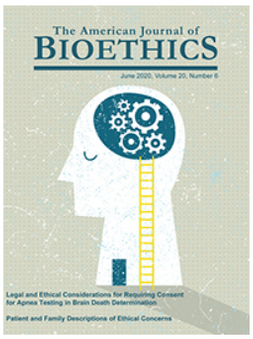

The American Journal of Bioethics

ISSN: 1526-5161 (Print) 1536-0075 (Online) Journal homepage: https://www.tandfonline.com/loi/uajb20

Beyond the Apnea Test: An Argument to Broaden the Requirement for Consent to the Entire Brain Death Evaluation

Erin Paquette, Joel Frader, Seema Shah, Robert C. Tasker & Robert Truog

To cite this article: Erin Paquette, Joel Frader, Seema Shah, Robert C. Tasker & Robert Truog (2020) Beyond the Apnea Test: An Argument to Broaden the Requirement for Consent to the Entire Brain Death Evaluation, The American Journal of Bioethics, 20:6, 17-19, DOI: 10.1080/15265161.2020.1754523

To link to this article: https://doi.org/10.1080/15265161.2020.1754523

Published online: 22 May 2020.

Submit your article to this journal

Article views: 326

View related articles

View Crossmark data

Citing articles: 5 View citing articles

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=uajb20

THE AMERICAN JOURNAL OF BIOETHICS
2020, VOL. 20, NO. 6, 17–19
https://doi.org/10.1080/15265161.2020.1754523

Taylor & Francis
Taylor & Francis Group

OPEN PEER COMMENTARIES

Check for updates

Beyond the Apnea Test: An Argument to Broaden the Requirement for Consent to the Entire Brain Death Evaluation

Erin Paquette[a,b], Joel Frader[a,b], Seema Shah[a,b,c], Robert C. Tasker[d,f], and Robert Truog[e,f]

[a]Ann & Robert H. Lurie Children's Hospital of Chicago; [b]Northwestern University Feinberg School of Medicine; [c]Stanley Manne Children's Research Institute; [d]Harvard Medical School; [e]Harvard Medical School, Global Health & Social Medicine; [f]Children's Hospital Boston, Critical Care Medicine

In their article, *Legal and Ethical Considerations Requiring Consent for Apnea Testing in Brain Death Determination*, Berkowitz and Garrett (2020) argue that informed consent for apnea testing is legally and ethically required based on two primary arguments—that the apnea test has associated risk as a medical procedure and that individuals have a right to refuse medical treatment. They highlight persistent controversies in the determination of death by neurologic criteria (DNC), including rejection of the concept of brain death, which has led to several legal challenges and proposed solutions (Shah, 2018; Veatch and Ross, 2016). Many court cases have framed legal objections to the determination of DNC in terms of consent. As Berkowitz and Garrett (2020) discuss, courts have ruled differently on whether consent is required, with some courts remaining silent on the issue and settling the case on other grounds (Leemputte and Paquette 2019).

We agree that obtaining consent to determine DNC is ethically required and—at minimal—legally permissible (Truog and Tasker 2017a, 2017b; Leemputte and Paquette, 2019). Obtaining consent for the determination of DNC respects diverse viewpoints about brain death without requiring a change in the definition of death by providing a functional mechanism through which individuals can express objection to the concept of brain death. However, to fully recognize this important motivation for requiring consent, we raise three considerations that support requiring consent for the entire examination needed to determine DNC, rather than the apnea evaluation alone. First, we argue that the ethical justification for seeking consent requires seeking consent for the entire evaluation. Second, we advance that statutory and common law legal justifications for consent also support obtaining consent for the full evaluation. Finally, we discuss the importance of not conflating the *justification for requiring consent* with the *obligations of adequate disclosure* when consent is obtained.

Narrowly framing the argument for consent in terms only of the apnea evaluation, rather than the entire evaluation for determining DNC, risks clinicians attempting to utilize ancillary evaluations to replace the apnea test, as the authors acknowledge. Berkowitz and Garrett (2020) respond to this concern by arguing that the use of ancillary evaluations to evade a consent requirement is not consistent with the spirit of existing clinical guidelines, which stipulate that ancillary testing should not be used to replace the neurologic examination unless it cannot be performed. But, if there is significant resistance to consent for the apnea evaluation, these consensus guidelines could simply change. Counter to this position, is the fact that such a manipulation would change the pretest probability of the ancillary test. That is, we simply do not know the sensitivity and specificity of an ancillary test in correctly identifying the state of DNC when performed in a population on the basis of deciding to omit the apnea test.

More problematic is that narrowly conceptualizing consent in terms of the apnea evaluation weakens the ethical basis for obtaining consent. Berkowitz and Garrett (2020) appeal to the risk associated with the apnea evaluation as an important component that justifies consent. Risk associated with the evaluation, however, does not capture the strongest ethical justification for seeking consent. Instead, a more robust ethical justification follows from recognizing the need to respect the moral status of the person and right to autonomous choice (Beauchamp and Childress 2019). In the setting of an incapacitated individual, surrogates assert this right on behalf of the incapacitated

CONATCT Erin Paquette erin.talati@northwestern.edu Department of Pediatrics, Northwestern University, Chicago, IL, USA.
© 2020 Taylor & Francis Group, LLC

person. Whether reflected through the individual directly or through a surrogate, respect for moral status requires that the individual (or surrogate) expresses a choice over what is done with his/her body. Full recognition of this status requires consent to perform any component of an examination, including the entire clinical determination of DNC.

Some might argue that consent is not necessary for the clinical determination of DNC, specifically under this framework, because it is an examination to which an individual agrees when they provide consent for general treatment (Beauchamp and Childress 2013). However, this view ignores that another essential element of consent includes the right to withdraw that consent at any time. Thus, even if consent were provided for examination on a general consent form, at any point, the patient could decline an examination be performed.

Alternatively, as Berkowitz and Garrett (2020) put forth, others might suggest that a diagnosis of respiratory failure does not require consent so that a determination of DNC should not. Yet, while a diagnosis of respiratory failure does not require consent, a physical examination of the lungs as part of such a diagnosis does require consent. Consent may be implied if the patient does not reject the evaluation, but if the patient or their surrogate opposed the examination, ethically, the physician could not forcibly perform the examination to make the diagnosis. Likewise, the not-yet-dead patient—who retains moral status prior to the determination of DNC—should not be denied the opportunity to refuse the entire clinical evaluation or any of its component parts.

Legal justifications for seeking consent further support a broader view of the consent requirement. The Uniform Determination of Death Act (UDDA 1981), as it is adopted in the majority of states, requires that death be declared in accordance with "accepted medical standards." This is interpreted to reflect the medical guidelines describing the various components of the neurologic examination that comprise the evaluation for DNC. But, as described above, any physical examination in usual medical practice requires the patient's consent. Thus, the statutory requirement for performing the examination in accordance with "accepted medical practice" implies that consent should be obtained as it would be in any other medical encounter.

Common law evolution of the concept of informed consent also provides strong justification for obtaining consent for the entire examination needed in the determination of DNC, rather than the apnea evaluation in particular. The earliest cases describing informed consent directly contemplated whether a medical examination could be performed above an individual's objection. In refusing to compel an examination to determine the extent of a work-related injury, the Court in *Union Pacific v Botsford* (1891) described the "right of every individual to the possession and control of his own person, free from all restraint or interference of others." In *Mohr v. Williams* (1905) the Court found that "any unauthorized touching of the person of another … constitutes an assault and battery," essentially establishing a requirement for consent for any examination performed on the body. Finally, the landmark case of *Schloendorff v. Society of New York Hospital* (1914) aligned the legal basis for consent with its ethical foundations in autonomy, grounding consent in respect for the body, and holding that, "every human being of adult years and sound mind has a right to determine what shall be done with his own body." Together, these cases create a strong legal foundation for requiring consent for the entire evaluation leading to the determination of DNC.

