGETATIVE STATE; MINIMALLY CONSCIOUS STATE; CONSCIOUSNESS; BRAIN-IMAGING; TERMINATION OF TREATMENT

1. The Vegetative State

In recent decades, modern medical technology and resuscitation techniques have increased the number and prevalence of neurological syndromes involving severe cognitive and motor disabilities, of which the vegetative state (VS), first defined in 1972, is a prime example. VS refers to a post-comatose state secondary to profound brain damage, typically due to hypoxia or traumatic brain injury. Patients in this state recover to the extent that they usually do not require respiratory support. However, although they appear to be awake—they typically open their eyes and have discernible sleep-wake cycles—they fail to provide any evidence of awareness of self or environment. Such patients may respond to painful stimuli with reflex flexion but show no voluntary response to external stimuli. Such patients appear to be in a form of “eyes-opened unconsciousness” where there is a disassociation between wakefulness and awareness.¹

The diagnosis of VS is based upon repeated failed attempts to elicit voluntary responses from a patient. VS patients should be distinguished from patients in the minimally conscious state (MCS) where brain damaged patients do manifest some intermittent or minor responses to stimuli such as visual tracking, or responding to questions with a gesture or word.² The distinction between VS and MCS is prognostically important. Patients who are in VS for more than 6 months after sustaining brain injury have only a very small chance of significant recovery.³ Prognosis for patients in MCS is more variable.⁴

¹Multi-Society Task Force on PVS, 1994, Howsepiam 1996.
²Bernat 2007.
³When VS has been present continuously for a period of more than 12 months following traumatic brain injury or 3 months with a non-traumatic cause, the likelihood of recovery is extremely small and the patient is described as permanently vegetative.
⁴Bernat 2007.
2. Brain imaging evidence for consciousness?

VS is diagnosed when there is no evidence of consciousness. But as science advances, new ways of detecting consciousness are emerging. Recent neuroimaging research in VS patients has provided evidence that some patients previously diagnosed as in VS might in fact possess some measure of consciousness despite the absence of any observable behavioural evidence.

A number of brain imaging studies in VS patients have already 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 say their name. The findings of a recent functional magnetic resonance imaging (fMRI) study by Owen and his colleagues are even more striking. This study focused on a 23-year-old woman who sustained a severe brain injury in a traffic accident. After an initial comatose state, this patient 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, signs considered diagnostic of a vegetative state.

In one of their experiments, the authors tried to engage the patient in two mental imagery tasks by asking her either to “imagine visiting the rooms in your home” or to “imagine playing tennis.” Patterns of brain activation observed using fMRI during each task were highly suggestive of an active mental performance relevant to the task: the observed brain activation patterns were the classic neural correlates of these two mental imagery tasks, and statistical parametric maps of brain activation were indistinguishable from those recorded from a group of conscious control subjects.

What is established by these findings? The authors of the study are quite categorical: they claim that this pattern of response “confirmed beyond any doubt that she was consciously aware of herself and her surroundings.” They consequently argue that their work raises the serious possibility that other brain damaged patients have been wrongly diagnosed as being in the VS.

If these claims are correct, then these findings have far reaching implications. Some of the most heated moral disputes of our time have revolved around legal right to die cases of patients diagnosed as VS such as the 1975 Karen Quinlan case, and, most recently the 2005 Terry Schiavo case. There is little doubt that Owen et al’s findings and forthcoming fMRI research will have a major impact on future moral and legal disputes about withdrawal of life-sustaining treatment from brain-damaged patients. These findings will be cited as evidence for uncertainty in assessments of whether or not patients who appear to be unconscious are truly unconscious, and it seems likely that fMRI will increasingly be called upon to provide definitive evidence of the absence or presence of consciousness in disputed cases.

Although these findings are likely to have such an impact on legal practice, we ought to be very cautious in drawing their moral implications. To start with, as Owen et al themselves insist, there are many cases of VS where the brain damage is so extensive that it is extremely unlikely that there are any prospects of consciousness. These would plausibly include some

5See e.g. Di 2007, Perrin 2006, 2006. 6Owen 2006, 2008. 7Owen 2006. As Owen et al point out, only positive results can provide definite evidence for the presence of consciousness. Negative results are compatible with the patient being conscious but not responding at that given moment, or to that stimulus, or not in a way detectable by that particular technique. 8For discussion of these legal cases, see Fine 2005, and Cranford 2005. The early disputes concerned the withdrawal of respirators and other sophisticated life-sustaining technology, but recent ones centered on the withdrawal of artificial nutrition and hydration.
of the most controversial cases in past disputes. Second, and perhaps more importantly, Owen et al’s interpretation of their findings is itself highly controversial. fMRI studies of learning under anaesthetic and the phenomenon of priming show that highly complex cognitive processing can occur even at a non-conscious level. Some have accordingly argued that the results of Owen study similarly provide only evidence of brain processing occurring at a subconcious level. This issue remains very much in dispute but it may be resolved in the near future by further empirical research and conceptual discussion. For example, Owen and colleagues are proposing to repeat their experiments to test whether VS patients can use mental imagery to reply to a series of questions with verifiable answers. A positive performance on such a test by VS patients would presumably offer very strong evidence for consciousness although failure would not, of course, establish its absence. Further empirical data is forthcoming, then, though we still lack a satisfying account of what would count as adequate criteria for the presence of consciousness. We shall now argue that there is something else we still lack: an account of why and how consciousness matters.

