

Intentional action and the post-coma patient

Zoe Drayson

The presence or absence of intentional action plays an important role in the clinical diagnosis of patients emerging from coma states. Judgments about a patient's capacity for intentional action have traditionally focused on the nature of their bodily movements, but recent neuroimaging data suggests that brain activity can be indicative of intentional action. I will suggest that this change of focus, from the interpretation of motor behaviour as intentional (bodily) action to the interpretation of neural activity as intentional (mental) action, raises philosophical issues that have not yet been addressed in the literature. In particular, I am concerned that the notion of 'intentional mental action' is not as straightforward as the notion of 'intentional bodily action', which leaves the neuroimaging experiments open to different interpretations. Furthermore, I suggest that the kind of neuroimaging data that would settle the question of interpretation will be difficult to obtain, partly due to the nature of intentional mental action.

Claims about intentional action in post-coma patients have important implications. A patient's capacity for intentional action is used diagnostically as an indication that the patient has some form of awareness of their surroundings; in other words, they are *conscious* of the world around them. Attributions of conscious awareness are not to be taken lightly: they bring with them legal and ethical ramifications that affect patients, their families, and clinicians.

In the first part of the paper I introduce the clinical diagnoses of coma and post-coma states, and demonstrate their reliance on the notion of 'purposeful behaviour'. In the second part of the paper, I show how purposeful behaviour relates to the philosophical notion of intentional action, and how intentional action is used as an indicator of consciousness, both in the everyday case and the clinical setting. In section three, I explore the ways that neuroimaging technology can be used to argue for the presence of consciousness. I focus on the experimental paradigm which claims to show that patients diagnosed as vegetative are in fact capable of intentional action, where this is suggested not by overt bodily movements but by neural activity. The claim, issuing from the Owen et al. (2006) experiment, is that patients are performing intentional acts of imagination and therefore can be attributed conscious agency. In the fourth part of the paper, I raise questions about the interpretation of neural activity as intentional action, focusing on the notion of imagination as an intentional mental action. Drawing on the work of Mele (1997) and Peacocke (2007), I suggest that being in a mental state with a particular content can sometimes be an intentional action and sometimes not be an intentional action. Furthermore, I argue that the current neuroimaging data is neutral between

these two possibilities, and so does not establish that patients are performing an intentional act of imagining. As a result, Owen et al. (2006) are not in a position to attribute consciousness to patients until they have supplemented their data with an experiment that controls for this problem. But there seem to be both practical and philosophical issues that would make designing such an experiment extremely difficult. I conclude by suggesting that the neuroimaging experiments would benefit from further philosophical work on the nature of intentional mental action.

1. Medical diagnosis: coma and post-coma

After severe brain injury, a patient can be in a *coma* for several days or weeks. Those patients who recover may progress initially to a *vegetative state*, then to a *minimally conscious state*, before eventually regaining the ability to interact and communicate with those around them. Each of these clinical states has different diagnostic criteria.

In a *coma*, the patient is completely unresponsive and unarousable, as if in a deep sleep. They don't respond to any kind of stimuli (e.g. the sound of their name or the voice of a loved one), and don't show even reflex responses to painful stimuli, or loud noises. The patient may be unable to breathe normally, and their body may show stereotypical posturing. There is usually damage to the brainstem and/or cortex (Plum and Posner 1982, 5).

When a patient emerges from a coma into a *vegetative state*, they seem to 'wake up': their eyes open and close, and they display a sleep-wake cycle. Vegetative patients have at least partial preservation of hypothalamic and brain stem autonomic functions: they breathe spontaneously and have stable circulation. They may also show reflex responses like swallowing (Plum and Posner 1982, 6). However, it is essential to the diagnosis of vegetative state that the patient does not display any signs of 'purposeful behaviour'. This means that there is "(a) no evidence of non-reflexive response to sensory stimulation; (b) no evidence of awareness of the self or the environment; (c) no evidence of language comprehension or expression" (Royal College of Physicians 2003).

Patients who go on to develop reproducible signs of purposeful behaviour, however inconsistently, are diagnosed as being in a *minimally conscious state*. This might consist merely in being able to reach for objects or fixate on moving objects, as long as the behaviour is "reproducible or sustained long enough to be differentiated from reflexive behavior" (Giacino et al. 2002). Recovery from a

minimally conscious state is considered complete when the patient is able to functionally communicate, verbally or otherwise.