We have argued that both the ethical and legal justifications for obtaining consent for the evaluation leading to the determination of DNC require consent beyond the apnea evaluation. This distinction is critical for two reasons. First, it is important to appeal to the ethical and legal foundations for consent. The need for consent does not arise because a procedure, examination, or intervention is risky. Rather, the need arises because of a duty to respect autonomous choice of a moral agent, either through the individual directly or his/her surrogate.

Second, the clinician's disclosure obligations follow from the reasons to seek consent in the first place. Limiting consent to the apnea evaluation would limit the obligations around disclosure to the elements of the risks associated with this evaluation. The clinician may not feel compelled to disclose the consequences of the remainder of the examination, including the implications of a determination of DNC, which should be a part of the standard disclosure to all patients/surrogates when a diagnosis of DNC is suspected.

This is not to diminish the significant risks of the apnea evaluation. Indeed, they are particularly relevant because exposure to these risks can induce the state that the examiner is looking to find. As Berkowitz and Garrett (2020) describe, the rise in partial pressure of arterial blood carbon dioxide that occurs with apnea can, in theory, lead to an increase in intracranial pressure, placing an already injured brain at greater risk of swelling and progression to DNC, if it were not present before. As one's moral agency ends at death, exposure

to the apnea test runs the risk of creating a situation where the individual loses the ability to express any preferences through his/her surrogate.

Narrowly defining the need for consent in terms of the apnea evaluation alone would remove the obligation to obtain consent in situations in which the examination truly could not be performed due to patient instability. Even if the patient's surrogate would object to the remainder of the examination, if consent were only required for the apnea evaluation, the remainder of the examination could arguably proceed without consent and a declaration of DNC could be made with the aid of ancillary testing. This end run around consent requirements does not respect the moral agency of the patient, which holds no weight after a declaration of death has been made. Undermining requirements for consent in highly contentious, value-laden areas of medical care risks further damage to already tenuous trust between clinicians and many in the public.

Resolution of the complex challenges to the determination of DNC will require reassessing many of our current practices and beliefs. Seeking consent to perform the evaluation provides a mechanism to respect dissent to the concept of DNC without requiring a change to the definition of death, which could have implications in other situations. We agree with Berkowitz and Garrett (2020) that there are strong ethical and legal reasons to require consent for the examination that may lead to the determination of DNC. However, to respect these ethical and legal foundations, consent should be obtained for the entire examination and not unique to the apnea evaluation.

DISCLOSURE STATEMENT

Dr. Truog serves as a paid consultant on data safety monitoring committees for Sanofi and Covance.

REFERENCES

Beauchamp, T. L., and J. F. Childress. Respect for autonomy. In: *Principles of Biomedical Ethics*. 7th ed., 101–149. New York: Oxford University Press; 2013.

Beauchamp, T. L., and J. F. Childress. 2019. *Principles of Biomedical Ethics*, 8th ed. New York: Oxford University Press.

Berkowitz, I., and J. Garrett. 2020. Legal and ethical considerations requiring consent for apnea testing in brain death determination. *The American Journal of Bioethics* 20(6): 4–16.

Leemputte, M., and E. Paquette. 2019. Consent for conducting evaluations to determine death by neurologic criteria: A legally permissible and ethically required approach to addressing current controversies. *Curr Pediatr Rep* 7: 152–162.

Mohr v. Williams. 1905, 95 Minn. 261, 104 N.W. 12.

Schloendorff v. Society of New York Hospital. 1914. 105 N.E. 92.

Shah, S. 2018. Rethinking brain death as a legal fiction: Is the terminology the problem. Defining death: Organ transplantation and the fifty-year legacy of the harvard report on brain death, special report. *Hastings Center Report* 48(6): S49–S52. doi: 10.1002/hast.955.

Truog, R. D., and R. C. Tasker. 2017a. Counterpoint: Should informed consent be required for apnea testing in patients with suspected brain death? Yes. *Chest* 152(4): 702–704. doi: 10.1016/j.chest.2017.05.032.

Truog, R. D., and R. C. Tasker. 2017b. Rebuttal from Drs Truog and Tasker. *Chest* 152(4): 705–706. doi: 10.1016/j.chest.2017.05.033.

UDDA. 1981. *President's commission for the study of ethical problems in mediine and biomedical and behavioral research. Defining death: A report on the medical, legal and ethical issues in the determination of death*. Washington, D.C.: U.S. Government Printing Office.

Union Pacific Railway Co. v. Botsford. 1891. 141 U.S. 250.

Veatch, R., and L. Ross. 2016. *Defining death: The case for choice*. Washington, DC: Georgetown University Press. Response to key challenge to DNC.

Philosophers' Imprint

VOLUME 17, NO. 20
OCTOBER 2017

Death and Consensus Liberalism

Jeremy Williams

University of Birmingham

© 2017 Jeremy Williams
This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 3.0 License.
<www.philosophersimprint.org/017020/>

1. Introduction

Theories of public reason propose moral constraints on political action. They have, then, an inherently practical focus. Each such theory articulates a test or standard of public justification which is purported to show, roughly speaking, when political institutions, laws, or policies have been legitimately established, and when they ought to be abolished, repealed, or reformed. Depending on the details of the test which a given theory proposes, the upheavals it requires to political practice could be extremely radical—and not necessarily welcome. The full evaluation of any theory of public reason must therefore depend on whether the results of applying it are sufficiently plausible or palatable. Rawls acknowledged this point explicitly. After developing his seminal idea and ideal of public reason in *Political Liberalism*, he concluded that 'whether this or some other understanding of public reason is acceptable can be decided only by examining the answers it leads to over a wide range of the more likely cases.'[1]

Despite Rawls's counsel, however, philosophical investigation into the implications of public reasoning for concrete political questions remains surprisingly rare. Most discussion of public reason is pitched at a fairly high level of abstraction, consisting of reflection on whether, in general outline, particular accounts are coherent or compelling, and of analysis of the best ways to specify their various theoretical components (such as the conception of a reasonable person or that of a justificatory reason). To the extent that proponents of public reason consider in depth what their theories portend for the resolution of particular policy problems, their approach tends to be to handpick issues as illustrative case studies, on which their theory is taken to provide clear, attractive guidance. In this paper, my aim is to contribute to remedying this general omission of the literature. I do so by drawing out some hitherto unexplored practical implications of public reason under the Rawlsian conception, which remains, in spite of fierce competition, the dominant brand of public reason liberalism on the market.

1. Rawls 2005: 254.

The Rawlsian conception provides the archetype of the so-called *consensus* model of public reason. Consensus theories are distinguished by their restriction of the reasons that may be invoked in justifying exercises of political power to those that can be recognised as in some sense good or acceptable grounds for political action by all qualified members of the community. On the Rawlsian version of the consensus model, political decisions must be justified, more specifically, with reference to the moral reasons given by only the limited set of political values and concepts that all citizens will share if they are (by Rawlsian lights) reasonable. This is not to say that the Rawlsian view altogether disallows appeal to controversial reasons. For it permits justification on the basis of competing views about the interpretation, applicability, and relative importance of citizens' shared values, as adduced from within their reasonable 'political conceptions of justice'.[2] The theory does, however, significantly constrain political disagreement by specifying that reasons drawn from reasonable citizens' religions, metaphysics, conceptions of the good, and other aspects of their 'comprehensive doctrines' cannot count as justificatory. And it imposes upon all individuals who share in the exercise of the state's coercive power, including ordinary voters, a moral 'duty of civility' to advocate and lend support to only political initiatives that are publicly justified according to the foregoing standards.[3]

2. A political conception of justice is, roughly, one that takes account of only those moral values that constitute shared political values, and that tries to specify and order these in a sufficiently precise way to provide determinate answers to political questions. See Rawls 2005 at, e.g., 386.

