ABSTRACT

A Permanent Vegetative State (PVS) is a condition in which one’s consciousness is irreversibly lost even though one can spontaneously breathe and can stay alive without continuous external support. Many philosophers believe that PVS-patients lack a significant moral status. I call their position “PVS-exclusion.” One popular argument for PVS exclusion is that upon entering PVS, persons lose their personal identity and with it their moral status. In this paper, I argue that PVS-exclusion cannot be motivated by considerations of personal identity, for the following reason. If considerations of personal identity support PVS-exclusion, they also imply that individuals in temporary vegetative states lack a significant moral status. This result is morally counterintuitive and goes beyond what PVS-exclusionists would wish to affirm, and it cannot be avoided if we follow any standard theory of personal identity. I conclude that if personal identity matters to moral status, PVS-patients have a significant moral status.

A permanent vegetative state (PVS) is a condition in which an individual irreversibly (given our current medical knowledge) loses consciousness. PVS typically involves a significant loss of functioning in the cerebrum, but unlike brain-death (as it is usually defined), it does not involve damage to the brainstem. Since in PVS the brainstem often remains intact, PVS-patients can typically breathe spontaneously, have sleep/wake cycles, cough when their throat is irritated and swallow food when it is placed in their mouths.¹

Many philosophers believe that individuals in PVS have no significant moral status, and so no moral justification is needed to unilaterally withdraw treatment from them, for example, or to use their organs for donation.² Let us call this position “PVS-exclusion.” My aim in this paper is to criticize a

common argument in support of PVS-exclusion, according to which individuals in PVS have no significant moral status because they are no longer identical to any person.³

The argument for PVS-exclusion I have in mind relies on the following normative principle:

**The Numerical Identity Principle:** If, and only if, an individual at time $t_1$ is numerically identical to an individual who is a person at time $t_2$ (where $t_2$ may be any time) that individual has significant moral status at $t_1$.

According to the Numerical Identity Principle, what matters for moral status is whether one was, is or will be a person (however we define the term). PVS-exclusionists can rely on this principle and point out that most definitions of personhood imply that consciousness is at least a necessary condition for personhood. One simply cannot do the things persons do (respond to reasons as such, have second-order volitions, have the capacity to care, have a first-person perspective, and so on) in the complete absence of consciousness.⁴ Moreover, it seems initially plausible to argue that one cannot remain numerically the same individual when one’s personhood is lost. When personhood ends, one ceases to be.

Putting all this together, we can reconstruct the argument for PVS-exclusion as follows:

**The Argument from Personal Identity:**


2) Individuals in PVS are not persons.

³ Supporters include McMahan; Soifer; Warren; Wikler; Wrigley.

⁴ Which is not to deny that these capacities require a great deal of subconscious psychological activity.
3) Individuals who are not persons at \( t_1 \) are not numerically identical to any individual who is a person at \( t_2 \) (where \( t_2 \) may be any time).

\[ \text{Therefore:} \]

4) Individuals in PVS are not numerically identical to any individual who is a person at any time (from (2) and (3)).

\[ \text{Therefore:} \]

5) **PVS-exclusion:** Individuals in PVS do not have a significant moral status (from (1) and (4)).

In this paper, I do not question the Numerical Identity Principle, nor will I deny the claim that individuals in PVS are not persons. My argument focuses on premise (3): “Individuals who are not persons at \( t_1 \) are not numerically identical to any individual who is a person at \( t_2 \).” I argue that if we accept this premise, PVS-exclusionists cannot differentiate the moral status of individuals in PVS and individuals in temporary conditions of unconsciousness, and that this would be a reductio of PVS-exclusion.

Let us define an individual in Temporary Coma (TC) as an individual in a vegetative state whose current responsiveness to the environment is identical to that of an individual in PVS (i.e. no responsiveness), but whose brain is such that, if her body is kept in good condition for a certain period of time, she would recover and be able to lead satisfying daily life. TC is not a medical category, but an ethical category. No doctor would tell family members that if only taken good care of, their beloved sibling or parent or child, who became unconscious due to accident or cardiac arrest, would recover. Nor can our current medical knowledge guarantee a recovery in such cases. However, there are patients of which the mentioned conditional is true, so there are patients who are in TC, even if we do not know who they are. For those patients, I assert the following:

*Patients in TC have a significant moral status.*
I take this claim as a bedrock moral judgment. It expresses the view that if it is possible for a patient to recover if she receives good treatment, we owe her that good treatment (assuming the burden this would impose on other patients and on the system would not be excessive, etc.). Killing her while she is in her temporary vegetative state would be murder. Taking her organs for donation without her consent would be morally wrong. I believe that PVS-exclusionists would agree with all this. Their thesis denies only patients in permanent vegetative states a significant moral status. For this reason, I take it that any argument in support of PVS-exclusion that also implies that patients in TC lack a significant moral status would be rejected by PVS-exclusionists. It would be, in other words, a reductio of their view. My argument is that premise (3) leads to such a reductio.

2. Personal Identity

When can an individual be said to be numerically identical to an individual who is a person? According to the Biological Criterion of Personal Identity, “if $X$ is a person at $t_1$, and $Y$ exists at any other time, then $X=Y$ if and only if $Y$'s biological organism is continuous with $X$'s biological organism.”5 The biological criterion does not support PVS-exclusion: the biological organism of a PVS-patient is biologically continuous with the biological organism of the person her body once embodied. The only biological difference between a pre-PVS person and a PVS-patient is in the cerebrum, which is hardly enough to make them biologically discontinuous.

PVS-exclusionists need a different criterion of personal identity if they wish to affirm the non-identity of a PVS-patient and the person who once had the same name. The criterion that most clearly suggests itself is The Psychological Criterion of Personal Identity: “$X$ at $t_1$ is the same person as $Y$ at $t_2$ if and only if $X$ is uniquely psychologically continuous with $Y$, where psychological continuity consists in

---

overlapping chains of strong psychological connectedness, itself consisting in significant numbers of
direct psychological connections like memories, intentions, beliefs/goals/desires, and similarity of
character.⁶ A PVS-exclusionist can appeal to this criterion and say: Since a PVS-patient has no
consciousness, she has no psychological continuity with any person.

But this reasoning implies the following: while in the vegetative state, TC-patients also have
no consciousness, and so, according the Psychological Criterion, are not numerically identical to any
person. Together with the Numerical Identity Principle, this would entail that they lack significant
moral status, which is exactly the result a PVS-exclusionist would wish to avoid.

A PVS-exclusionist who wants to rely on the Psychological Criterion would probably wish to
revise it or supply it with some other metaphysical principle to support the claim that a permanent
loss of consciousness ends personal identity while a temporary loss does not.

Here is one possible way to do this. Imagine that Sarah is in TC because of an accident. After
a few weeks, she regains consciousness. Let us say that Sarah remembers who she was before the
accident. She remembers her family, friends, and deeply held values. I think it should be open to a
supporter of the psychological criterion of personal identity to say that, given the psychological
continuity (at the content level, even if not at the temporal level) between pre-accident and post-
recovery Sarah’s psychological states, they are the same person.⁷

Now, there is a widely-held metaphysical doctrine according to which identity admits of no
temporal gaps. Once one ceases to be, one can never go back into being. We cannot take a break from
existence and jump back into it. Nothing can.⁸ So, if post-recovery Sarah and pre-accident Sarah are

---

⁶ Ibid, section 2.1.
⁷ This is a revision of the psychological criterion as stated above, because it gives up on “overlapping chains of
strong psychological connectedness,” but one could make the case that this is a reasonable revision of the
criterion.
⁸ For a vivid illustration and discussion, see Peter van Inwagen, “The Possibility of Resurrection,” *International
numerically identical, there is no time in between in which Sarah’s numerical identity was cut. Sarah remained one and the same Sarah during her temporary vegetative state, despite not having any psychological continuity during that time.⁹

How can this help to support PVS-exclusion? Compare Sarah with Eve. Eve had an accident that caused her to be in PVS. Eve can never recover from her vegetative state. She will never remember who she was before the accident, because she will never be conscious again. In her case, she will never have future psychological states that will “tie up” her numerical identity with who she was before the accident. She will never have those crucial memories. Given the absence of consciousness while in PVS, and assuming the Psychological Criterion of Personal Identity, we should conclude that Eve lost her numerical identity in the accident. The metaphysical no-gaps principle that allowed us to ascribe personal identity to Sarah during her vegetative state does not apply to Eve. Eve’s vegetative state is not followed by conscious memories of who she was before the accident, memories that would have allowed us to say that pre-accident and post-recovery Eve are the same person, so Eve’s vegetative state is not a phase to which we should attribute continued personal identity. This is how the Psychological Criterion, supplied with a metaphysical no-gaps principle, will differentiate Sarah and Eve. A Biological Criterion, by contrast, would take both Sarah and Eve to be identical to their past persons.