3. The Moral Significance of Consciousness

Resistance to withdrawal of such life-preserving treatment is sometimes defended by appeal to the supposed Sanctity of Life. The sanctity of life might require preserving such patients’ lives even if it is not in their interest, indeed even if they can no longer be said to have interests. However, to a very large extent disputes about VS have tended to revolve, not around the question whether these and similar patients are still alive in some biological sense—something no one seriously denies, given that VS patients are not fully brain dead and their brain stem continues to maintain vegetative functions—but around the question whether these and similar patients could still be said to be conscious or have irreversibly lost the capacity for consciousness. It is only to this question that Owen et al’s findings might contribute something new, and we shall therefore focus here on the moral significance of consciousness and its relevance for questions about the treatment of severely brain damaged patients.

Far more than in the debate about abortion, many people accept that the moral questions about the permissibility of withdrawing life-preserving treatment from such patients would be settled if we knew whether they are still conscious. Indeed, one of the justifications for accepting a brainstem definition of death, such as that employed in the UK, rather than a whole brain definition of death, such as employed in the US, is that death of the brainstem necessarily involves loss of the potential for consciousness through the destruction of the reticular activating system and other pathways. Because of the importance attributed to consciousness, it has sometimes been denied, often against the plain scientific evidence, that some patients diagnosed as being in VS really are in that state. Those who oppose the termination of the lives of these patients have claimed either that these patients have in fact exhibited signs of consciousness, or that we do not know with sufficient certainty that they are not conscious. These disputes about consciousness are clearly not disputes about the sanctity of human life.

It is accepted by many, including both sides in these disputes, that the presence of consciousness, or of a capacity for consciousness (for simplicity’s sake we shall not distinguish the two in what follows), marks a crucial moral boundary separating conscious

---

9 As they write, in many cases “standard clinical techniques, including structural MR imaging, may be sufficient to rule out any potential for normal activation, without the need for fMR imaging.” (p. 1101).
10 See e.g. Deeprose 2005.
11 See e.g. Greenberg 2007.
12 Personal communication.
13 See the Vatican statement on the VS (Congregation For The Doctrine Of The Faith, August 2007).
beings from other entities. We shall call this principle *The Moral Significance of
Consciousness (SC)*.

Given that there is general agreement on SC, it may seem that no new ethical issues are
raised by the possibility that some patients diagnosed as being in VS may in fact still be
conscious. This, it might seem, is only an empirical question, if a very difficult one. Once
appropriate scientific criteria are developed for identifying consciousness on the basis of
brain activity alone, and we can tell whether a given patient is conscious or not, then this
should by itself settle the fundamental ethical issues. This impression is mistaken. The
difficulty *would* have been merely empirical if it were clear what it is doctors, ethicists and
others who make value judgements are referring to when they claim that consciousness is
morally significant. But this is not at all clear. A quick perusal of the voluminous literature
on consciousness reveals that there are very many competing accounts of what might be
meant by consciousness. To mention just some of these, philosophers of mind distinguish
and debate the relation between self-consciousness, transitive consciousness and intransitive
(or creature) consciousness, as well as monitoring, phenomenal and access consciousness.
As Ned Block remarks in a recent review, “[t]he concept of consciousness is a hybrid or
better, a mongrel concept: the word ‘consciousness’ connotes a number of different concepts
and denotes a number of different phenomena.”14 This diversity is not just a problem for
philosophers of mind. It is a problem for ethics and medical practice because this means that
we cannot tell whether SC applies to a given case without first identifying the morally
relevant concept of consciousness. Even if Owen *et al* are correct to claim that their brain
damaged patient is conscious in one sense this finding might have no moral difference if she
isn’t conscious in the relevant sense. Indeed we cannot just assume that the consensus view
about consciousness in neuroscience or philosophy of mind will converge on the morally
relevant concept. It might be that the sense of consciousness most fruitful for scientific
inquiry happens to be distinct from the one that makes a moral difference.

It is beyond the scope of this paper to offer an exhaustive taxonomy of concepts of
consciousness let alone to undertake the important task, lying in the borderline area between
ethics and philosophy of mind, of relating these to the principle of SC. What we shall do
instead is assume, with many others, that the central concept of consciousness is that of
phenomenal consciousness, and explore the ways in which this assumption might help to
explain SC (in what follows, we’ll therefore use ‘phenomenal consciousness’ and
‘consciousness’ interchangeably). Once we have clarified SC we shall, in the next section,
turn to consider the ethical significance of neuroimaging studies of brain damaged patients.

The notion of phenomenal consciousness is notoriously difficult, perhaps impossible, to
define. We might say that a state is phenomenally conscious if it has a phenomenal
character, if there is something ‘it is like’ to be in that state.15 Paradigm states with
phenomenal character are sensations: the felt pain of a muscle cramp or the way it feels like to see the redness of an apple. Phenomenal consciousness shouldn’t be confused with
self-consciousness, that is, possession of the concept of self and the ability to use this
concept in thinking about oneself. Self-consciousness is a sophisticated cognitive capacity
possessed only by most adult humans and perhaps by some of the higher primates, whereas
the capacity for phenomenal consciousness is arguably something humans share with many
lower animals. It is plausible that possession of self-consciousness makes a moral
difference, but it isn’t plausible to think that it is the concept of consciousness assumed by
SC.16 In disputes about brain-damaged patients like Terry Schiavo, what is at stake is not

14Block 2002.
15Nagel 1974.
whether she still had a sophisticated understanding of herself as a person with a past and future, but whether there was anything it was like to be her or whether ‘all is dark within’.