3. See Rawls 2005 at, e.g., 217ff. On the so-called 'wide view' of public reason, which Rawls came to endorse in his final works on the subject, the duty of civility is somewhat relaxed: citizens are permitted to invoke their comprehensive doctrines in public argument, subject to the 'proviso' that a case in public reason can be produced 'in due course' that leads to the same conclusion they seek to defend (Rawls 2005: 462). The wide view is somewhat controversial among Rawlsians (for a rejection of it, see Hartley and Watson 2009). But because it does not give decision-makers any latitude to depart from the policy prescriptions that would be issued by public reason alone, it will make no difference to the argument of this paper whether or not it is assumed to be in force.

Any proposal that important decisions be taken on the basis of an artificially restricted set of reasons is likely to invite concern that the ability of political agents to reason and choose well will be unduly inhibited. Accordingly, one crucial test for Rawlsian public reason is whether it is, as Rawls himself puts it, *complete*. Rawls defined completeness as a matter of whether the content of public reason — that is, the total set of ideas, arguments, and principles from which public justifications are to be composed — is sufficient to generate 'a reasonable answer to all, or nearly all' the political questions for which the use of public reason is required.[4] He hypothesised that his conception of public reason is indeed complete in this sense, though he did not attempt to *show* it, and his defenders have not, I believe, met that challenge either.[5] The reasons for this are perhaps understandable.

4. See Rawls 2005 at, e.g., 244ff or 454. Rawls further advises (at 246) that whether an answer counts as reasonable for purposes of evaluating the completeness of public reason is to be 'judged by public reason alone'. In other words, public reason is not to be understood as incomplete just because it generates answers to which we object from our comprehensive or all-things-considered moral outlook. Rather, completeness is undermined only by cases in which the public reasons available for addressing a political question are so sparse that either no answer can be reached, or we can produce only answers that are disqualified as incompatible with relevant political values or principles of public reason itself. Interestingly, meanwhile, in *The Law of Peoples*, one of Rawls's comments might be taken to suggest, differently, that a model of public reason should be accounted incomplete not merely in the event that it fails to orient political decision-making, but also in the event that we cannot reconcile ourselves to its guidance in reflective equilibrium. For discussion, see Williams 2016: 3–4. At certain points below I will have occasion to discuss the possibility of public reasoning's leading us to policy conclusions that are defective in the latter sense. But to forestall confusion, I shall throughout use the term 'completeness' in only the first sense given above (the sense also standardly attributed to it in the secondary literature), on which it concerns, narrowly, whether public reason enables deliberators to argue their way to any resolution to the political problems put before them at all.

5. To be sure, there exist defences of the completeness of Rawlsian public reason. But their general strategy is to argue (a) that the (hitherto unmet) burden of proof lies with those who dispute public reason's completeness to make their case, and (b) that insofar as public reason fails to provide a basis for decision-making, there are nonetheless ways for citizens to select between the policy options before them that do not involve resorting to non-public

Demonstrating that a conception of public reason is complete seems a highly daunting — if not Sisyphean — task, involving delving into the minutiae of a vast number of political problems, to establish, in each case, what answers can be justified by public reasons alone. Showing that a conception of public reason is *in*complete, on the other hand, may be a more manageable undertaking. If a sufficient number of examples can be found in which public reason proves inimical to forming at least one reasonable conclusion, then the charge of incompleteness is substantiated without the need to exhaustively catalogue the outcomes of public reasoning for other questions. Yet, while doubts about the incompleteness of Rawlsian public reason have been often voiced, critics have thus far carried out relatively little of the necessary philosophical spadework.[6] The question of the completeness of Rawlsian public reason therefore remains crucially unsettled.

In speculating about the political issues that would be most likely to make revealing test cases for the incompleteness objection, commentators have typically alighted on the field of *bioethics*. For it contains many questions that appear to turn on precisely the sort of deep, longstanding philosophical debates that public reason requires citizens to put to one side. The question most commonly identified in the literature as raising the spectre of incompleteness is that of abortion. Having considered the implications of Rawlsian public reason for that controversy elsewhere,[7] however, in this paper I address an important — and thus far widely overlooked — bioethical problem that arises at the other end of human life. This is the problem of how to define and diagnose the death of a person, or determine at what point the clinical and legal practices conventionally associated with death, such as the removal of vital organs, may take place. My thesis will be that this is a matter on which public reason does indeed have a grave incompleteness problem. Public reason is indeterminate, I aim to demonstrate, between a broad range of legal definitions of death (at least bracketing the socially contingent effects which candidate policies might have on third parties).[8] I also aim to go beyond existing articulations of the incompleteness objection, moreover, by examining what the Rawlsian view implies about how decision-makers ought to respond to indeterminacies in public reason. Insofar as the route to a reasoned choice between competing criteria of death is indeed foreclosed, I shall contend, public reason requires that selection among policy options proceed in an unacceptably arbitrary fashion.

Before I begin, three clarifications about the scope of my argument and conclusions. First, to reiterate, my target is the Rawlsian view (and hence, in what follows, the terms 'public reason' and 'consensus liberalism' always refer to the Rawlsian versions thereof, unless otherwise specified). But who, for present purposes, counts as a Rawlsian? I take my critique to apply to consensus theorists who believe that public justification must proceed on the basis of the reasons given by political values and concepts shared by reasonable citizens, and who follow Rawls in identifying who the reasonable are. As I understand their respective positions, Jonathan Quong and Andrew Lister both fall into this category, for instance, despite their various innovations on Rawls's original theory.[9]

reasoning. See especially Williams 2000 and Schwartzman 2004. I examine claim (b) in section 8, below.

6. The most detailed attempt in the earlier literature to advance the incompleteness objection through sustained analysis of particular political controversies appears in Greenawalt 1988: chs. 6–8. Greenawalt's principal examples of alleged incompleteness are the problems of abortion and animal welfare. His argument predates *Political Liberalism*, and thus does not respond to the mature version of Rawlsian public reason, as developed there and in subsequent essays (though for a new argument to comparable effect, also focusing on abortion, and published only after the present article was completed, see Kramer 2017: ch. 3. I comment on Kramer's argument further in n. 47 and n. 76, below.) I have argued previously, moreover, that in the case of the problems Greenawalt cites, the charge to which Rawlsian public reason is vulnerable is not incompleteness but something else. See Williams 2015. Especially if I was right, the incompleteness objection still stands in need of substantiation of the kind I aim to provide here.

7. In Williams 2015.

8. Note that 'indeterminate' is a term of art within the public reason literature, on the precise meaning of which see the text around n. 26, below.

9. See Quong 2011 and Lister 2013.

Second, the case I make here against the Rawlsian view, thus defined, is admittedly pro tanto. I believe that the results of applying the model to the problem of death are sufficiently unwelcome that one would be warranted in abandoning it on the basis of my argument here alone. But I shall not attempt to convince the committed Rawlsian who believes that the difficulties identified do not outweigh the various merits of their theory. This is, then, a contribution to a wider critique of consensus liberalism's consequences for political practice. I reflect further on the implications of my findings for the future of consensus liberalism in the paper's conclusion.