Notice the importance of so-called ‘purposeful behaviour’ for differential diagnosis between the vegetative state (in which it is absent) and the minimally conscious state (in which it is present). For the clinician confronted with a post-coma patient, their first diagnostic goal is “is to determine whether the patient retains the capacity for a purposeful response to stimulation, however inconsistent” (Monti et al. 2010). If this capacity is present, the patient will be considered minimally conscious. The diagnosis of vegetative state, on the other hand, “depends crucially on there being no reproducible evidence of purposeful behavior in response to external stimulation” (Owen et al. 2006). To understand why ‘purposeful behaviour’ is enshrined in the diagnostic criteria, one has only to look at the diagnostic labels: the capacity for purposeful behaviour is what enables the patient to be considered *conscious*, albeit minimally so. And attributions of consciousness bring with them a host of clinical and moral considerations: “Such a capacity [for purposeful behaviour], which suggests at least partial awareness, distinguishes minimally conscious patients from those in a vegetative state and therefore has implications for subsequent care and rehabilitation, as well as for legal and ethical decision making.” (Monti et al. 2010)

In the following section I’ll explore this diagnostic criterion of purposeful behaviour in terms of its relation to the philosophical notion of intentional action. I’ll also examine the link between attributions of purposeful behaviour and attributions of consciousness, and say more about how to understand the notion of consciousness involved.

2. Intentional action and consciousness

The notion of ‘purposeful behaviour’ found in the diagnostic criteria is what philosophers call ‘intentional action’: voluntary or controlled movements, which can be contrasted with involuntary or reflex movements. To perform an intentional action is to be *agent*, to possess *agency*. When I pick up the ringing telephone receiver, for example, or fill a glass with water, these are intentional actions directed towards a goal: I am answering the phone, I am fetching a drink. By contrast, when I develop goose-bumps on my skin in cold weather, or when my pupils dilate in dimmed light, these are not intentional actions: they are things that happen to me rather than things I do. There may be an evolutionary purpose to such reflexes, but they are not directed towards my own goals.

This raises the question of how we attribute the movements of my body to me-as-agent, rather than as something that happens to me. The generally accepted view is that we think of behaviour as the action of agent when it is best understood as the behaviour of the thinking, reasoning agent *as a whole* rather than merely as the behaviour of the person's component parts. This is the idea that there should be some sort of integration between our actions and our other mental states: "What it is for an action to be assigned to the agent herself rather than one of her components is for it to be suitably integrated into her cognitive economy" (Bayne 2013, ms. 5).

When I answer the ringing telephone, for example, I'm in a position to use the information that the phone is ringing in different ways. I can tell someone else the phone is ringing, I can make decisions about whether to answer it, I can wonder who is calling. By contrast, when my pupils dilate in response to the light, the information about the amount of light needn't be available to me: I can't report it, remember it, or make decisions on the basis of it. The information is only available to be used by my eyes, and is not integrated into my general cognitive operations. This idea of cognitive integration is closely related to what Block (2007) calls "cognitive accessibility" and Chalmers (1997) calls "availability for global control". (I am not proposing a set of necessary and sufficient conditions for determining when information is cognitively available, and there may well be difficult or borderline cases.)

How does intentional action relate to consciousness? If you see me answer the ringing phone, you can infer that I can hear it: my awareness of the sound is the best explanation of my intentional action of answering it. By contrast, we can't draw inferences like this from reflex behaviours: my skin can develop goose-bumps without me noticing that the temperature has dropped, and my pupils can dilate without me noticing that the light has dimmed. The claim is not that intentional action requires consciousness, but merely that "the exercise of intentional, goal-directed agency is a *reliable guide* to the presence of consciousness" (Bayne 2013, ms.1, my italics). Bayne (2013) characterizes this position as treating "agency as a marker of consciousness":

"Consider the infant who reaches for a toy, the lioness who tracks a gazelle running across the savannah, or the climber who searches for a handhold in the cliff. In each case, it is tempting to assume that the creature in question is aware—that is, consciously aware—of the perceptual features of their environment (the toy, the gazelle, the handhold) that guide their behaviour."
(Bayne 2013, ms.1)

When we say that agency is a marker for consciousness, what do we mean by "consciousness"? "Consciousness" is used in a variety of different ways, to pick out phenomena including wakefulness, attention, self-awareness, and subjectivity. But the notion of consciousness that relates to intentional action is the transitive sense: consciousness as consciousness *of* something: the

ringing telephone, the running gazelle. The transitive sense of consciousness might be glossed as responsiveness to a stimulus, or awareness of a stimulus. For the rest of this paper, I shall use the term ‘consciousness’ to refer to this sense of conscious awareness of aspects of one’s surroundings.