The distinction between phenomenal consciousness and self-consciousness draws attention to an important general point: not only can an animal possess phenomenal consciousness without having any kind of self-consciousness, it might possess phenomenal consciousness while having no, or only minimal, cognitive and motivational states and capacities. We shall use ‘sapience’ to refer to possession of such states and capacities. Sapience in our sense involves a rich mentality that is likely to imply consciousness in another, non-phenomenal sense—consciousness understood as global access to information, or, in Block’s terms, access consciousness. An animal may possess phenomenal consciousness despite having little or no sapience. The relation in the other direction is more controversial. Some philosophers hold that full-blown sapience can exist even when phenomenality is totally absent—as in the notorious thought experiments about zombies. As we shall see below, this seemingly far-fetched possibility is not entirely irrelevant to the interpretation of Owen et al’s fMRI findings.

Once we see that phenomenal consciousness and sapience can come apart, it might seem mysterious why it is so commonly thought that consciousness morally matters. Arguably, the moral significance of consciousness is grounded in the moral significance of interests. It is a truism that interests matter morally. The claim then is that consciousness explains why certain things have interests whereas others, even if valuable (like plants or works of art), do not.

To understand this claim, we need to clarify the relation between consciousness and interests. This relation might depend on what interests are, itself a difficult philosophical question. According to hedonistic theories, interests consist of having pleasant experiences and being happy. According to desire satisfaction theories, what matters is having our desires satisfied. According to objective good theories, certain activities are intrinsically good—for example, developing deep personal relationships and talents, gaining knowledge, and so on. In this paper we needn’t decide between these three theories. And it is plausible that all three capture important aspects of well-being. This view is implicit in Plato’s remarks on well-being in the Philebus, where he suggests that the good life for a man could neither consist only in pleasure nor only in wisdom. Socrates first rejects a life of pure pleasure:

…but without true judgement you could not judge at the moment of enjoyment that you are experiencing it, while without the power of calculation you could not even calculate that you will get enjoyment in the future; your life would be that not of a man, but of a sea-lung or one of those marine creatures whose bodies are confined by a shell.

---

16 Hassun and Kreigel 2007 argue that personhood in the morally significant sense requires consciousness and that consciousness requires having a self concept—essentially, self-consciousness or something close. Consciousness in the morally relevant sense we are discussing needn’t imply personhood, but in any case we find Hassun and Kreigel’s derivation of self-consciousness from consciousness very suspect. Lower animals can feel pain—an intrinsically bad aversive conscious state—without possessing anything like a self concept, indeed plausibly without grasp of any concept.

17 This notion of consciousness is also close to Bernard Baars 1997 theory that conscious representations are ones that are broadcast in a ‘global workspace’ and Daniel Dennett’s 1991 suggestion that consciousness is ‘cerebral celebrity’. In order not to generate unnecessary confusion, we’ll use ‘consciousness’ to refer only to phenomenal consciousness.

18 We’ll understand ‘interests’ to be roughly synonymous with what affect a being’s good, welfare or well-being.

19 On some views, all are aspects of well-being. But even those who endorse only one of these theories will typically accept the relevance to well-being of the elements cited by the competing theories. Thus hedonic states would matter on most objective theories, and desire-satisfaction theories typically assume that the value of hedonic states can be reduced to that of the satisfaction of desires to prolong or end sensations. Many of the things objective theories claim to be good are also things that people actually desire. And conversely objective good theorists needn’t deny that the satisfaction of at least some desires is part of an agent’s good, whether directly or through the hedonic satisfaction it can generate.

J Med Philos. Author manuscript; available in PMC 2011 December 18.
and then a life of pure wisdom:

The question is now whether any of us would be content with a life in which he possessed thought, intelligence, knowledge and perfect memory for everything, but had no sense of pleasure, great or small, nor of pain either, a state of perfect intelligence to all such feelings.\textsuperscript{21}

The good life, Plato concludes, consists in an appropriate ‘mix’ of pleasure and wisdom.\textsuperscript{22}

We’ll therefore just keep all three theories on board and draw their implications to our questions as we go along. We’ll accordingly distinguish between an entity’s experiential, desiderative and objective interests.

So let us consider what moral significance consciousness would have in light of each of these theories. Consider first objective theories. The typical goods listed by objective theorists—for example, deep personal relationships, knowledge, achievement, development of talents—clearly require a high degree of sapience but they do not obviously require phenomenal consciousness.\textsuperscript{23} The same is true of desiderative interests. Although some desires have a phenomenological aspect—think of how it feels to crave something—it needn’t feel like anything to want something. And on desire-satisfaction theories, our interests are promoted not when we feel psychological satisfaction but when our desires are satisfied in a logical sense—when we desire that \( p \) and \( p \) is actually the case.

It might seem that at least pleasure and pain are good or bad only because of their phenomenal character—because of what it intrinsically \textit{feels like} to have hedonic experiences. Even this, however, might be too quick. ‘Pain’ can refer to a certain kind of bodily sensation, or to the felt state of disliking a sensation. But ‘pain’ can also refer to bodily damage, or to the representation of bodily damage in the brain or to strong motivation elicited by such a representation.\textsuperscript{24} These latter states can take place without phenomenal consciousness. Would such states still count as suffering and call for relief? Many will plausibly deny that this makes sense, and we agree. We shall assume that pain in the sense that matters requires phenomenality—that for a representation of bodily damage or a motivation to get rid of it to morally matter, there must be something it’s like to be in that state.\textsuperscript{25} But even this claim is by no means universally accepted.\textsuperscript{26} In any case, unless hedonism is correct, the conceptual tie between phenomenal consciousness and hedonic states only establishes that consciousness is a necessary condition for experiential interests. It leaves it open that beings that do not enjoy phenomenality might still possess non-experiential interests.