Third and finally, the argument of this paper, if sound, might be taken to provide indirect support not only to the various strands of comprehensive or ethical liberalism (which is where, for what it's worth, my own loyalties lie), but also to consensus liberalism's emerging competitor within the public reason fold: the innovative recent *convergence* liberal view.[10] On convergence liberalism, few restrictions are imposed on the reasons by reference to which citizens may evaluate the case for political action (beyond, principally, the requirement that those reasons be *intelligible*). But the view holds, demandingly, that a law or policy is only publicly justified, and permissibly imposed, if there are from each reasonable perspective sufficient grounds to endorse it, or not to veto it.[11] Whether convergence liberalism does indeed derive comparative advantage from my argument will depend, I think, on whether it too runs afoul of objections that target its problematic implications in practice. I suspect that it does. But this is a matter for another day.

2. Determining death: the political problem

It will help to preface my argument with some background regarding how the conceptualisation and clinical determination of death arose as a matter of political concern.[12] Prior to around the mid-twentieth century, the limitations of medical science were such that the irreversible loss of heart and lung function was inevitably swiftly followed by complete loss of neurological functioning, and vice versa. Thus, there was no apparent reason for dissatisfaction with the traditional *cardiopulmonary* criterion of death, under which, if the patient's heartbeat and breathing ceased and could not be restarted, (s)he was declared dead. The advent of modern respirators and other medical technologies from the 1950s onwards, however, made it possible to indefinitely sustain the cardiopulmonary functioning of patients whose heart and lungs did not work independently, even in the face of permanent loss of consciousness or of brain activity generally. The practice of continuing treatment to patients under such conditions seemed, in the eyes of many observers, to involve an objectionable squandering of scarce resources and facilities, including not only medicines, hospital beds, and machinery, but organs which, given new transplantation techniques, could provide others with immense benefits — especially if taken from a heart-beating donor. Debate thus ensued, within and beyond the medical community, over whether the medico-legal understanding of death might be amended so as to facilitate more timely organ procurement, withdrawal of life support, and so forth, while protecting physicians from accusations of misconduct, or indeed murder.

The result of that debate was a widespread legal shift around the world from the 1960s, away from exclusive reliance on the cardiopulmonary criterion, and towards recognition of the idea of *brain death*. Brain death is standardly defined as the irretrievable cessation of functioning of the brain as a whole, including both the 'higher brain', in which consciousness is generated, and the 'lower brain', or brainstem, which is responsible, inter alia, for controlling autonomic bodily functions and reflexes such as respiration, heartbeat, blood pressure, and vomiting. While a large number of states have enshrined brain death in law, however, it remains controversial. Opposition comes primarily

10. I am grateful to an anonymous referee for suggesting that I acknowledge this possibility.

11. The chief architects of the convergence view are, of course, Gerald Gaus and Kevin Vallier. See especially Gaus 2011 and Vallier 2014.

12. For a longer, informative account, see DeGrazia 2005: 115–24.

from two sources: from those who, for various moral, philosophical, and religious reasons, support a return to the cardiopulmonary standard, and from those who favour adoption of a more radical *higher brain death* criterion. On the latter, death occurs upon the permanent loss of function of those regions of the brain in which the capacity for consciousness is realised, even if, because the brainstem survives, the patient's somatic functions continue spontaneously, without the need for medical assistance beyond basic intravenous hydration and nutrition.

The continued controversy over the correct understanding of death has resulted in legal clashes, on both sides of the Atlantic, in which — for instance — parents have resisted attempts to disconnect their brain-dead children from ventilators, or remove their bodies to the mortuary, on grounds of their (typically religious) conviction that they were still alive.[13] The practical significance of the choice between criteria of death is not, moreover, confined to the medical sphere. For in addition to deciding when physicians ought to be permitted to remove life support or organs from their patients, and when a human body may be autopsied and disposed of, we also need to know, say, under what conditions the crime of murder or manslaughter has taken place, when individuals and corporations may be sued for wrongful death, when the posthumous confiscation or reallocation of a person's property may take place, and when to change a surviving spouse's marital status to widowed.

These are *political* matters, for which the decisions reached will be backed up by the state's coercive power. Indeed, most, if not all, are *fundamental* political matters, in the Rawlsian sense that they represent, or are inextricably bound up with, so-called 'constitutional essentials' and 'questions of basic justice'. The latter fact is significant, because it means that even if — as Rawls himself proposed — the use of public reason is mandatory only when fundamental matters are at stake, public reason will inevitably be called upon to resolve the problem of how death is to be legally construed.[14] Our question is whether it has the resources to do so.

13. For a description of two recent such cases, see Brierley 2015.

3. Death and personal identity

We have immediate grounds to suspect that it does not. For what our death consists in, and the conditions under which it occurs, appears to be a *metaphysical* question. At first sight, the statement that someone has died seems equivalent to the statement that she has ceased to exist. For the dead are, as we say, no more. (Let us accept this claim for now; I return to it in section 5.) To know when someone ceases to exist, we need to know what is involved in her — and our — *continuing* to exist over time. And that in turn depends on our fundamental nature, or what kind of entities or substances we essentially are. The truth about our fundamental nature and persistence conditions is what theories of *personal identity* seek to establish. There is, however, no such theory that is or could stably remain non-contentious among Rawlsian reasonable citizens.

For illustrative purposes, consider just three of the most prominent (secular) families of views about personal identity in contemporary metaphysics.[15] First, on so-called *biological* or *animalist* accounts, each of us is essentially a human animal or organism, whose persistence over time consists in the continued functioning of the body as an integrated unit (or perhaps in its performing certain specified critical functions). Second, under *mind essentialism*, we are instead fundamentally minds, or beings with the capacity for consciousness, who are distinct from our bodies or organisms (albeit closely related to and dependent upon them), and whose existence over time consists in the continued

14. For Rawls's view that the duty to employ public reason applies only when addressing fundamental political questions, see Rawls 2005 at, e.g., 214–15. For the view that the duty applies in political justification generally, see, e.g., Quong 2011: ch. 9.

15. The literature on personal identity — even as restricted to the three canvassed views — is too vast to survey here. For three of the most influential proponents of these particular approaches, however, see, respectively, Olson 1997, McMahan 2002: ch. 1, and Parfit 1987: part III.

functioning of those regions of the brain responsible for generating conscious mental states. Third, on *psychological* approaches, our essential nature is not that of a merely conscious subject but of a more complex psychological being, whose existence over time requires a certain minimum degree of continuity in the contents of our mental lives, such as our memories, beliefs, intentions, and so forth. There are no obvious grounds for thinking that a reasonable citizen could not endorse any of these positions.[16] For reasonableness on the consensus liberal understanding is, in a nutshell, a matter of subscribing to the set of core normative beliefs which Rawlsians take to constitute the foundational commitments of a democratic society. These are the belief that one's fellow citizens are free and equal in their moral standing, that society should be organised as a fair scheme of cooperation among them, and — most controversially — that, owing to the existence of the so-called 'burdens of judgement', and consequent 'fact of reasonable pluralism', all should practice reciprocal restraint in public justification.[17] Each of the foregoing metaphysical views seems fully compatible with these political commitments. And, crucially, those views imply — or can be developed in ways that imply — strikingly different conclusions about the conditions under which we die.

Animalism, for instance, has been variously interpreted as compatible with the idea of brain death, and as ruling it out and requiring a return to the traditional cardiopulmonary criterion.[18] What is at stake in this debate is whether, absent the survival of the brain, the residual somatic functioning of which a human organism on artificial life support can be capable is sufficient for it to be considered alive.[19]

16. I explore in the next section whether there are any other less obvious reasons why they could not do so, arising out of the minutiae of the political conception of the person which the reasonable must endorse.

