Can one still speak meaningfully of personhood today?

Challenging Farah and Heberlein's argument

Bernard Baertschi

1. Introduction

The concept of 'personhood' or of 'what it means to be a person' has been a central tenet in the question of moral status for a long time. A person, Boethius claimed in the 5th century, is an individual endowed with reason (or rationality). Aquinas, Locke and Kant adopted this definition. Its centrality has been contested by several moral philosophers, notably by utilitarians (they believe that the property to possess moral status is not reason, but sentience) and by environmental ethicists. Recently, neuroscientists and neuroethicists have levelled a new charge. In a paper entitled 'Personhood and Neuroscience: Naturalizing or Nihilating?' published in 2007, Martha Farah and Andrea Heberlein argued that this concept, even if it has played a major role in our moral tradition (it is linked with moral agency and responsibility, and then 'a foundational concept in ethics' (2007: 39)) should be dispensed with. They think that its neural basis consists in an automatic response from an innate brain module or network that is dedicated to face recognition and therefore is alien to morality (and often even to reality – they speak of an 'illusion'). For these authors, we would be well advised to turn to a conception of moral status grounded in interests, a conception akin to utilitarianism.

In my talk, I will examine the charge. Conceding that the concept of 'personhood' is based on an innate brain module, I will question the idea that this fact discredits it as a criterion for moral status. I will argue that having a basis in a brain module is not a problem for a moral concept and that a psychological property like rationality has no intrinsic feature that could exclude it from playing a moral role. Moreover, a property like 'having interests' does not fare better, despite what the authors seem to think. In fact, the gist of their argument or objection has nothing to do with neuroscience, but with the classic claim that 'personhood' is an unusable concept because it is loose; an objection I will also answer.

Nowadays, utilitarian thinkers seem to hope that neuroscience will buttress their position. Greene (2008) has developed arguments going in the same direction not in relation with moral status, but with normative theories. However, I don't believe they have succeeded till now. Philosophical arguments still stand true when normative questions are debated.

2. Farah and Heberlein's argument

Etymologically 'person' comes from the Latin 'persona' and from the Greek 'prosopon'. In the beginning, these words referred to a theatre mask, and before a face ('what is before the eyes of somebody'). Also, it is in this sense that neuroscientists use it when they speak of 'prosopagnosia', for a condition where a patient does not recognise the faces of the people around him anymore, even if she is still able to identify them with the aid of the tone of their voices or through their figure – a condition caused by a lesion of their inferotemporal cortex. Prosopagnosia is a very disturbing condition, psychologically and socially, because face recognition is central to our relations with the people we live with. For Farah and Heberlein, the psychological phenomenon of face recognition is linked with the ethics of personhood, that is with the crucial role the moral status of person plays in most of our ethical approaches. A view I will name 'personism', following Jean-Yves Goffi (2007).

For Farah and Heberlein, if being a person is so important to us, it is because face recognition is innate and very soon into play: 'Evidence for the innateness of the person–non-person distinction comes from the behavior of newborn infants. Johnson et al. showed that newborns tested within 30 minutes of birth show a greater tendency to track moving face-like patterns with their eyes than other patterns of comparable complexity or symmetry. This finding implies that, prior to virtually any opportunity to learn, the human brain is equipped with a general representation of the appearance of the human face' (2007: 43). To buttress their claim, they still mention the case of a boy: 'Another demonstration of innateness in person processing comes from the study of a boy who sustained visual cortical damage, including damage to the fusiform face area, in his first day of postnatal life. Despite his relatively preserved ability to recognize non-face objects, he never acquired the ability to recognize faces. In other words, a certain region of cortex is destined for face recognition as early as age 1 day, and other regions, which are capable of recognizing inanimate objects, cannot take over this function.' (2007: 43)

At first sight, it seems that these observations do not prove anything: the boy does not recognize faces, but he knows that he is before persons. Of course, because as we have said before, there are other means to gain access to persons. But Farah and Heberlein think that cases like this allow us to conclude that it is on the basis of our capacity to recognise faces that we have built our view of moral status. Briefly said, we are so wired that we separate spontaneously and naturally persons and non-persons. A consideration that could also be important for understanding the condition of autistic persons.