It is important to stress that the idea of agency as a marker for consciousness does not require a conceptual connection between consciousness and intentional agency, or that we’re infallible in our attributions of either. And it is certainly not to identify conscious awareness with intentional agency. It is simply to acknowledge that intentional agency is an indicator (albeit an indirect, fallible indicator) of consciousness. And of course, to say that intentional action is a marker for consciousness is not to say that intentional agency is the only way of making attributions of consciousness. Most of our everyday attributions of consciousness are based on what people *say* about their experiences of the world: “I can hear a car alarm”; “I can smell the sea”; “I can’t see the letters at the bottom of the eye-chart”. When we make such claims, we are giving *introspective reports* on our own conscious states. Using introspective verbal reports how they are aware of the environment around them, which allows us to reliably and accurately attribute consciousness to them. There are also non-verbal ways to report one’s awareness: one can point, for example, and one can nod or shake one’s head in answer to questions. Psychology experiments often ask for non-verbal responses such as hand-raising or button-pressing to indicate when a subject has consciousness of a certain stimulus.

Introspective report can be understood as a special kind of intentional action: one that acts as a ‘gold standard’ for assessing whether (and how) people are conscious of their surroundings. Some people would go so far as to say that introspective report is the only way to measure consciousness (see discussion in Bayne (2013) for examples) but it seems clear that in everyday life we do in fact use ‘mere’ intentional action (i.e. intentional action that doesn’t include introspective report) for attributing consciousness.

Now that we are clearer on the way our attributions of intentional agency and conscious awareness work in everyday life, let’s consider the clinical setting. How does the clinician determine whether any sort of patient is consciously aware of their surroundings? Introspective report is the first option: the doctor might ask the patient what they can see or hear and record their verbal report. Of course verbal report isn’t always an option: there might be damage to the larynx or to the tongue, or to speech-processing centres in the brain. Here, the clinician will ask the patient to report in a non-verbal way: by squeezing their hand to indicate that they can hear them, for example. But doctors also rely on intentional actions that are *not* introspective reports: a patient who turns their head at

the sound of their name, or who reaches to grasp the mug in front of them, will be judged to be *conscious of* the relevant stimulus in the same way that we normally attribute consciousness awareness on the basis of intentional action. Wherever the patient is displaying intentional agency, we are justified in taking this as an indicator of conscious awareness. This idea of agency as a marker of consciousness is built into the diagnostic guidelines: patients are judged to be in a minimally conscious state on the basis of their capacity for intentional action, not their capacity to give introspective reports.

The patient is judged to be in a vegetative state rather than a minimally conscious state when they do not display the capacity for intentional action. But the fact that they do not meet the diagnostic guidelines for consciousness does *not* mean we can rule out their being conscious. This would be a risky move, because there are clear clinical cases where intentional action is minimal or absent but the presence of conscious awareness can be established. There are two well-known clinical cases where such a scenario occurs: the phenomenon of awareness during surgery, and the condition of ‘locked-in syndrome’. In the former case, a patient is given a general anaesthetic and a muscle-paralysing agent prior to surgery. The first removes consciousness, and the second prevents bodily movement. If the anaesthetic wears off before the paralysing agent, the patient can be fully conscious but unable to communicate in any way. Afterwards, they are able to relate their experiences. In the case of locked-in syndrome, the patient has full consciousness but a complete or partial inability to produce motor behaviour – often, the only muscles they can control are those governing eye-movements. Unless someone notices that they are trying to communicate with their eye-movements, or specifically asks them to respond in that way, their behaviour may be interpreted as mere reflex rather than intentional action. Such cases as these tell us that it is possible for consciousness to exist in clinical states where there is little or no suggestion of intentional action. This should make us wary of thinking that the lack of intentional action in the vegetative state should be taken as evidence that the patient lacks consciousness: the absence of intentional action is by no means a guarantee that there is no consciousness. In the next section, however, I’ll explore how neuroimaging data has recently been used to supplement traditional methods of attributing consciousness to post-coma patients.