17. See Rawls 2005: 48–66.

18. Or some variant thereof. For an 'updated' cardiopulmonary standard, see DeGrazia 2005: 147–49.

19. For differing perspectives, see, e.g., DeGrazia 2005: 142–49, Bernat 2006, and President's Council on Bioethics 2008: ch. 4.

Animalist proponents of brain death typically claim that an organism whose brain is destroyed cannot function as a sufficiently integrated unit to be deemed alive, or that it is incapable of functions that are conceptually basic to biological life. For opponents of neurological criteria, however, it is absurd to claim that an organism that remains capable — as some brain-dead patients are — of such complex, coordinated activities as growth, sexual maturation, fighting infection, or gestating a fetus, is dead.

Although many animalists have thought that we can die by suffering complete brain failure, few if any would say that the irreversible loss of the capacity for consciousness alone is sufficient for death. Rather, on the standard animalist account, an individual whose higher brain is destroyed, but whose brainstem continues to mediate the autonomic functions of the body, remains alive, but enters a *persistent vegetative state*. Mind essentialists, by contrast, claim that such a person *is* dead, and hence endorse the higher brain criterion of death. On their view, the circumstances of one's death might be such that one leaves behind an organism that continues to function in various ways, with or without mechanical assistance. But since we are not fundamentally organisms, they contend, but minds that exist in some form of association *with* our organisms, we should think of the nonconscious animal that persists after higher brain death merely as a person's discarded vehicle, or living corpse.[20]

Finally, consider the psychological approach. Some of its proponents have argued that, like mind essentialism, it supports the higher-brain-death criterion.[21] But, as others have argued, some versions of this approach appear to yield a still more radical — if not rather unsettling — understanding of death, whereby one of us might cease to exist even *prior to* the permanent cessation of consciousness.[22] This is possible because the psychological approach holds that we cease to exist

20. See especially McMahan 2002: 423–55.

21. See, most famously, Green and Wikler 1980.

22. See, e.g., McMahan 2002: 43–55, and DeGrazia 2005: 127 et circa.

when the level of psychological continuity required for diachronic personal identity has broken down, for which loss of consciousness, though sufficient, is not strictly necessary. The psychological account implies, for instance, that in a science fiction scenario in which our memories and other psychological features are completely erased by some machine, we cease to exist, even if a conscious subject persists throughout the process. And outside the realm of science fiction, some psychological theories might imply that certain forms of dementia, whether brought on abruptly by injury, or progressively by disease, involve sufficiently dramatic erosion of our psychological capacities and characteristics as to be incompatible with our survival.

The precise implications of a psychological theory regarding when we cease to be depend upon its details. If a sufficiently weak degree of psychological connectedness is held to be enough for identity — or any degree at all — then the theory's implications for the point at which we cease to exist may be indistinguishable in all real-world cases from those of mind essentialism. Many psychological theories, however, hold that the psychological connectedness required for identity over time is of a more demanding level, of which only a *person* — in the Lockean sense of a self-conscious, thinking being — is capable. These theories hold, by implication, that our fundamental nature is that of a person in the foregoing sense. Those who endorse this view might be thought to be constrained to accept that death for us occurs immediately upon the loss of the higher cognitive endowments that make us Lockean persons. Yet, this is not necessarily so. For such a psychological theorist might think that, if our cognitive capacities are diminished below the level required for Lockean personhood, and we correspondingly dip below the threshold of psychological connectedness needed for identity, we do not cease to exist all at once, but fade out of existence gradually, as the remaining vestiges of our mental lives are extinguished — a process that only terminates at or around the final cessation of consciousness.[23] Nonetheless, if a psychological account holds that our existence cannot be a matter of degree, and depends on a level of psychological continuity which only Lockean persons can possess, then it does indeed seem committed to the conclusion that there is a chance of our ceasing to exist in a way that leaves a subject of basic consciousness behind for a non-negligible period of time. And while critics typically maintain that it is an embarrassment to the psychological approach insofar as it is thus committed, there still seems nothing unreasonable, given the understanding of 'reasonableness' advanced above, in a proponent of the psychological account's accepting or positively embracing this conclusion.

The foregoing discussion, while based on only a small sample of relevant metaphysical theories, indicates the existence of a striking degree of reasonable disagreement over the conditions under which we die, stemming in turn from reasonable pluralism over our fundamental nature and persistence conditions. To be sure, the scope of this disagreement is not unlimited. For all agree in particular (or so I shall assume) that, if a person suffers irreversible failure of cardiopulmonary function, causing the disintegration of his brain and body, death has occurred.[24] The question is whether we *must* await cardiopulmonary failure before pronouncing the patient dead, if some prior neurological standard has already been satisfied. For the state to require that we wait for the satisfaction of a later standard is for it to coercively restrain those who would perform the various death-related activities earlier, and who will in many cases think that delay is not merely a mark of suboptimal public policy, but a threefold betrayal: of the family, whose grief is pointlessly prolonged; of those in desperate need of the patient's organs or other resources; and of the memory, values, dignity in death, and so on, of the patient him or herself.[25] Conversely,

23. Cf. Parfit 1987: 323.

24. I discuss further limits on reasonable disagreement about death in the next section.

25. As an illustration of such sentiments, consider for instance the inscription on the grave of Nancy Cruzan, whose family engaged in a high-profile battle in the U.S. court system to have her life support discontinued after she fell into an irreversibly nonconscious state: 'DEPARTED JAN 11, 1983/AT PEACE DEC 26, 1990'. Cited in McMahan 2002: 423.

for the state to endorse an earlier standard requires restraint of those whose moral convictions still direct them to treat the patient as a living person, and for whom a premature declaration of death will be taken to evince, primarily, an abominable disregard for the latter's still-operative basic rights. Under consensus liberalism, what is crucial is that those who stand to be coerced in these ways can, no matter how vehemently they dissent from the law on death, nonetheless be said to have received a proper public justification for it.

Such justification will not be possible, however, if, in order to reason one's way to a conclusion about how death should be legally defined, one has no choice but to take sides, explicitly or implicitly, between reasonably rejectable understandings of our essence and identity. Instead, insofar as the required deliberative route to a policy conclusion is blocked by the Rawlsian requirement of neutrality between reasonable metaphysical doctrines, the determination of death will be a question on which the rules of public reasoning produce *indeterminacy*. I use the term 'indeterminacy' here in a technical sense attributable to Gerald Gaus.[26] Public reason is indeterminate in this sense when the considerations to which it permits appeal fail to provide deliberators with sufficient warrant to choose one way or another between the options on the table. Indeterminacy so defined is to be distinguished from what Gaus calls *inconclusiveness*, which occurs when the admissible reasons enable decision-makers to reach multiple competing conclusions, but no further public reasons can be adduced that would facilitate agreement over which is best or most reasonable. Put in further Gausian terms, inconclusiveness occurs when two or more options have public justifications that are neither *defeated* (that is, refuted by some publicly eligible reason) nor *victorious* (proven beyond reasonable doubt).[27] In cases of indeterminacy, by contrast, we cannot even get *that* far: whatever political conclusions our full, comprehensive perspective might have enabled us to draw, the *public* reasons on hand do not add up to 'the minimum degree of proof required for either justified acceptance or rejection' of any relevant policy alternative.[28]

Consensus liberals have sometimes argued that inconclusive justification is an endemic feature of political life, which a polity can accommodate without abandoning the ideal of public reason.[29] Insofar as public reason is found to be indeterminate, however, I believe consensus liberalism faces a more serious challenge. This is because, as we shall see in more detail later, in cases of inconclusiveness it is consistent with Rawlsian values for us to select between policy options via the familiar devices of democratic politics, such as majority voting. But should public reason prove indeterminate, the deadlock will be breakable only by resort to rather more unusual and unappealing procedural mechanisms. Before considering, however, how damaging it would be to public reason should it prove indeterminate on the matter of defining death, our more immediate task is to confirm whether the appearance of indeterminacy observed so far is confirmed on further inspection. We need to confirm, in other words, whether there is any viable, non-metaphysical form of reasoning about death that is generally available to Rawlsian deliberators. Over the next four sections, I will argue that there is not. The terms of citizens' duty of civility, we shall see, prohibit them from publicly invoking, or factoring into their decision-making, precisely the considerations needed if they are to reliably reach even inconclusively justified verdicts in this complex and morally fraught policy area.