This is a piece of psychology. But what should we think of this natural tendency ethically speaking? Is its moral impact justified? A personist will not be very happy with that: every human being possesses a face, but every human being is not a person, because every human being is not endowed with rationality – think of some PVS patients or of anencenphalic babies. Moreover, every genetically human being does not possess a face – think of embryos – and there exist non-human beings who

are persons, as Boethius already said, for example angels and God. Engelhardt mentioned ET in the same context (1986: 107). Farah and Heberlein also deem that this natural tendency can lead us astray when they state: 'The human face is a powerful trigger cue that activates the whole person network, and this may be what makes it hard for many of us to dismiss the personhood of a vegetative patient or a fetus' (2007: 45). Our innate capacity to recognise faces is obviously a source of confusion because it compels us to grant the moral status of a person to everyone possessing a face, and this is not a criterion more appropriate than, for instance, species belonging ('speciesism' is a charge often voiced against such criteria).

Farah and Heberlein extend the critique against personism as such. They think that believing that the world is divided between persons and non-persons is an illusion from a moral point of view, and even an entrenched illusion, like all the illusions that are grounded in an innate brain mechanism. In their opinion, ethics should not be built on such a foundation and we will be well advised to substitute a better one for it, centred on the protection of the interests of all the individuals who possess interests – a kind of utilitarianism.

Is it appropriate from a moral point of view? In order to answer this question, I will first examine more thoroughly Farah and Heberlein's objection. Then I will ask if their proposal is morally better than the personist one, because even if their objection could be correct, their own proposal could suffer from shortcomings, too.

Farah and Heberlein's argument consists of a threefold charge: the concept of a person rests on an illusion, it is arbitrary and it is a categorical concept, ill suited to the gradual character of our psychological and moral life. Of course, they do not deny the existence of persons in the sense of beings endowed with rational psychological states, but they deny the relevance for ethics of a moral status named 'person' or 'personhood'.

3. A reply to the charge of illusion

'The first relevant feature of the person network in the brain is its separateness from the systems representing other things. We suggest that this feature is responsible for the illusion that persons and non-persons are fundamentally different kinds of things in the world, despite our inability to draw a principled line between them. This illusion may come from the operation of two separate and incommensurate systems of representation in the brain for persons and for things in general, in contrast to a common distributed representation' (Farah and Heberlein, 2007: 45).

Our brain divides the world in a manner it is not truly divided: the brain does not carve out the world at its joints. Consequently, concerning personhood, 'like visual illusions, it is the result of brain mechanisms that represent the world nonveridically under certain circumstances.' (2007: 45) Of course, there exist clear cases of persons – adults, for example – but others are not clear, and their number is growing with the progress of medicine (embryos, fœtuses, comatose persons, psychopaths, etc.). Therefore, to believe that the world is tightly divided between

persons and non-persons is a mistake, of the nature of an illusion, because it is a mistake that is not responsive to arguments or reform. Consequently, we should better ground our ethics on a firmer foundation, like interests.

Is the charge convincing? We could already doubt that the illusory character of the distinction could be an argument against personism, in the sense that the illusion is not that there exist persons, but that the divide is sharp. But that is only one aspect of the illusion that I will discuss later (it is the third charge); another one is that we extend personhood to beings that are not able (or not able anymore) to have interests, like some comatose individuals, granting them moral status they should not have.

I could reply in biting the bullet: why not grant them this status and consider, like the Law, all born human beings as persons? Pragmatically, it could be a wise stance to adopt, protecting vulnerable people, allowing abortion and embryo research. But most personists would be dissatisfied with this reply, some because they believe that human beings without a face are also persons and some others because they separate personhood and humanity since rationality is not necessarily linked with the latter. Consequently, another reply will be preferable. In short, my answer will be that personhood could well be anchored in a brain module for face recognition, but that this link is only contingent and has for long been severed by moral philosophers adopting personism. In order to understand the gist of this reply, I should examine in more length how the attribution of moral status is working.

When we ask someone: 'Why do you grant moral status to *X*', she usually points to some property of *X*. For example: 'Human beings have moral status because they are rational beings', or 'Animals are morally considerable because they are sentient (i.e. they can feel pleasure and suffer)'. 'To be rational' or 'to be sentient' are properties conferring moral status. But how and why? What is the nature of these properties and why are they capable of conferring moral status?