3. Neuroimaging and the vegetative state

Recall that the most reliable way to make attributions of consciousness is on the basis of the subject’s introspective report. In the absence of introspective report, however, it seems that attributions of consciousness can be made on the basis of the subject’s capacity for intentional

action. In the case of patients in a vegetative state, there are no signs of intentional action. So in order to attribute consciousness to such patients, we must either find a new way of attributing consciousness that doesn't rely on intentional action, or find a way to demonstrate that such patients are capable of intentional action after all. Neuroimaging data has been used to support both these projects. The first (or 'direct') approach uses neural data to argue that patients can be conscious in the absence of a capacity for intentional action, and thus relies on neural activity providing a new sort of indicator of consciousness. The second (or 'indirect') approach argues that neural activity, like bodily movement, can be evidence of intentional action and therefore an indicator of consciousness.

(Notice that the two approaches have slightly different consequences for the diagnostic criteria. If the first approach works, it will show that patients can be minimally conscious without meeting the 'purposeful behaviour' criterion for the minimally conscious state. If the second approach is successful, however, it would mean that the relevant patients would be re-diagnosed as minimally conscious rather than vegetative, in light of displaying 'purposeful behaviour' by way of their neural activity.)

I'm interested mostly in the second approach, but it's worth saying a little about the problems faced by the first approach. The first approach is direct: it relies on drawing inferences directly from neural activity to the existence of consciousness, without going via any considerations of intentional action. Using neuroimaging techniques like positron-emission tomography (PET) and functional magnetic resonance imaging (fMRI), data has been collected about the brain activity of subjects in certain situations. It has been demonstrated that patients in the vegetative state show different levels of activity when they are being read to, depending on whether the reader is a stranger or a family member (de Jong et al. 1997). Similarly, the brain of the vegetative patient often responds differently to photos of familiar faces and to photos of people they've never seen (Menon et al. 1998). But it would be unwise to conclude from such neural data that the patient is conscious of the stimuli: it is well known that perceptual processing can take place without consciousness, for example when a subject is asleep or under deep anaesthesia (Gross et al. 2004, Quandt et al. 2004). Further imaging studies, however, have shown that the vegetative brain shows a differential reaction to speech and to acoustically-matched but meaningless noise (Owen et al. 2005), and that parts of the brain associated with semantic processing are more active when presented with ambiguous words compared to non-ambiguous words (Rodd et al. 2005). This sort of semantic processing has often been assumed to require consciousness, but the evidence is less than conclusive: semantic processing has been found in unconscious subjects under anaesthesia (Munte

et al. 1999) and in coma patients (Kotchoubey et al. 2005). It is generally accepted that these results are “suggestive [but] not unequivocal evidence that a person is consciously aware” (Owen et al. 2006), given everything we know about the amount and variety of human cognition that can take place without conscious awareness.

Attributing consciousness to patients in the vegetative state directly on the evidence of neural activity thus seems problematic. This has led several researchers to the second approach, which involves attempting to interpret neural activity in terms of intentional action, and if successful, using intentional action as an indicator of consciousness in the standard way. Thus far, we have been concerned with intentional actions in the form of bodily movements: a hand movement that can be an act of grasping, for example, or a head movement that can be an act of watching. This is the standard interpretation of the notion of ‘purposeful behaviour’ that appears in the diagnostic criteria. But neuroimaging technology gives us access to the *behaviour of the brain*: if some neural behaviour counts as intentional action, then we might be able to ascribe consciousness to subjects on the basis of their neural activity rather than their bodily activity. This approach is also known as the ‘argument from volition’ (Shea and Bayne 2010).

The canonical version of the argument from volition is due to Owen et al. (2006), and uses an established fMRI task involving mental imagery. It is well known that spatial imagery tasks and motor imagery tasks activate distinct areas of the brain in healthy subjects. When subjects imagine that they are playing in a game of tennis, the supplementary motor area (SMA) is activated; when subjects imagine that they are moving from room to room in their house, very different brain areas are activated: the parahippocampalgyrus (PPA), posterior parietal-lobe (PPC), and lateral premotor cortex (PMC). Owen et al. (2006) took one patient who had been diagnosed as being in a vegetative state, and places them in an fMRI scanner. The patient was played pre-recorded instructions to imagine playing tennis or to imagine moving around the rooms in their house. Thirty seconds after each instruction, the patient was told to relax until the next instruction, thirty seconds later. Each imagery task and the rest condition were repeated ten times. The same experiment was done on twelve healthy control subjects, with the expected results: ten consecutive 30-second bursts of activity in the relevant brain areas, separated by thirty seconds each. Importantly, the researchers found the same patterns of neural activation in the vegetative patient as in the healthy controls: the patient diagnosed as being in a vegetative state exhibited neural responses that were “indistinguishable from those observed in healthy volunteers” (Owen et al. 2006).