4. Death and the political conception of the person

The natural place to begin our inquiry is by asking whether a conclusion about the determination of death could be derived from the Rawlsian 'political conception of the person' (hereinafter 'PCP'). For the purpose of the PCP is precisely to fulfil the role in democratic deliberation

26. See Gaus 1996: 151–58. For informative further discussion of the distinction, see Schwartzman 2004: 193–98.

27. Gaus 1996 at, e.g., 151.

28. Gaus 1996: 153.

29. See especially Schwartzman 2004.

that comprehensive conceptions of the person typically perform in ordinary moral reasoning.

Let us first take stock of the PCP's main features. Like the other political concepts and values on which public reasoning depends, Rawls presents the PCP as one of the 'fundamental ideas' that characterise the tradition of democratic thought, and are latent within the 'public political culture' of a democratic society.[30] It represents, he thinks, the distinctive way in which democratic citizens view themselves and their peers. Indeed, the PCP understands a person *in terms of* citizenship: as an individual who can take part in public life in virtue of her possession, to a sufficient degree, of certain cognitive capacities and moral sensibilities — namely, the 'moral powers' of rationality and reasonableness.

Rawls sometimes formulates the PCP in such a way as to imply that, unless an individual possesses the relevant endowments at a given time, she is not a person at that time. For instance, he writes that 'we think of persons as rational and reasonable, as free and equal citizens, with the two moral powers and having, at any given moment, a determinate conception of the good, which may change over time.'[31] When the PCP is understood in this way, many human beings, such as children, or those who were once cooperators, but whose mental capacities are now too diminished, do not qualify. And a requirement that public reasoning be informed by this version of the PCP would accordingly seem at serious risk of generating a raft of unpalatable conclusions concerning the rights and permissible treatment of those excluded. On other occasions, however, Rawls observes a more capacious understanding of the person at work within the democratic tradition. On the latter, 'we say that a person is someone who can be a citizen, that is, a normal and fully cooperating member of society over a complete life.'[32] In allowing that a person need only fulfil the role of citizen over the course of a complete life, this formulation of the PCP draws in members of society who do not yet possess the requisite capacities but will, and those who no longer possess them but did.[33] Call this the inclusive PCP, in contrast to the exclusionary variant identified at the top of this paragraph.

As I understand it, it is the inclusive PCP that consensus liberalism accepts, and requires reasonable citizens to endorse. Were this not so, Rawls could not coherently claim, for instance, as he does, that children are equal beneficiaries of justice from any reasonable perspective.[34] Moreover, Rawls does not merely stipulate that the more inclusive formulation applies — he justifies consensus liberalism's adoption of it on the ground that it is needed to 'go with' the democratic conception of society.[35] On the latter, society is understood as a collective enterprise of a certain scope. It is not a mere 'association' of the kind that one is free to join or leave at will once one reaches 'the age of reason', thereby acquiring or divesting oneself of its package of rights and obligations.[36] Rather, a democratic society is 'a more or less complete and self-sufficient scheme of cooperation, making room within itself for all the necessities and activities of life, from birth until death.'[37] 'We add the phrase "over a complete life"' when specifying the extent to which persons must be able to participate in public life, Rawls tells

30. Rawls 2005: 14 and 29–35.

31. Rawls 2005: 481–82 (emphasis added, footnote deleted).

32. Rawls 2005: 18 (emphasis added).

33. To be sure, it continues to exclude those permanently incapable of participation in social life. Yet while I believe this is a matter of concern, it falls outside the aims of this paper to consider the consequences of that residual exclusion for public reasoning here.

34. Rawls 2005: 474.

35. Rawls 2005: 18.

36. Rawls 2005: 41 et circa.

37. Rawls 2005: 18. In addition to being a 'complete' social system in the foregoing sense, Rawls suggests here and elsewhere that, for the purpose of developing a political conception of justice, it is appropriate to model a democratic society as 'closed' — that is, without inward or outward migration — such that entry and exit are by birth and death *only*. But the stipulation of closedness, he stresses (at p. 12), can only be a temporary theoretical convenience (unlike, assumedly, the characterisation of a democratic society as complete).

us, to reflect this fact about the bounds of the societal relationship.[38] Rawls's view, thus, appears to be that the inclusive PCP is the more faithful rendering of the conception of ourselves presupposed by the distinctive democratic mode of societal organisation.

Accordingly, in what follows my argument will be predicated, unless otherwise indicated, on the assumption that it is the inclusive PCP that is found among reasonable people's stipulated shared commitments, and hence lies within the content of public reason. I shall, however, consider in the paper's concluding section the possibility of amending the PCP to the exclusionary formulation as a means of combatting the indeterminacy problem I am in process of outlining. Suffice it to say for now that, given the implications of adopting the exclusionary PCP, the assumption that consensus liberalism subscribes instead to its inclusive sibling is not unfavourable to the theory.

Because it includes the idea of a citizen's leading a complete life, without specifying a point of terminus, it would seem correct to say — if it does indeed turn out to be the case that public reason is indeterminate on what death consists in — that the location of that indeterminacy is the PCP itself. There is, however, one part of Rawls's discussion of the PCP that might be taken to suggest, if obliquely, that this piece of theoretical machinery contains further features that can facilitate a resolution of the problem of legally defining death.

I have in mind Rawls's seldom-discussed suggestion that the PCP incorporates the idea of a citizen's 'public, or institutional, identity, or their identity as a matter of basic law'.[39] His brief remarks about public identity reveal that he conceives of it as having synchronic and diachronic dimensions.[40] That is, it provides a standard for the identification of citizens by the state *at a time*, and for their re-identification *over time*, for such purposes as determining their legal rights and share of resources. As with the PCP generally, Rawls suggests that the conception of public identity relevant to public reason is distinctively democratic; other kinds of society, he says, may employ different understandings of when their members continue or cease to be persons, or the same persons, under law. Unfortunately, however, Rawls does not say what he thinks the democratic criterion of public identity might be. Instead, he only illustrates it with an example: under the relevant criterion, he says, someone undergoing religious conversion does not become a different person, or cease to be a person, and conversion is accordingly irrelevant to our legal rights. But while Rawls demurs on the question of what the democratic criterion of public identity consists in, it would be natural to anticipate that, if it exists, it could be used to yield a corresponding public criterion of a person's death or final exit from social relations. This would presumably be uncovered, as is done with metaphysical accounts of personal identity, by following the identity relation forward in time to the point at which it ceases to hold between the person in question and anyone in the future.