What is the nature of these properties? They are *intrinsic* properties. An intrinsic property, says Warren, is a property a being possesses independently of its environment: 'A thing's intrinsic properties are those which it is logically possible for it to have had were it the only thing in existence' (1997: 21). Intrinsic properties are contrasted with *extrinsic* ones, that is, with properties a being possesses because of its environment. The two main kinds of extrinsic properties are relational ones and instrumental ones. A human being possesses intrinsic properties (e.g., reason), relational ones (e.g., he is loved) and instrumental ones (e.g., he is a waiter). If we have some hesitation about categorising a property (think of 'to be blue': is it intrinsic or relational?), a syntactic test will give the right answer: intrinsic properties are monadic, i.e. they have only one place free, whereas extrinsic ones are polyadic, i.e. they have more than one place free: $F(x)$ is intrinsic, $F(x,y)$ is extrinsic (Bochenski, 1959: 43 and 51). Therefore, 'to be blue' is intrinsic and 'to be bluer than' is extrinsic (relational). Nevertheless, I will not dwell on this logical point. Let us then admit that moral status is conferred by intrinsic properties and not by extrinsic ones.

How do such properties confer moral status? Not all intrinsic properties confer moral status; they must also be valuable. These properties ground values; intrinsic properties ground intrinsic values, and external properties external values. Following G. E. Moore, we can say that a being's intrinsic value is the value he possesses by virtue of his intrinsic (or internal) properties (Rabinowicz and Rønnow-Rasmussen, 2000: 33; Baertschi, 2012: 9-10). Technically, we say that values *supervene* on properties.

A meta-ethical observation is still in order. This thesis does not have any privileged relationship with moral realism. It states that intrinsic value supervenes on intrinsic properties, but remains silent on the nature of this relationship. It can be read in a realist spirit, but in an antirealist spirit as well if, for instance, we understand it in the following manner: the intrinsic value of *X* is projected on *X* on the basis of its intrinsic properties. This observation has some weight here because Farah and Heberlein seem to object to personists that they are moral realists, a view in their minds incompatible with the projective nature of our brain module.

This presentation of moral status shows clearly that what is at stake is the *value* of the individuals, a value grounded in intrinsic properties bearing *values*. We value persons for their rationality, not for the fact they have a face; or we distinguish persons *because* we value rationality, and not because we value human faces. It is possible that the psychological origin of this process resides in our sensibility to faces, but this fact has no essential relation with our moral judgment, it is only contingent, and often misleading because the class of individuals with a human face and the class of individuals endowed with reason are not the same, even though they overlap.

4. A reply to the charge of arbitrariness

Conceptually, personhood could be the marker of moral status as the possession of interests could be (because having interests is also a bearer of values). But is it an adequate one? Farah and Heberlein do not think it is because it is arbitrary. It is their second charge.

This charge is nevertheless a little ambiguous and covers two different objections. The first is that the property of rationality is a gradual one and that we are unable to say precisely where the threshold is, and the second is that rationality covers a large number of different properties and that it is arbitrary to pick one instead of another. The authors illustrate both objections in discussing Fletcher's view: 'Joseph Fletcher proposes 15 criteria for personhood. He begins with intelligence, and makes an admirably straightforward effort to specify the dividing line between persons and non-persons by referring to intelligence quotient (IQ) scores: "Below IQ 40 individuals might not be persons; below IQ 20 they are definitely not persons." The problem with this criterion is that, while it is explicit and precise, it is also arbitrary. His other 14 "marks of personhood" include traits and capacities similar to the ones already mentioned as well as a few additions and elaborations. They are: self-awareness, self-control, sense of time, sense of futurity,

sense of the past, capacity to relate to others, concern for others, communication with other persons, control of existence, curiosity, change and changeability, balance of rationality and feeling, idiosyncrasy, and neocortical function' (2007: 38). In this section, I will examine the second charge (arbitrariness of property), and in the following the first charge (arbitrariness of graduality).

Should we possess conscience, consciousness, critical interests, capacity for language, for abstract thinking, ability to choose on the basis of reasons or to have a life plan in order to be a person? All these properties, and those mentioned by Fletcher, can be put under the umbrella of reason or rationality, but it seems possible to have one without having the others. Traditionally, reason has been put forward in contradistinction to sentience, rationality being the mark of human beings and sentience the mark of animals. But we know today that some animals possess some rational capacities listed above.