So far we’ve tried to explore the ways in which particular types of interests might presuppose consciousness. This did not take us very far. But perhaps there is a more direct connection between interests and consciousness. Here is one suggestion. When interests are promoted or set back, these are not just ways in which things in the world might go impersonally better or worse. It’s someone’s \textit{good} that is being promoted or set back. It may be impersonally good to prevent a great work of art from being needlessly destroyed, but we promote interests for someone’s sake.\textsuperscript{27} What is the difference between these two forms of

\begin{footnotesize}
\begin{enumerate}
\item Plato, \textit{Philebus} 21 C 1-12.
\item 2121 D 11-3, E 1-3.
\item 2261 C 5-10.
\item The value of aesthetic appreciation might be an exception.
\item See the papers in Aydede 2006.
\item One of the authors defends this claim in XX.
\item For a general denial of the moral significance of phenomenal consciousness, see Carruthers 1999; for denial of its relevance to the badness of pain, see Clark 2006.
\item See Kamm, 2006, chapter 7.
\end{enumerate}
\end{footnotesize}
value? How is it that certain states of affairs matter, not impersonally, but in relation to someone? The difference seems to be that a painting is just an object but a squirrel or infant or adult human all have a points of view, a subjective take on things. In this way, possession of consciousness—of a subjective standpoint—might be a general condition for an entity’s having interests. It’s hard to give a substantive argument to support this suggestion, but think of how awry it seems to say ‘He led a good life—but there was absolutely nothing it was like to live that life’ (a thought echoed, perhaps, in Plato’s rejection of pleasureless wisdom).

We’ve tried to sketch an account of SC and what might ground it. Notice, however, that we have only considered the claim that consciousness is a necessary condition for the existence of interests. It is a further claim that it is a sufficient condition. After all, as we have pointed out, consciousness might still be present when cognition and motivation are entirely absent or present only in minimal form. Indeed, it is doubtful that a mental life consisting only of a bare stream of consciousness—a sequence of random and hedonically-neutral sensations—could be said to involve interests of any kind. A being can’t have desires, and thus desiderative interests, without a sufficient degree of cognitive capacity. Nor can one possess objective interests such as the interest in friendship or knowledge in the absence of such capacities (indeed many objective goods seem to require self-consciousness, not phenomenal consciousness). What about experiential interests? One cannot enjoy or suffer without being phenomenally conscious, but it’s far from obvious that mere possession of phenomenal consciousness implies that one has the capacity to experience pain or pleasure. A being that lacked both cognitive capacities and the capacity to feel pleasure and pain might be a being without interests despite possessing phenomenal consciousness. On the other hand, the capacity to experience hedonic states seems to require little or no sapience. Imagine that bits of brain are artificially assembled to form only that part of the brain needed to sustain physical pain, and that a single moment of pure agony is then generated in this brain part before it is again disassembled. In this example, we cannot really talk of a person or even animal that is suffering the pain. Yet it still seems that this moment of pain is bad—not impersonally bad but bad for someone, even if that someone is only a transient and rudimentary subject of consciousness.

There is another important point to make about the scope of SC. Even if phenomenal consciousness is a necessary condition for interests, we should not simply assume that the fact that an entity has irretrievably lost the capacity for phenomenal consciousness immediately implies that it no longer has any interests. This implication does go through for experiential interests, but it is not as obvious with respect to desiderative and objective interests.

In order to have a point of view on things, it might be enough if one were conscious at some point in time. But once this condition is met, it might be that some interests extend beyond the extinction of consciousness. Take desiderative interests. A person’s desire that p could be satisfied even when he no longer exists, pari passu if he exists without consciousness. And some objective interests may have an equally tenuous tie to the presence of consciousness. Now we do not want to enter here into the dispute over whether people can benefit or be harmed even after they have died. It suffices to note that there are those who hold this view. Presumably, the case for interests extending only beyond the loss of the capacity for consciousness would be stronger. We do not want to endorse this claim, merely to note that acceptance of SC needn’t directly imply that beings no longer have (non-experiential) interests when consciousness is irreparably extinguished.
4. Consciousness, Sapience and the Value of Life

We’ve accepted that consciousness makes a moral difference, and that it makes this difference through its tie to interests, but we’ve deliberately left it open what difference it makes. It is often assumed that the difference it makes is that we have strong moral reasons to sustain the life of conscious patients, reasons we would not have if they had permanently lost the capacity for consciousness. In this section we will argue that, stated thus, this assumption is far too simple.

Let us return to severely brain damaged patients and to what neuroimaging might tell us about their mental lives. A brain damaged patient might be phenomenally conscious or she might not be. And such a patient might enjoy significant sapience or she might not. This gives us four possibilities, two where the patient isn’t conscious and two where she is.

Under which of these does Owen et al.’s patient, and those like her, fall? That’s a difficult empirical question which we cannot answer but which may be answered by further empirical research. What we want to do in this section is to consider the ethical implications of each of the four possible answers.

I. Neither consciousness nor sapience

This is the condition VS patients are assumed to be in: a condition where the brain still maintains the body’s vegetative functions but a mental life is totally absent. There is nothing it is like to be in this condition, nor is there any real degree of what we called sapience; someone in this condition can’t be said to believe or want anything. Note, however, that being even in this condition is compatible with the occurrence of fairly sophisticated information-processing in some areas of the brain—such as the processing of visual or semantic information observed in some neuroimaging studies of VS patients. Importantly, such mental activity would not only be non-conscious—that is, taking place in the absence of phenomenal consciousness—but also sub-personal—too localized and disjointed to merit ascription of mental states to the person as opposed to ascription of information-processing to localized sub-systems in her brain.