So far so good. But what happens if a TC-patient has not recovered?

Consider the following case: While in a vegetative state, there is an earthquake that destroys the hospital and kills Sarah (who is still in TC). Up to the earthquake, Sarah has a significant moral status: it is morally impermissible to kill her or withdraw life-support from her. By implication of the

⁹ This might seem to support the Biological Criterion of personal identity, but it does not have to: we affirm personal identity based on the memories and other mental states of the post-recovery person, but supply it with another general metaphysical principle that is true of all objects.
Numerical Identity Principle, Sarah, while in TC, is numerically identical to pre-accident Sarah. But Sarah will never recover. She will never have the memories crucial for personal identity. So in what way is Sarah different from Eve?

I imagine that a PVS-exclusionist would like to appeal to a modal difference between Sarah and Eve at this point. *Had* Sarah recovered she would have been identical with her pre-accident person, and, up to the point of her death, it is *possible* for Sarah to recover. It does not matter if Sarah *actually* recovers to determine her moral status. If Sarah recovers, she would be the same individual she once was only if during the time she was in a vegetative state she did not lose her identity. Sarah has to exist without gaps during her vegetative state for her recovery to be *her* recovery. Sarah’s actual recovery and memories cannot *make it retroactively the case* that she retained her identity all along. If she hadn’t retained her identity all along, no memory of the recovered Sarah could give her her identity back, because there can be no gaps in existence. So it is the *possibility* of recovery, the possibility of having the memories crucial for personal identity, not the actual recovery, that retains the personal continuity of a TC-patient during her vegetative state.

So: if it is *possible* for a patient in a vegetative state to be uniquely psychologically continuous with a person, she is numerically identical to that person. Does this revision of the Psychological Criterion help PVS-exclusion?

Consider the following case. Sarah is in TC and there is no earthquake. However, Sarah’s evil doctor decides to kill her. He is a very competent and strong-willed killer: he never fails to kill someone when he intends to do so. He prepares a powerful poison and is just about to pour it into Sarah’s feeding tube. Let’s pause. What is Sarah’s moral status right now? Presumably she has a significant moral status. What Dr. Evil is about to do is murder. But at this moment in time, it is impossible for Sarah to recover. Once the doctor intends to kill her, and especially once he stands with the poison above her, her fate is doomed: it is inevitable that she will die. But if we accept the revised
psychological criterion of personal identity, at this moment Sarah lost her numerical identity with pre-accident Sarah and lacks a significant moral status for this reason. If we follow the (modal) Psychological Criterion, it is quite permissible for Dr. Evil to put the poison in her feeding tube, and his action is not a murder. This cannot be right.\textsuperscript{10}

Appeals to the possibility of recovery cannot explain the moral difference between TC-patients and PVS-patients that the PVS-exclusionist needs. It could be just as impossible for a TC-patient to recover as it is for a PVS-patient.

At this point, a PVS-exclusionist is likely to protest: it is possible for Sarah to recover. Dr. Evil might have a heart attack just before pouring the poison in her feeding tube. An earthquake might kill him before he manages to kill Sarah (and the earthquake might miraculously not kill Sarah). Dr. Evil might (although that never happened before) listen to the voice of conscience. But let us assume that these are very unlikely scenarios. If the only sense in which it is possible for Sarah to recover is if one of these scenarios happens, then in that sense of “possible” it is also possible for Eve to recover: maybe a medicine will be developed that will allow people in PVS to recover? When we define someone to have PVS we do so in relation to our current medical knowledge. To be in a PVS is not to be in a condition in which it is metaphysically impossible for someone to recover, but in which it is medically impossible, given our current and projected medical knowledge.\textsuperscript{11}