This difficulty is often raised against personism. Nevertheless, I think that it is not fatal at all, because it is grounded in a double mistake. First, as many commentators have observed, many concepts have no necessary and sufficient conditions without any consequence for the existence of their referent (think of plants and even mountains; Churchland, 2007: 54-55). Second and more importantly for our topic, in ethics and concerning the question of moral status, we do not need a definition of what a person is, that is a necessary and sufficient condition of personhood, and less a grasp of the essence of personhood (if there exists such a thing); we do not even need a necessary condition, but only a sufficient one, that is a criterion. And we have many, including the list of mental properties put forward by philosophers from Boethius on – Roskies speaks of a 'cluster-concept' (2007: 56). These criteria ought to be directly or indirectly observable (since they consist mainly in faculties, that is in dispositions to behave rather than behaviours proper). The only requirement is that they will show some degree of convergence because they are all indicators of personhood. Practically, this means that we will require that a being possesses at least one of the rational properties on the list in order to be granted the status of a person. As Blackford comments: 'Different philosophers have considered different cognitive capacities to be of metaphysical or moral importance', but that did not mean that they 'worked with differing conceptions of personhood' (2007: 70).

5. A reply to the charge of graduality

As I just said, a person must possess at least one of the rational properties on the list. But these properties are gradual; therefore, which amount of it? Where is the threshold? It seems impossible to set it non-arbitrarily and even resorting to neuroscience does not solve the problem, as Farah and Heberlein claim: 'Relevant clinical observations and neural network modeling indicate that the change in psychological capabilities would be gradual and would in general lack the kinds of qualitative transition points that could be used as non-arbitrary places to draw a line between persons and non-persons […]. Thus, for defining personhood the devil is

just as much present in the neurological details as in the psychological ones' (2007: 40). The failure of neurology to give precision to what is psychologically loose is important in Farah and Heberlein's paper, because it is natural to hope that the progress in neuroscience could remedy the vagueness of our psychological criteria. As they state: 'We believe that this empirical, neuroscience-based approach to defining personhood will eventually be successful in translating the psychological criteria discussed earlier into neurological criteria' (2007: 40); however, as we have seen, this hope has been frustrated because neuroscience does not take us further, since it stumbles over the same difficulties: we are unable to say 'which cortical systems in which combinations are critical and how much functionality is required of each of those systems'. Hence a more modest project: 'The real contribution of neuroscience to understanding personhood may be in revealing not what persons are, but rather why we have the intuition that there are persons' (2007: 40). And we know the answer: face recognition.

The difficulty of graduality stems from the requirements of ethics and, more generally, of normativity. If the psychological discourse lives very well with gradual properties, it is not the case with the normative one: we must know if an individual is a person or not in order to know how we should treat her and if she possesses rights. Or course, it is possible to set a threshold, but it will necessarily be arbitrary. On this point, Farah and Heberlein are right.

They are, but is it damaging for personism? In a sense, it is not very surprising that normative distinctions and psychological ones do not completely match. Their requirements are not the same and we know that it is not possible to pass easily from facts to norms. 'Person' is indeed a special concept: it has normative implications, but is not in itself a normative concept. Rather, it is an ontological one, saying what there is, not what there should be, but with a normative impact. In itself, it could be a gradual concept, but its function in ethics prevents that we use it as a gradual one. It could be a problem if we want to anchor ethics in our human reality – and how could we give up the idea to anchor it if we want to propose an ethics suitable for the human beings we are?

Nevertheless, these observations would compel us to give up personism only if we had a better option. Farah and Heberlein think that they have one, stating that 'rather than ask whether someone or something is a person, we should ask how much capacity exists for enjoying the kinds of psychological traits previously discussed (e.g., intelligence, self-awareness) and what are the consequent interests of that being' (2007: 46). Since they focus on interests, they describe their view as a 'more utilitarian approach', but this is a mistake, because their proposal is essentially the same as the proposal of the personists. As Glannon asks: 'For what or for whom should we maximize the good' (2007: 69), that is the satisfaction of interests, and more directly, Perring: 'Whose interests count' (2007: 68)? Moreover, when faced with objections, personists too should descend to the level of interests, even if they often use the language of capacities (the basis of many interests, like critical interests in the sense of Dworkin (1993: 218)). In a sense, Farah and Heberlein acknowledge it,

since they say just after the passage quoted above: 'Of course, this view requires deciding how these traits should be defined and ranked in importance and whether to consider a being's potential, or only actual, status. In other words, many similar problems arise as in discussions of criteria for personhood.' (2007: 46)