This suggestion, while intriguing, faces an obvious problem. For the PCP is expressly designed to stand apart from longstanding philosophical controversies over the nature of our identity and existence, not to provide a basis for wading into and resolving them. Rawls says that the problem of personal identity

> raises profound questions on which past and current philosophical views widely differ and surely will continue to differ. For this reason it is important to develop a political conception of justice that avoids this problem as far as possible.[41]

He claims on behalf of the PCP that

> [i]f metaphysical suppositions are involved, perhaps they are so general that they would not distinguish between the metaphysical views … with which philosophy has

38. Rawls 2005: 18.

39. Rawls 2005: 30.

40. See Rawls 2005: 30–32.

41. Rawls 2005: 32 n. 34.

> traditionally been concerned. In that case they would not appear to be relevant … one way or the other.[42]

And he implies that the idea of public identity in particular is general enough to be acceptable to citizens with a broad range of metaphysical commitments, saying 'all agree, I assume, that for purposes of public life, Saul of Tarsus and St. Paul the Apostle are the same person. Conversion is irrelevant to our public, or institutional, identity.'[43] This would all be an extraordinarily misleading way of presenting the PCP, if the truth were that it came with determinate and contentious commitments regarding when a person should be taken by the state to have ceased to be.

It might be replied that in the foregoing quotations Rawls somewhat overstates the degree of metaphysical equidistance required in the specification of the PCP. What is needed is not, as Rawls seems to suggest, general acceptability to those who hold one of the competing views in the philosophical debate, but neutrality among those views that are *reasonable*. As I noted above, reasonableness, on the Rawlsian understanding, is a matter of acceptance of certain central holdings of the democratic tradition: as Rawls himself puts it, public reason 'does not trespass on citizens' comprehensive doctrines so long as those doctrines are consistent with a democratic polity.'[44] If, then, it could be confirmed that particular commitments regarding our identity, persistence conditions, and death are latent within the basic moral framework of such a society, then the fact that certain philosophical perspectives are incompatible with these commitments would be no obstacle to incorporating them into the PCP. The question, then, becomes one of whether there are indeed any such commitments identifiable within the tradition of democratic thought.

There is some initial cause for optimism here. For the democratic tradition, as glossed by Rawls, does indeed appear incompatible with

42. Rawls 2005: 29 n. 31.

43. Rawls 2005: 32 n. 34.

44. Rawls 2005: 490.

at least some understandings of personal identity and their practical implications that individuals might conceivably hold. Suppose, for instance, that on one view whenever we fall asleep we cease to exist, and the individual who wakes up is a different person. Proponents of this view might take it to have a range of unusual implications — for instance, that it is wrong to hold a person, Y, morally responsible for what his physically and psychologically continuous predecessor, X, did the previous day, or that it is wrong to burden X for the sake of benefits to Y.[45] This seems a paradigmatically unreasonable position, insofar as it conflicts with the idea of personal responsibility, and of the pursuit and refinement of a conception of the good over a prolonged period of time, that Rawls identifies as part of the democratic view of what it means to be free.

It is not enough, however, for the PCP to provide guidance at the margins of the debate by ruling out certain idiosyncratic outlying views. If the PCP is to be the source of a solution to public reason's apparent indeterminacy on death, it must also provide grounds for choosing among the criteria on which the public and philosophical debates have centred, such as those described earlier in this paper. Yet, try as I might, I cannot see how it could do so. For to the extent that it is possible to discern an understanding of a person's public or institutional identity within democratic public culture at all, it is too loose or inchoate to do the necessary work. The best way to confirm this seems to be to attempt to evaluate the accounts of identity, existence, and death described in section 3 on the basis of their liberal or democratic credentials. If one does this one sees that, whatever one might make of their respective philosophical merits, there is none among them that it would be remotely plausible to impugn on the basis that they are insufficiently in keeping with a democratic polity. These views are, as I have suggested, objects of *reasonable* disagreement.[46]

45. For discussion of the ethical implications of a view of this sort, see Olson 2010.

46. Matthew Kramer, I anticipate, would object to this statement. In new work (Kramer 2017: ch. 3), he offers a critique of Rawlsian public reason that has, if I understand it aright, strong affinities with the incompleteness objection,

It should not, then, be a surprise that Rawls failed to identify what the conception of citizens' institutional identity found in democratic public culture consists in: insofar as it exists, it is too coarse-grained to articulate with any precision. It is, then, too coarse-grained to settle the question of whether, for example, the loss of the capacity for consciousness, or self-consciousness, or psychological continuity, is compatible with a person's survival. Liberal democracy, as a system of ideas, is simply not, so to speak, *complete* or *comprehensive* in the required respect.

5. Death as a biological concept

The PCP, I have argued, is of scant help in enabling public reasoners to reach a determinate conclusion about the definition of death. It may seem to some Rawlsians, however, that I have been looking for an answer to our question in the wrong place. Death, it might be said, it not something that happens only to persons, but to all life. It is therefore a *biological* concept. Thus, the appropriate way for a democracy governed by public reason to arrive at a legal criterion of death is for it to treat the question as a scientific rather than a philosophical one, to be resolved in accordance with the evidence and conclusions put forward by the relevant experts.

If death is a scientific concept, it is a heavily disputed one. There exists no consensus, either among members of the public or the scientific community, over how our death is best defined in theory, or which criterion of death should be adopted in practice. This, however, poses a serious difficulty for the suggestion that the political problem of death should be resolved by appeal to science. For while it is indeed permissible under the rules of public reason to draw on scientific evidence and expertise, there are significant caveats.

Rawls addresses the place of science in public reason while setting out what he refers to as the 'guidelines of inquiry'.[47] The purpose of these guidelines is to further regulate the way in which citizens evaluate the applicability and implications of their abstract political values and principles in the concrete circumstances they face, and especially their use of empirical evidence and predictions in so doing. In essence, the guidelines of inquiry impose a general constraint, over and above consensus liberalism's headline requirement of non-reliance on reasonably rejectable comprehensive doctrines, on political appeals to arcane or specialist academic ideas that are opaque to, or contentious among, ordinary citizens. 'As far as possible', Rawls says when describing the guidelines, 'the knowledge and ways of reasoning that ground our affirming the principles of justice and their application to constitutional essentials and basic justice are to rest on plain truths now widely accepted, or available, to citizens generally.'[48] Thus, citizens may not invoke 'elaborate economic theories of general equilibrium, say, when these are in dispute.'[49] And they may likewise appeal to the 'methods and conclusions of science' only when 'not controversial'.[50] Indeed, in a striking passage, Rawls suggests that the reasoning of

and which he articulates primarily with reference to abortion. Perhaps Kramer's central claim (for which see especially pp. 110, 115, and 144–46) is that, where the PCP fails to specify whether certain beings fall within its scope, we cannot say of those involved in the dispute over the moral or metaphysical status of those beings whether their perspectives are reasonable. We can pronounce on their reasonableness, he thinks, only when we have resolved the philosophical debate between them. For only then will we know whether the beings at issue are 'in fact moral persons' (Kramer 2017: 115), and hence which of the disputants envisage treating them consistently with the values of interpersonal freedom and equality. *Pace* Kramer, however, further moral and metaphysical argument of the ordinary kind cannot retroactively transform the content of the reasonable. For reasonableness is just what consensus liberalism stipulates it to be. The perspectives on death described in section 3 are properly accounted reasonable, I contend, in that they are compatible with all those commitments about persons and their relations that Rawlsian reasonable citizens, *qua* liberal democrats, are definitionally required to accept. Their disagreement is on a question which the PCP, as one element of those commitments, fails to settle.

47. Rawls 2005: 223–26.

48. Rawls 2005: 225.

49. Rawls 2005: 225.

50. Rawls 2005: 224.

scientific experts regarding the risk to the population from a nuclear accident is non-public in the same way as the reasoning of a religious group concerning some article of faith.[51] These restrictions rule out selecting a legal definition of death on the basis of scientific testimony, just as surely as they rule out doing so on the basis of clerical authority.