\textsuperscript{10} This is not like the case of abortion Harman describes in her seminal paper on abortion. Harman relies in that paper on something very similar to the Numerical Identity Principle, what she calls The Actual Future Principle. The Actual Future Principle says that one has a significant moral status if one will be a person. So it is simply the Numerical Identity Principle, applied to the future and not mentioning the past. The Actual Future Principle (and so The Numerical Identity Principle) indeed support a very liberal view in relation to abortion. In the abortion case, according to Harman, the act of aborting an early fetus determines its moral status, because in the case of the early fetus, it is still an open question whether it will ever be a person. In the case of someone in PVS, it is usually a fact that she was once a person, and no action in the future can change that. So the Numerical Identity Principle can support a very liberal view of abortion but a pretty conservative view of PVS. See also, Harman Elizabeth Harman, "Creation Ethics: The Moral Status of Early Fetuses and the Ethics of Abortion," \textit{Philosophy & Public Affairs} 28, no. 4 (1999).

\textsuperscript{11} McMahan (2002) disagrees. In the longer version of this paper I discuss his objection.
A PVS-exclusionist will probably say: “but the brains of an individual in TC is different from the brain of an individual in PVS! An individual in TC has a brain that allows for recovery, while an individual in PVS has a brain that (given our current medical knowledge) cannot recover. So the reason why it is impossible for Sarah to recover in the Dr. Evil case is very different from the reason why it is impossible for Eve to recover. This surely has to count for something!”

But does it? We should care about brain structures if they correspond to ethically significant categories. The categories we now examine are “possible to recover” and “impossible to recover,” which PVS-exclusion might propose as a morally relevant difference between TC-patients and PVS-patients. But the difference turns out not to be morally relevant: it might be equally impossible for a TC-patient to recover as it is for a PVS-patient. Why should the reasons for the impossibility make any difference?

But a PVS-exclusionist will insist: Sarah, even as Dr. Evil attempts to kill her, can in principle recover. She has the kind of brain that, if she will not be killed and if she will be taken care of, will allow her to recover. While it is impossible for her to recover, it is possible for her in principle to recover, even as Dr. Evil is attempting to kill her. For an individual in PVS, it is impossible in principle to recover. So it is not just a difference in the reasons for impossibility, it is a different kind of impossibility.

But why should this distinction between kinds of impossibility matter for the attribution of numerical identity? Recall how we got here: we started with actual recovery, and observed that in such a case, it makes sense to attribute continuity to the individual based on her memories and other mental states. We then said that it would make sense to do so only if while unconscious, the individual maintained her personal identity. And we said that an individual maintains her personal identity because we have a sense that if that individual actually recovered, she would be the very same individual (given the memories that she would have). That possibility is absent when it is impossible for one to recover. Does it matter whether this possibility is absent “in principle” or just simply absent?
It seems to me that it could matter only if the different kinds of impossibility introduce different strengths of modality. For example, if in a case in which it is impossible in principle for someone to recover, the modality of “impossible” is much stronger than in the Dr. Evil case, perhaps the strength of the modality can be a non-arbitrary difference between TC-patients and PVS-patients. A PVS-exclusionist might want to say: A TC-patient is numerically identical to the person she once was, while a PVS-patient is not, because it less impossible for a TC-patient to recover compared with a PVS-patient. The impossibility of recovery of a TC-patient in a Dr. Evil scenario is a weaker kind of impossibility than the impossibility of recovery of a PVS-patient.

But if we agree that it is not metaphysically impossible for a PVS-patient to recover, it is hard to see why the impossibility of Eve’s recovery is “stronger” than Sarah’s. After all, it is not entirely impossible for Eve to recover, and given that Dr. Evil is a really efficient killer, it is very much impossible for Sarah to recover.

1. Conclusion

To conclude, considerations of personal identity, which are common in discussions of PVS, do not support the claim that TC-patients and PVS-patients have a different moral status. In fact, leading criteria of personal identity suggest that either both TC-patients and PVS-patients are numerically identical to past persons, or both are not.

We are now in a position to construct an argument from personal identity in favor of the significant moral status of PVS-patients, based on the Numerical Identity Principle: Either both TC-patients and PVS-patients are numerically identical to the persons they once were, or both are not; TC-patients have a significant moral status, so, if the Numerical Identity Principle is correct, they are
numerically identical to the persons they once were; so PVS-patients are also numerically identical to the persons they once were; so PVS-patients have a significant moral status.