The mere occurrence of complex information processing does not endow a being with any kind of moral status—otherwise it would be morally wrong to switch off a computer. Might there still be some value in the preservation of such sub-personal cognitive and affective processes, not in their own right, but because they are the remnants of the mental life of a person? (Think of such sub-personal responses to hearing one’s name or the name of one’s beloved.) To us it seems highly doubtful that there is any moral reason to sustain a life just in order to allow such sub-personal processes to continue to occur.

Since a person in this condition has no phenomenal consciousness, she can feel no pain or pleasure—she has no experiential interests. Notice that some of the sub-personal correlates of pain and pleasure might still occur in her brain in response to stimuli but, as we’ve remarked earlier, it’s hard to think of unfelt pain as amounting to any kind of genuine suffering—to anything bad for the person. And, in the absence of sapience, she can have no desires. So she also lacks any present desiderative interests.

On some views of well-being, however, she might still have some desiderative and objective interests grounded in her past conscious and sapient life. In particular, there is the possibility of desiderative and objective interests that have to do with her continuing to exist in such a vegetative state. However, if there are such interests, they are likely to point in the direction of discontinuing life-sustaining treatment. Take first desiderative interests. Several surveys reveal that a vast majority of people would prefer not to be given life-sustaining treatment if
they were in a non-reversible vegetative state. And these preferences might themselves reflect recognition of an *objective interest* in not continuing to exist in a state that has no personal meaning and that could be seen as degrading to one’s dignity as a rational being. Furthermore, to the extent that persons in such a condition might still be said to have interests, these interests (both desiderative and objective) are likely to include the interests of close others, interests that in many (though not all) cases will support ending the patient’s life.

We need not take a stand on whether interests can persist beyond the extinction of consciousness (or death); in their absence there may not be positive reasons, grounded in the patient’s interests, to end her life, but nor would there be such reasons to sustain it. And even those who baulk at thinking of persons’ interests extending beyond the loss of consciousness might accept that these past desires and objective considerations generate moral reasons to end a person’s life, reasons we have out of respect for that person.

II. Significant sapience without consciousness

If VS is understood to imply the absence of *phenomenal* consciousness, then the possibility that significant sapience will persist in the absence of consciousness is, strictly speaking, compatible with VS. Nevertheless, so far as we know it has not yet been discussed. This is not entirely surprising: the possibility that sapience might come apart from phenomenal consciousness is controversial, and there are difficult questions about what would justify ascribing sapience without consciousness, indeed, even about the very coherence of this possibility. Still, it is a possibility that cannot be just ruled out, and neuroscience has already provided us with many examples of aspects of the mental which, despite seeming inextricably intertwined in normal subjects, can nevertheless come apart in certain forms of brain damage.

Let us try to be clear about what significant sapience without consciousness would amount to. It would involve the presence of cognitive and motivational processes that are sufficiently extensive and systematic to merit not just ascription of local information processing in some area of the brain but ascription of genuine person-level mental states such as beliefs and desires. But this mental activity would take place without phenomenality—strictly speaking, there would be nothing it is like to be such a person. Now this possibility is obviously a conceptual relative of the possibility of zombies, but it is not as far fetched as that philosophers’ thought experiment. A zombie is supposed to be physically identical to a person yet lack any phenomenal consciousness. The possibility we are considering requires no such thing. It only requires that it be possible that the neural machinery required for phenomenal consciousness would be damaged while the brain areas responsible for cognition and motivation remain largely intact. This possibility may ultimately prove to be incoherent but at this stage of knowledge we certainly can’t just rule it out. Indeed notice that when Owen *et al* write of their patients that “her decision to cooperate with us by imagining particular tasks when asked to do so represented a clear act of intention that confirmed beyond any doubt that she was consciously aware of herself and her surroundings,” the evidence they are citing is really evidence for sapience (indeed, for access consciousness), and only indirectly, if at all, evidence for phenomenal consciousness.

---

28Frankl, Oye, and Bellamy 1989 studied the preferences of 200 adult inpatients regarding life-sustaining treatment in the context of four outcome scenarios. They reported that only 6 percent of the studied population would want life-sustaining treatment if they were in a vegetative state presumed to be permanent. Emanuel *et al* 1991 surveyed 405 outpatients and 102 members of the general public in Boston, asking them about treatment preferences in four hypothetical scenarios. About 80 percent of the studied population indicated that they would not want life-sustaining treatment if they were in a VS presumed to be permanent.
What would be the moral status of a patient in such a state? This is a difficult and largely unexplored question in ethics, considered until now largely in discussion of artificial intelligence. Some views of moral status or interests do explicitly require the capacity for consciousness but do not clearly identify the relevant notion of consciousness as that of phenomenal consciousness. Others ground them only in possession of certain cognitive or desiderative states, and can therefore be interpreted as accounts of moral status or interests that require only possession of sapience. However, given that the authors defending these views do not explicitly consider the possibility of sapience without phenomenal consciousness, we doubt that such a reading is warranted.29

We only have space to make some tentative suggestions. Given the absence of phenomenal consciousness, we believe that such a person would not have experiential interests—she might have states that are functionally very similar to pleasure and pain, but there would be nothing it feels like to be in those states. We do not think that such states would count as intrinsically good or bad but, as noted earlier, not everyone agrees.30 A patient in such a state would have genuine desires and preferences. Would these generate desiderative interests? The answer to this question depends on whether, as we’ve suggested, phenomenal consciousness is required if a person is to have a point of view, that is for the satisfaction of some desire to be a benefit for someone. It might even be thought that the desires and preferences of someone in this condition should be taken into account out of respect for that person, whether or not they generate genuine interests—indeed it might be that phenomenal consciousness is not required for personhood and the moral standing it implies. These are all questions that require further exploration.31

What, finally, about desiderative interests that might be generated by past desires? To a large extent the situation here is no different than that presented by the absence of both consciousness and sapience. There is however a minor complication. In the surveys we cited earlier, people were envisaging being in a vegetative state with neither consciousness and sapience. Would they have different wishes if they expected sapience to be preserved? We think that this is unlikely, but this is a matter for empirical inquiry.