What do these results show? Owen et al. (2006) interpret the results as demonstrating that the patient “retained the ability to understand spoken commands and respond to them through her brain activity” (Owen et al. 2006). The patient's capacity to respond is the key element of this interpretation, because by responding that the patient appears to be performing an intentional action. And intentional action is a marker of consciousness: when someone responds appropriately to a stimulus, this is a good indicator that they are aware of the stimulus. Owen et al. (2006) explicitly draw this conclusion when they say of their patient that her “imagining particular tasks when asked to do so represents *a clear act of intention*, which confirmed beyond any doubt that she was consciously aware of herself and her surroundings” (Owen et al. 2006, my italics).

A further example of the argument from volition can be found in a later paper reporting similar findings, again based on the spatial and motor imagery fMRI tasks. Reporting their findings, Monti et al. (2010) claim that of their 23 patients who were diagnosed as being in a vegetative state, four patients had “*voluntary*” responses in which they “were shown to be able to *willfully* modulate their brain activity through mental imagery” (Monti et al. 2010, my italics). After attributing intentional action to the patients in this way, Monti et al. (2010) then draw the inference from intentional action to consciousness, claiming that their data “provided clear evidence that the patient was aware” (Monti et al. 2010, my italics).

While I think the neuroimaging data is important and interesting, I do not share these researchers' certainty regarding the conclusions they draw. I agree with the conditional claim that *if* neuroimaging revealed intentional action, we would have good reason to attribute consciousness to the patients in question. But as the experiment stands, I do not think enough has been done to establish that the patients are performing intentional actions. In order to show why, I will need to say more about intentional action and its relation to mental imagery.

4. Mental imagery as intentional action

When we think about everyday examples of intentional action, the obvious cases are overt bodily actions such as kicking or waving. In the clinical case too, the sorts of 'purposeful behaviour' that is looked for in patients generally involves motor responses, such as reaching for an object or nodding in agreement. As Soteriou (2009) notes, the same focus on overt bodily action is found even in philosophical discussions of intentional action:

“The ordinary, common-sense notion of action is one we most naturally associate with overt bodily actions [...] and one finds this reflected in the emphasis placed on bodily action in the philosophical literature.” (Soteriou 2009, 232)

But the category of intentional action is not exhausted by overt bodily action:

"Some actions are 'overt': they essentially involve agents moving their bodies. Others are not: there are mental actions - for example, solving a chess problem in one's head, or deliberating about whether to accept a job offer." (Mele 1997, 231)

Indeed, this distinction between bodily and mental actions is what makes the argument from volition (outlined in the previous section) possible. People are diagnosed as in a vegetative state rather than a minimally conscious state precisely because they do not perform overt bodily actions; but what the neuroimaging experiments purport to show is that these patients might be able to perform mental actions. In the experiments detailed above, the intentional mental action in question is that of imagining oneself in a particular scenario. The argument from volition therefore relies on the claim that the patients involved are performing an intentional (mental) action of imagination, despite their inability to perform intentional (bodily) actions of the sort normally used to diagnose post-coma patients.

For the argument from volition to work, the neural data must be interpretable as an intentional mental act of imagining. How would we go about establishing this? One issue concerns specifying what it is for an intentional action to be an intentional *mental* action: as Soteriou (2009) acknowledges, "[s]pecifying what it is that makes an action a mental one, as opposed to a physical or bodily one, is not a straightforward matter" (Soteriou 2009, 232). He points out that the distinction is most often made clear in terms of examples of mental actions such as calculating in one's head, or trying to remember something, and concludes that what is common to such activities is that "they do not seem to require for their successful performance the performance of any *overt* bodily action" (Soteriou 2009, 233). This seems to allow that patients who are unable to perform overt bodily actions could nonetheless be attributed intentional mental actions.

But the more pressing issue, it seems, is specifying what it is that makes a mental state an intentional mental *action*. Not all mental states (or mental events, or mental occurrences) look to be intentional actions: my belief that Paris is the capital of France doesn't seem to be something I do, for example, and neither my fear of heights nor my headache seem to qualify as intentional actions. Some philosophers go so far as to claim that there is *no such thing* as intentional mental action, or at least that much of our mental activity should not be classed as intentional action. Strawson (2003), for example, suggests that most of our thought is neither voluntary nor purposeful in the sense required for genuine intentional action:

"Obviously thought involves—is—mental activity, but activity, whether mental or volcanic, does not always involve intentional action. And if we consider things plainly, we find, I think, that most

of our thoughts just happen. In this sense they are ‘spontaneous’. But they are not actions—not at all. Contents simply spring up—the process is largely automatic and largely involuntary.”