6. Conclusion

Why, then, all this fuss about personhood? Farah and Heberlein think that personhood is a veil that prevents us from seeing many normative relevant points, and they hope that the language of interests will not: 'Having understood the need to set aside intuitions about personhood and having avoided the distraction of seeking criteria for personhood, we can work more productively on assessing and protecting the interests of all' (2007: 46). I think that they are partially right, in the sense that their arguments strike not personism as such, but a kind of 'mystical' one.

Some personists claim that the status 'of "being a person" should not depend on whether one has or does not have certain capacities (e.g., intellectual capacities)' (Gastmans and De Lepeleire, 2010: 81). So, on what should it depend? The 'mystical' personists speak as if they believed that it depended on nothing (we should respect persons 'just as they are'); but in fact, they suggest social properties like 'belonging to human society' or 'being related to other human beings'. Sometimes they also put forward the property 'to be human' or 'to belong to the human species'. These are views very near the one depicted by Farah and Heberlein and linked more or less consistently with the recognition of the human face. However, if Farah and Heberlein are right to think that some personist theories are grounded in an unconscious and automatic brain module – a ground unsuited for a normative thesis about moral status – the main personist tradition is not. Even if there was a link between the human face and personhood, the link has been largely loosened, even severed, for a long time in non-mystical personist ethics. It was already done with Boethius, substituting a rationalist interpretation of personhood to a speciesist one (Baertschi, 2014).

I said in my introduction that utilitarian thinkers seem to hope that neuroscience will buttress their position and I added that I don't believe they have succeeded till now and that philosophical arguments still stand true when normative questions are debated. I hope that my analysis confirms this statement: determining the moral status of human beings necessarily leans on a human property, and even a psychological one. Neuroscience can help us tell what are the possible candidates and teach us why we tend to pick one rather than another; but neuroscience remains unable to tell what is *the normatively adequate one*.

References

Baertschi B. 2012. The Moral Status of Artificial Life. Environmental Values, 21: 5–18.
Baertschi B. 2014. Human Dignity as a Component of a Long-Lasting and Widespread Conceptual Construct. Journal of Bioethical Inquiry. DOI

10.1007/s11673-014-9512-9.
Blackford R. 2007. Differing Vulnerabilities: The Moral Significance of Lockean Personhood. The American Journal of Bioethics, 7(1): 70–71.
Bochenski J. 1959. A Precis of Mathematical Logic. Dordrecht: Reidel.
Churchland P. 2007. The Necessary-and-Sufficient Boondoggle. The American Journal of Bioethics, 7(1): 54–55.
Dworkin R. 1993. Life's Dominion: An Argument About Abortion, Euthanasia, and Individual Freedom. New York: Knopf.
Engelhardt HT. 1986. The Foundations of Bioethics. Oxford: Oxford University Press.
Farah M and Heberlein A. 2007. Personhood and Neuroscience: Naturalizing or Nihilating? The American Journal of Bioethics, 7(1): 37–48.
Gastmans C and De Lepeleire J. 2010. Living to the Bitter End? A Personalist Approach to Euthanasia in Persons with Severe Dementia. Bioethics, 24(2): 78–86.
Glannon W. 2007. Person, Metaphysics and Ethics. The American Journal of Bioethics, 7(1): 68–69.
Goffi J-Y. 2007. Personne, Personnalisme, Personnisme. In R. Mache. La Personne dans les Sociétés Techniques. Paris: L'Harmattan, p. 13–41.
Greene J. 2008. The Secret Joke of Kant's Soul. In W. Sinnott-Armstrong. Moral Psychology. Cambridge MA: MIT Press, p. 35–79.
Perring C. 2007. Against Scientism, For Personhood. The American Journal of Bioethics, 7(1): 67–68.
Rabinowicz W and Rønnow-Rasmussen T. 2000. A Distinction in Value: Intrinsic and for Its Own Sake. Proceedings of the Aristotelian Society, 100: 33–51.
Roskies A. 2007. The Illusion of Personhood. The American Journal of Bioethics, 7(1): 55–57.
Warren MA. 1997. Moral Status. Oxford: Oxford University Press.