III. Consciousness with minimal or no sapience

Clinically, this is the most probable state that a person emerging from VS of long duration would enter. It’s what is known as the minimally conscious state (MCS), a state of “profound neurological dysfunction but in which awareness is present, at least to some extent, and at some times.”32 The only difference from existing cases diagnosed as MCS is that in existing cases there is behavioural evidence for the presence of consciousness, and here (by assumption) only brain imaging evidence.33

We assume that patients in MCS have phenomenal consciousness but possess only the most minimal sapience.34 Actually, to refer to such patients as ‘minimally conscious’ is highly misleading. These patients do not necessarily have less phenomenal consciousness than we...
do. Indeed it is far from clear that we can coherently speak of phenomenal consciousness as a matter of degree. What these patients have only to a minimal degree is sapience.\(^{35}\)

What is the moral status of such patients? In most respects, it is essentially the same as that of patients in the latter stages of dementia, the main difference being that most demented patients have lived a full human life and arrived at their present stage fairly gradually, whereas a patient can enter the MCS abruptly and at a young age.\(^{36}\)

Such patients have, at best, only the most rudimentary desires, and they clearly do not have enough sapience to enjoy most objective goods—the goods of friendship, knowledge, achievement and the like. Do they at least have experiential interests? As we noted earlier, although there are no experiential interests \textit{without} phenomenal consciousness, the mere presence of phenomenal consciousness doesn’t by itself imply a capacity to feel pleasure and pain. It is however plausible to assume that such patients do feel pleasure and pain—that they do have experiential interests.

These experiential interests generate moral reasons to alleviate pain as well as to make the lives of MCS patients as comfortable and pleasant as possible, so long as they are alive. It is a separate question, however, whether there are any good moral reasons to \textit{sustain} the life of such patients. Consider this question first only in light of the patient’s experiential interests. If such a patient is experiencing great agony which we cannot relieve, it seems that continuing to exist is a source of harm and of no benefit. What, however, if we were confident enough that the patient does not suffer significant pain but instead experiences highly pleasant states of consciousness? Would this, in itself, be a reason to keep this patient alive for as long as we can? This is not an easy question. Such a patient clearly benefits from going on living in the sense that this means that, over time, she will enjoy more experiential goods. But this patient does not possess self-consciousness or a desire to go on living, and little to no psychological connectedness over time. It is thus not clear that she would be significantly harmed if her life ended earlier than was possible.\(^{37}\) And, given that, considerations of distributive justice may tell against continuing to sustain the life of such a patient at great cost.

Things get more complicated when we also consider the person’s desiderative and objective interests. Again, her desiderative interests (assuming she has any) would not be generated by any present desires but by past ones. Would people want to go on living in such a state? There is no specific survey data available that we are aware of, though many people would prefer not to go on living in the similar state of extreme dementia, and complete advance directives or living wills to that effect.\(^{38}\) A number of medical experts have indeed expressed the view that it is preferable to be in the VS (i.e. to permanently lose the capacity for consciousness) rather than live on in the MCS. For example, Nelson and Cranford claim that “being kept alive in the minimally conscious state may be far worse for the individual than being maintained in the vegetative state”;\(^{39}\) and Ashwal and Cranford write that

---

\(^{35}\) Sapience clearly is a matter of degree, in both dispositional and occurrent senses. First, an organism can possess cognitive and conceptual capacities to various degrees and, consequently, enjoy different levels of awareness of self and environment. Second, an organism ability to actually exercise these capacities can vary over time: the sapience of an extremely tired person at the moment before falling asleep is impaired in obvious ways. Sapience in the MCS seems to be severely impaired in the second and, probably, also the first sense.

\(^{36}\) We are assuming that our imagined MCS patient has little or no prospects of recovery. Neuroscientists have recently used deep brain stimulation to induce modest improvement in an MCS patient (Schiff 2007), but this line of research is still in its infancy.

\(^{37}\) For a related point see Wilkinson 2005.

\(^{38}\) Schiff 2000.

\(^{39}\) Nelson and Cranford 1999, p. 448.
if there were a better understanding of MCS, especially the critical issues of consciousness and likelihood for pain and suffering, a broader consensus would develop, that being in a permanent MCS would actually be worse than being in a permanent VS. In other words, just as VS is considered to be ‘a fate worse than death,’ being in a permanent MCS is a fate worse than VS.40

These remarks can be understood simply as a preference—as a desiderative interest that may or may not be shared with others. It is an empirical question how widespread this preference is. Understood as a response to an objective interest, the idea seems to be that remaining conscious, and thus in one sense continuing to exist as a psychological entity, yet losing all of the aspects of sapience that constituted you as a person with particular projects and attachments, is more meaningless and degrading than to lose consciousness completely and thus stop existing as a psychological entity—recall Plato’s dismissive remarks about the lives of ‘marine creatures’. However, if this preference is really grounded in the fear that one’s experiential interests would be compromised—that one would go on suffering pain—then this preference would not be reasonable if we were able to reliably detect and alleviate pain in MCS patients. As noted above, it might even be possible to promote the experiential interests of such patients by inducing pleasant sensations.