(Strawson 2003)

This sort of argument is most appealing in the case of idle thought, daydreaming, and ‘stream of consciousness’ types of mental activity. In such cases, it can certainly seem to us that thoughts pop into our heads without any intentional action on our part. But such mental activity is only a small part of our mental lives: mental processes like deciding and calculating, for example, seem to be obvious cases of things we do rather than things that happen to us. But if even if we accept that some mental activities count as intentional mental actions, what the argument from volition needs to establish is that *imagining* qualifies as intentional action. Strawson considers imaginings to be one of the intuitively more obvious examples of a mental action:

“Imagination has a remarkable effortlessness or fluency, and can indeed seem like a paradigm case of action—something we do and do intentionally—and indeed like a matter of ‘basic’ action—something we do intentionally and do not do by doing anything else.” (Strawson 2003)

But even in this case, he suggests that we should not be swayed into thinking of imagining as an intentional mental action:

“even in the case of imagination there seems to be a sense in which the entertaining of content is not itself any sort of action, but is, rather, a kind of reflex response that we are prone to experience as something we do intentionally when it occurs (as it normally does) without any sort of resistance.”

(Strawson 2003)

One might want to acknowledge that Strawson is right about some cases of imaginings but not others. In the case where one finds oneself having an auditory imagining of a song, for example, this can seem like something that simply happens to us against our will, often frustratingly:

Strawson seems to be right that this sort of imagining is not an intentional action. But what about cases where we strive to remember a song, actively trying to entertain the relevant auditory imagining? Such cases look like a much more obvious candidate for intentional action. Peacocke (2009) puts forward a view of intentional mental action that allows for this distinction:

“Some types of mental event are such that instances of the type may or may not be mental actions. Such is the case with imagining. Imagining in your mind’s ear Beethoven's Hammerklavier Sonata may on a particular occasion be a mental action. On another occasion, that Sonata may equally come to your auditory imagination unbidden - your imagining may be a hindrance to what you are trying to do. In this respect, imagining as a type is like the bodily type of making marks on the carpet. When someone is making marks on the carpet, that may or may not be something she is trying to do.” (Peacocke 2007)

A similar idea is discussed by Mele (1997), who suggests a game that involves thinking of seven kinds of animals beginning with 'g'. Someone who completes this game has "performed an action describable as "thinking of (or mentally identifying) seven kinds of animal having a name starting with 'g'."", one such animal perhaps being 'goat' (Mele 1997, 234). Mele then suggests an alternative scenario:

"in the absence of a desire or intention to play my little language game, your reading this [...] caused a thought of 'goat' to come to mind. We can say, if you like, that you thought of 'goat' in this scenario. But, intuitively, your thinking of 'goat' was not an action of yours, much less an intentional action." (Mele 1997, 234-235)

I want to now apply the insights of Peacocke and Mele to the mental imagery task outlined in the previous section. When someone has mental imagery of themselves playing tennis, or mental imagery of the rooms in their house, they are certainly in a mental state of *imagining*. But that leaves open whether their imagining the relevant scenario is an intentional mental action or a mere mental 'happening'. Recall that for the argument from volition to work and consciousness to be attributed, the subject must be performing an intentional mental action of imagining rather than undergoing a mere mental happening. The data from the control subjects tells us that when they are they have mental imagery corresponding to playing tennis, they have neural activity in the SMA; and that when they have mental imagery corresponding to moving from room to room in their house, they have neural activity involving the PPA, PPC, and PMC. And it is true that they display such activity after having been instructed to imagine the appropriate scenario, where such imaginings were presumably performed as intentional actions. But – and this is the key problem with the data – we don't know what the neural activity would look like if these subjects were merely undergoing mental imagery of the relevant kind without trying to. What if subjects' neural activity showed the same distinct patterns (SMA activation for the tennis imagery; PPA, PPC, and PMC activation for the house imagery) when they simply happened to be having such imagery unbidden? If this were the case with the control subjects, then observing the same neural activity in the vegetative patient wouldn't license the conclusion that they were performing an intentional action.