Indeed, to the extent that there are genuine desiderative and objective interests in not continuing to exist in the MCS, these might conflict with the experiential benefits to the patient from continued existence, to the extent that we can ensure that she suffers little pain and that she enjoys sensory pleasures. However as we noted above, it is not clear that such possible benefits in themselves generate a strong interest in continuing to exist, and given the absence of any psychological continuity or even connectedness with her past, sapient self, these many not, in any case, be benefits that genuinely accrue to her, a temporally-extended psychological entity that may no longer exist.41

In the final month of her life, the parents of Terry Schiavo insisted that she was in MCS, not VS. But we’ve argued that the discovery of consciousness in patients diagnosed as in VS hardly settles the ethical questions on the side of continuing life-sustaining treatment. If the patient is in the MCS, it might rather be that we have no or only weak reasons to sustain her life, and some further positive reasons not to sustain it.

IV. Consciousness accompanied by normal or significant sapience

A patient in this state, externally indistinguishable from a VS patient, would be in an extreme and absolute form of the Locked-In State.42 Such a person recovers both consciousness and sapience. His rationality and memory are essentially preserved. He wakes up to find that, while perhaps still capable of a measure of perceptual input, he is utterly incapable of any kind of motor output. His ability to act in the external world, let alone communicate with others, has been lost. And unlike other locked-in patients, he is not even capable of eye movement and blinking which offer at least a minimal form of contact with (and influence on) the external world.

Like the patient in the MCS, a locked-in patient enjoys phenomenal consciousness and thus has experiential interests. He can feel great pain and, potentially, can enjoy pleasure. But unlike the MCS patient he has retained his intellectual faculties and the core of his self. And it might be thought that we are at least morally required to do our best to preserve the life of

---

40 Ashwal and Cranford 2002, p. 29.
41 See McMahan 2002 on ‘time-relative interests’.
42 Worryingly, the current authoritative medical definition of VS (Multi-Society Task Force on PVS 1994) defines this state in purely behavioural term. This has the absurd implication that even if we had conclusive evidence that a patient is totally locked-in, she might still count as VS on this definition—unless we understand ‘behaviour’ to include purely mental acts.
a patient in this state. This, however, is also far from clear. The issue revolves around whether such a life is really worth living.

It is arguable that such a life is even less worth living than that in the MCS. A patient’s ongoing interests in the MCS are essentially experiential. We can address these by trying to minimize physical pain and, perhaps, by trying to induce pleasure. Considered in itself there is nothing especially bad about such an existence even if, controversially, it could be said to go against the desiderative or objective interests of the temporally-extended person of which the MCS patient is a continuant. Things are very different in the case of the totally locked-in patient. Being aware of his condition and with little hope of communication with others, the patient cannot pursue most of his desires and personal projects and is cut off from most of the objective goods that make for a meaningful human life. Not only is he incapable of virtually all forms of agency, or of any meaningful social relations with others, he is also painfully aware that this is the case, meaning that his objective, desiderative and experiential interests are all frustrated.

Is this a life worth living? A less than completely locked-in patient recently described his experience:

Words can’t describe the situation I have been left in… but this is as close as I can get: an extremely horrific experience that I wouldn’t wish on my worst enemy. The incredibly immense frustration levels at times have eased slightly over the years because of physical and health gains I have made. If dying is as painless and peaceful as just drifting off to sleep, then there’s plenty of really very frustrating times that at a particular point I wished I wasn’t here any more… An itch is completely unbearable and incredibly frustrating because I can’t scratch it… I thought of suicide often… I never had enough courage then to go through with it. Even if I wanted to do it now I couldn’t, it’s physically impossible. I just have to deal with it the best I can… I only live for hope of recovery now. I can’t live like this for about another 40 years or so and will not.43

Some patients in the locked-in state who do have the capacity to communicate clearly do not wish to continue their lives in this state.44 However the capacity of human beings to adapt to their condition, no matter how adverse, has been amply documented in other contexts, and one survey suggests that a majority of locked-in patients would want to continue their lives even in this condition.45 Indeed in one now famous example, Jean-Dominique Bauby, the French locked-in patient, wrote the book The Diving Bell and the Butterfly one letter at a time by indicating the desired letter to a therapist by blinking.46 We should therefore not simply assume that patients in the locked-in state are unable to pursue personal projects and attain great objective goods.

The surveyed patients, however, are capable of communication. Bauby would not have been able to write his book otherwise.47 The totally locked-in brain-damaged patients we are now considering have no capacity for communication, no external agency and at most only limited (and completely passive) perceptual input. Many have also been in such a state for a long time. Their situation is thus far worse than someone in the worst form of solitary

43Chisholm and Gillett 2005.
44One doctor reports “In my 25 years in medicine, I’ve probably communicated with four or five patients who were locked in, spending many hours over many days with them to determine their wishes. None of these patients wanted to stay alive in that condition once they comprehended that they were going to be locked in.” (Fine 2005)
45Laureys et al 2005.
46Bauby 1998.
47Would there be even a point to trying? There might be rare exceptions. In J. L. Borges’s fiction ‘The Secret Miracle’, a playwright facing a firing squad finds that time has miraculously stopped, allowing him to complete his life’s work in his head (Borges 2000)
confinement. Their lives have gone very badly since entering this state and if it continues unaltered, may go on being very bad. It is far from obvious that such lives are still worth living. If so, then even if using fMRI we can establish that brain damaged patients still enjoy phenomenal consciousness and a significant measure of sapience, terminating these patients’ lives might be morally required, not merely permissible.