This raises a problem for the methodology of the experiment, but one which might be easily rectified. All that seems to be needed is a further control group: a group of healthy subjects who undergo fMRI scanning while undergoing the tennis and house imagery, *without* doing so an intentional mental action. If this gave us different neural activity from the intentional action case,

then the interpretation of Owen et al. (2006) would be warranted. This, however, is where practical concerns come in: how do we design such a neuroimaging task? How does one induce imaginative states with particular contents in subjects, without doing it in such a way that the subjects are purposefully imagining the appropriate the scenarios? It's not clear how we could ever instruct people to non-willfully imagine a scenario while in the experimental setting required to obtain the neural data, at least with any degree of reliability that the imaginings in question were not intentional actions. But until such an experiment can be designed, such a control study cannot be provided, and we have no reason to think that the neural activation patterns associated with tennis and house imagery are correlated with intentional acts of imagination rather than correlated with all imaginings of the same scenarios, including 'mere happenings'. And without being justified in interpreting the neural data as intentional mental action, we don't have evidence of agency to use as a marker for consciousness.

5. Conclusions

I have argued that we don't have sufficient experimental data to be justified in concluding that patients in the vegetative state displaying certain patterns of neural activity are thereby performing intentional mental actions. Without the attribution of intentional agency, we cannot use the argument from volition attribute consciousness to these patients. It is interesting to note that Owen et al. (2006) take themselves to have rejected the possibility of unintentional imaginings without conducting the relevant controls. They suggest that the design of the experiment, and the tasks involved, are sufficient to rule this out:

“Although it is theoretically possible that the mere instruction to imagine such actions triggers specific and automatic changes in brain activity, the complexity of the commands used here and the richness of the imagery that is likely to be required to produce a response that is indistinguishable from that of healthy individuals, make this possibility extremely unlikely.” (Owen et al. 2006)

Owen et al. (2006) are claiming that *either* such a neural response is an intentional action, *or* it is an automatic response. They further argue that it cannot be an automatic response, due to the nature of the task. But it's not clear that the dichotomy they set up is correct. It is true that in the case of overt bodily movement, we often assume that it is either an intentional action or a mere reflex. But once we start thinking about neural activity and intentional mental action, the appropriateness of the dichotomy starts to become strained.

To see this, consider the distinction between ‘automatic’ and ‘controlled’ thought processes that was introduced in the mid-1970s. As Bargh points out:

“an automatic thought process was initially defined as one that is capable of occurring without the need for any intention that it occur, without any awareness of the initiation or operation of the process, and without drawing upon general processing resources or interfering with other concurrent thought processes” (Bargh 1989, 3)

In strict opposition to automatic thought processes were placed those thought processes that were ‘controlled’:

“Conscious or controlled processes, on the other hand, were defined as those that are under the flexible, intentional control of the individual, that he or she is consciously aware of, and that are effortful and constrained by the amount of attentional resources available at the moment” (Bargh 1989, 3)

According to what Bargh calls “this dual-mode model of cognition” (Bargh 1989, 4), any particular mental process was either controlled or automatic, where these two modes were assumed to be mutually exclusive and exhaustive. As a result, any process that didn’t meet all of the criteria for being automatic was considered to be under intentional control. But this model of cognition has been seriously challenged by Bargh and colleagues, who argue that intention, efficiency, and control are separate issues. According to this view, a mental process that failed to meet some of the criteria for automaticity, for example, shouldn’t necessarily be classified as an intentional mental action (see Bargh 1989 for further discussion and examples).

If Bargh is correct to question the dichotomy between automatic and intentional mental processes, then it looks like Owen et al. (2006) are too quick to assume that their vegetative patients are performing intentional mental actions. They might be right to claim that their experimental tasks rule out the possibility that the patients’ imaginings are automatic (in the sense of reflex) processes, but this would not entail that the imaginings are intentional mental actions.

In conclusion, I suggest that the notion of intentional mental action would benefit from closer scrutiny, both in itself and in its relation to other kinds of mental processes and activities. It is all too easy to move from talking about intentional bodily action to intentional mental action, without questioning whether there are further issues raised in the process. If there is consciousness in the vegetative state, the argument from volition seems like the right approach to uncover it, but where the notion of ‘volition’ involved should be carefully explored.

Acknowledgements

Work on this paper began on David Chalmers' ARC-funded project on 'The Contents of Consciousness' at the Australian National University, and an early version of this paper was presented at the University of Western Australia in February 2012. In its current version, this paper was presented at the first Topoi conference in November 2012, where it benefitted immensely from the conference participants, in particular: Marcel Brass, Christiano Castelfranchi, Elisabeth Pacherie, Fabio Paglieri, Markus Schlosser, Corrado Sinigaglia, Bruno Verbeek, and Tillman Vierkant. Thanks are also due to Topoi's anonymous reviewers.