The dilemma this possibility presents is uniquely difficult because each such patient presumably has a view on the matter but at present there is no way of finding out what it is or of asking for their consent to the withdrawal of life-sustaining treatment. One consequence that should not be controversial is that we have strong reason to focus research not only identifying neural criteria for the presence of consciousness and sapience but also on the development of means of communication with what might be totally locked-in patients. Of course to the extent that such research is successful, it may also dissolve the dilemma by removing one of the central reasons for thinking such lives might not be worth living—the absence of relations with others, a strong human interest on all plausible accounts of well-being. However such advances are still in the future and such patients may be suffering great mental anguish. It may be wrong to prolong their suffering needlessly when it is unclear when and whether such means of communication would ever become available.

Even (or especially) those who find this pessimistic conclusion hard to accept would agree with us that we should ascribe great moral urgency to the relief of suffering in patients that might have until now been diagnosed as in the VS. Both MCS and totally locked-in patients are capable of feeling pain in the sense that matters morally. Indeed, for locked-in patients who are conscious of their state, the experience of physical pain is likely to be all the more great given that they have no means of acting to reduce or end it and given that others do not acknowledge their suffering. There is thus a strong moral imperative to provide painkillers or even anaesthesia to brain damaged patients suspected of being in the MCS or locked-in state. And even if a conscious brain damaged patient does not feel physical pain, he may still experience great mental suffering if he has retained significant sapience. Thus, not only should we seek to develop tests which provide evidence a person is in pain, we should also develop tests which are capable of providing objective evidence of mental distress. If there is some probability that a brain damaged patient is in such a state, doctors should consider administering anti-depressants and even sedatives.

5. Conclusion

Some doctors defend the ethical principle that “human life is to be preserved as long as there is consciousness and cognitive function in contrast to vegetative state or a condition of neocortical death.” This principle assumes that consciousness and cognitive function always go together. We now know that they do not. Nor does their presence have the implications assumed by the defenders of this principle. In this paper, we have argued that the common assumption that consciousness is morally significant is fundamentally unclear. After developing one plausible way of interpreting this assumption, we have argued that brain-damaged patients may have non-experiential interests even if they have lost the capacity for consciousness, but that it is very likely that these interests will point against the preservation of their life. We have further argued that even the presence of consciousness does not by itself show that it is in the patient’s best interest to go on living. Indeed, it is

---

48 In some locked-in patients, the brain damage might have impaired the capacity to experience bodily sensation, and on some theories of affect this may lead to a blunting of affect, potentially reducing such patients’ capacity to feel distress. Though anaesthesia might compromise the interests of such patients, on some accounts.

50 Stumpf 1986.
arguable that existence with consciousness but with minimal sapience is worse than non-conscious vegetative existence, and that such a state might in turn nevertheless be better than existence with consciousness and preserved sapience if one is totally locked-in.

As science advances and opens up new knowledge or possibilities, scientists and clinicians may struggle to apply concepts that are too vague or inclusive. Most countries in the world have followed the Harvard Ad Hoc Committee\(^{51}\) and have revised their definition of what constitutes death from one of whole organism death to brain death. Advances in the neuroimaging of severely brain damaged patients may put similar pressure on the concept of consciousness as used in law, clinical practice and normative ethics.\(^{52}\) It might be time to refine or revise the concept of consciousness if we are to adequately ascribe it and evaluate its ethical significance.

REFERENCES


Chisholm, Nick; Gillett, Grant. The patient’s journey: Living with locked-in syndrome. British Medical Journal. 2005; 331(9)


Cranford, Ronald; Facts, Lies; Videotapes. The Permanent Vegetative State and the Sad Case of Terri Schiavo. The Journal of Law, Medicine & Ethics. 2005

---

51Ad Hoc Committee of the Harvard Medical School to Examine the Definition of Brain Death, 1968.
52Eisenberg 2008 gives a useful survey of the bewildering range of different definitions of awareness in US state statutes concerning permanent VS. Worryingly, many are purely behavioural. None is sensitive to the conceptual and normative distinctions drawn in this paper.


Fine, Robert L. From Quinlan to Schiavo: medical, ethical, and legal issues in severe brain injury. Baylor University Medical Center Proceedings. 2005; 18(4)


Greenberg DL. Comment on “Detecting Awareness in the Vegetative State”. Science. 2007; 315(1221)


Laureys, Steven; Pellias, Frederic; Eeckhout, Philippe Van; Ghorbel, Sofiane; Schnakers, Caroline; Perrin, Fabien; Berre, Jacques; Faymonville, Marie-Elisabeth; Pankte, Karl-Heinz; Damas, Francois; Lamy, Maurice; Moonen, Gustave; Goldman, Serge. The locked-in syndrome: what is it like to be conscious but paralyzed and voiceless? In: Laureys, S., editor. Progress in Brain Research. Vol. 150. 2005.


Owen, Adrian M.; Coleman, Martin R.; Boly, Melanie; Davis, Matthew H.; Laureys, Steven; Pickard, John D. Detecting Awareness in the Vegetative State. Science. 2006; 313
Owen, Adrian M.; Coleman, Martin R.; Boly, Melanie; Davis, Matthew H.; Laureys, Steven; Pickard, John D. Using Functional Magnetic Resonance Imaging to Detect Covert Awareness in the Vegetative State. Archives of Neurology. 2007; 64(8):1098–1102. [PubMed: 17698699]