Bibliography

- Block, N. (2007) Consciousness, Accessibility, and the Mesh between Psychology and Neuroscience. *Behavioral and Brain Sciences* 30: 481-548.
- Chalmers, D. J. (1997) Availability: The cognitive basis of experience? N. Block, O.J. Flanagan & G. Guzeldere (eds.), *The Nature of Consciousness*. MIT Press.
- de Jong, B., Willemsen, A. T. and Paans, A. M. [1997]: 'Regional Cerebral Blood Flow Changes Related to Affective Speech Presentation in Persistent Vegetative State', *Clinical Neurology and Neurosurgery*, 99, pp. 213–16.
- Giacino, Joseph T., Stephen Ashwal, Nancy Childs, Ronald Cranford, Bryan Jennett, Douglas I. Katz, James P. Kelly et al. "The minimally conscious state Definition and diagnostic criteria." *Neurology* 58, no. 3 (2002): 349-353
- M.A.F. Gross, W. Nager, C. Quandt, T.F. Münte, S. Münte. Auditory processing in patients under general anaesthesia. An event related brain potential (ERP) study. *British Journal of Anaesthesia*, 93 (2004), p. 486
- Boris Kotchoubey, Jérôme Daltrozzo, Norma Wioland, Veronique Mutschler, Philippe Lutun, Niels Birbaumer and Albert Jaeger (2005), Semantic processing in a coma patient, *Grand Rounds* Vol 5, 37–41
- Menon, D. K., Owen, A. M. and Williams, E. J. et al. [1998]: 'Cortical Processing in Persistent Vegetative State', *Lancet*, 352, p. 200.
- Monti, M. M., Vanhaudenhuyse, A., Coleman, M. R., Boly, M., Pickard, J.D., Tshibanda, L., Munte S, Kobbe I, Demertzis A et al. Increased reading speed for stories presented during general anesthesia. *Anesthesiology* 1999; 90: 662–9.

- Owen, A. M., Coleman, M. R. and Menon, D. K. et al. [2005]: 'Residual Auditory Function in Persistent Vegetative State: A Combined PET and fMRI Study', *Neuropsychological Rehabilitation*, 15, pp. 290–306.
- Owen, A. M. and Laureys, S. (2010). 'Willful Modulation of Brain Activity in Disorders of Consciousness', *N Engl J Med*, 362: 579-89.
- Owen, A. M., Coleman, M. R., Boly, M., Davis, M. H., Laureys, S. and Pickard, J. D. (2006). 'Detecting Awareness in the Vegetative State', *Science*, 313, 1402.
- Owen, A. M., Coleman, M. R., Boly, M., Davis, M. H., Laureys, S. and Pickard, J. D. (2007). 'Using Functional Magnetic Resonance Imaging to Detect Covert Awareness in the Vegetative State', *Archives of Neurology*, 64, 1098–102.
- Peacocke, C. (2007). 'Mental action and self-awareness (I)', in McLaughlin, B. and Cohen, J. (eds.) *Contemporary Debates in Philosophy of Mind*. Oxford: Blackwell.
- Plum, F., & Posner, J. B. (1982). *The Diagnosis Of Stupor & Coma* (Vol. 19). Oxford University Press, USA.
- C. Quandt, W. Nager, M.AF. Gross, N. Löffler, S.A. Piepenbrock, T.F. Münte, S. Münte. Auditory perception during general anaesthesia with sevoflurane-event related potentials (ERP) indicate orienting mechanisms. *British Journal of Anaesthesia*, 93 (3) (2004), p. 488
- Rodd, J. M., Davis, M. H. and Johnsrude, I. S. [2005]: 'The Neural Mechanisms of Speech Comprehension: fMRI Studies of Semantic Ambiguity', *Cerebral Cortex*, 15, pp. 1261–9.
- Royal College of Physicians (2003). *The vegetative state: guidance and diagnosis on management*. London: RCP.
- Shea, N. and Bayne, T. (2010). 'The Vegetative State and the Science of Consciousness' *Brit. J. Phil. Sci.* 61, 459–484.
- Strawson, G. (2003). 'Mental ballistics or The involuntariness of spontaneity'. *Proceedings of the Aristotelian Society*, 103, 227-256.