Rawls's suggestion that complex and controversial scientific advice cannot be relied upon in public justification is made repeatedly, and is thus not a mere slip. But it might be argued that it is not a well-considered aspect of his view, which consensus liberals can safely jettison. Catriona McKinnon, for example, has proposed to amend the ideal of public reason to permit appeal to controversial scientific evidence and conclusions, within limits of reasonable disagreement to be determined by the community of relevant experts itself.[52] Her particular concern is that, absent such modification, public reason would not be fit for purpose in formulating policy on climate change. The importance of evidence-based policy-making does not in itself, however, show that we should amend rather than abandon consensus liberalism. To decide *that*, we need to know whether admitting controversial scientific submissions into democratic deliberation can be reconciled with the moral values animating the theory, and hence whether doing so would be more than an ad hoc amendment. Insofar as the relevant values condemn the oppressiveness of coercing people on the basis of claims to deference in judgement by supposed authorities whom they reasonably do not recognise, it is not obvious why scientific authority should not be, as Rawls suggests, regarded as of a piece with ecclesiastical and philosophical authority from the point of view of public reason.[53]

51. Rawls 2005: 220.

52. See McKinnon 2012: 21–30.

53. Some public-reason theorists may embrace this conclusion. Gaus, for instance (who is admittedly not a consensus liberal), has argued (2011: 251–53) that justificatory reliance on expert testimony is permissible only if the coerced have sufficient grounds, at the bar of their own evaluative standards, to accept that those offering it are indeed experts.

The role of science within public reason is an ongoing problem, which deserves attention that I cannot give it here. Fortunately, however, doing so is not required. For even if the best interpretation of the ideal of public reason permits appeal to disputed scientific expertise, the question of when we die is not, I believe, one that science can resolve under its own steam without the addition of controversial metaphysical premises.

This follows most clearly if, as we accepted provisionally in section 3, for someone to die is for her to cease to exist. While science specifies various candidate criteria of death (cardiopulmonary, whole-brain, neocortical, and so on), and is able to identify clinical investigations to confirm whether they have been met and assess the reliability of those investigations, it cannot tell us which criterion marks our ceasing to exist. For it cannot tell us our essential kind or persistence conditions.

This, however, is not the end of the matter. For while the view that death equals our ceasing to exist — sometimes referred to as the *termination thesis* — is plausible and widely held, it is controversial. Critics of the thesis contend that there is a conceptual wedge to be driven between death and nonexistence. It has been argued, for instance, that we ought to accept that a being can be first alive, and then cease to exist, without dying. The amoeba that ceases to exist by dividing, or the embryo in the womb that ceases to exist by fusing with its sibling, are alleged examples. Conversely, some people also think that a thing that is now living could die and yet continue to exist. This is not only a commitment of religious believers in an afterlife. For on some views it is also true to say that an animal or person that dies, rather than ceasing to exist, continues to exist as a dead animal or person.[54] Insofar as there is indeed a conceptual divide between death and nonexistence, such that to specify the conditions of our ceasing to exist is not straightforwardly, or pari passu, to specify the conditions of our death, this fact might be taken to throw into doubt the relevance of personal identity theory for the medico-legal criterion of death. Indeed, David

54. This view is associated in particular with Fred Feldman. See, e.g., Feldman 2000.

Shoemaker has suggested that the existence of such a divide *refutes* the relevance of personal identity to this public policy question. He says that it would be 'bizarre' to say that the amoeba that divides thereby dies, and that it is at least an 'open question' whether one would have died if one 'magically popped out of existence'.[55] And he concludes on that basis that '[c]easing to exist doesn't entail dying, and unless that's the case it seems that what's relevant for the definition of death remains independent of considerations of personal identity.'[56] It would be tempting to suppose that Shoemaker's argument must be helpful to the consensus liberal cause. One might reason that if, as Shoemaker avers, determining the conditions of our death is not a task for personal identity theory, then it must instead be a task for biological science. And if *that* is right, one might then naturally conclude, it suffices to show that the justification of laws or public policies relating to death can remain freestanding of controversial metaphysics, as public reason requires. I believe, however, that to reason in this way would be a mistake.

For a start, the mere fact (if it is a fact) that death and ceasing to exist are not equivalent ideas is insufficient to justify the conclusion that the definition of death 'remains independent of considerations of personal identity'. For it is possible that personal identity theory has an indispensable role to play in identifying the conditions under which we die *even if* the termination thesis is false. To determine whether this is indeed so, we need to know not only that the concepts of death and ceasing to exist diverge, but precisely *how*. Shoemaker does not provide an account of the necessary and sufficient conditions for a thing to be properly regarded as having died. Yet suppose that, in deference to people's intuitions about amoebas and embryos, say, we propose that death be understood as ceasing to exist by means other than fission or fusion. This would be to reject the termination thesis while retaining the relevance of personal identity to the definition of death.

55. See Shoemaker 2010 at, respectively, 487 and 488.

56. Shoemaker 2010: 488.

To be sure, it resolves Shoemaker's 'open question' of whether magically popping out of existence equals death in favour of the view that it does; but this point is at least arguable.[57] I emphasise that it is not my aim to defend the foregoing understanding of the death/nonexistence distinction, or any other.[58] Instead, the relevant point, given our concerns, is that to take a stand on the termination thesis — or, more broadly, to provide an account of the relationship between life, death, existence, and nonexistence — is itself to engage in metaphysical argument. To claim that the conditions of our death are to be obtained from biological science, on the ground that the question is freestanding of personal identity, given the nonequivalence of dying and ceasing to exist, is to rely on a metaphysical thesis that some citizens will reasonably deny. To coerce the latter on those grounds, therefore, would still be a violation of the terms of public reason.

In short, the claim that the conditions for our death can be identified without appeal to personal identity is not to be confused with the claim that this can be done without appeal to metaphysics. Indeed, to underscore the inescapability of metaphysics in this area, suppose that one were to publicly affirm (contrary, as I have argued, to the limits

57. It has been said in support of the termination thesis that, if someone ceases to exist, she must no longer be alive, from which it follows that she must have died. See, e.g., Luper 2016. Shoemaker (2010: 488) questions this, suggesting that it is plausible to think, of a person who magically ceases to exist, that she is now neither alive nor dead. Yet suppose we focus on the concept of *survival* rather than that of being alive. If someone ceases to exist, she fails to survive; but to say that someone did not survive seems equivalent to saying that she died.

58. For an account of what it means for something (whether a person or any other living thing) to die that is in some respects similar to — though considerably more nuanced than — the proposal mooted in the text, see Gilmore 2012. Gilmore argues that to die is to lose the capacity to live without undergoing certain kinds of fission, fusion, or metamorphosis. If I understand him aright, he thinks that his account does not rule out a continued role for personal identity theory in specifying the conditions of death for a person, since the candidate theories can vie for the status of the best explanation of what it means for someone to have the capacity to live. And even if I have misinterpreted him on that point, this is clearly a view that someone could reasonably hold.

of public reason) that it is not as minds, or psychological continuants, but as organisms that we die, and that death is not to be understood as the failure of preservation of numerical identity, but rather as the cessation of the somatic functionings required for an organism to be alive.[59] Even given these hefty assumptions, science cannot provide us with a definition of death unaided. For the question of what level and kinds of somatic functionings are required in an organism if one is to say that it is living is itself metaphysical: it remains outstanding even when one knows all the facts about the processes taking place within its body.[60] Just as science does not, for instance, independently settle the question of whether a fissioning amoeba dies or undergoes deathless annihilation, so it does not settle the question of whether or how far a living human organism is to be defined with reference to continued neurological functioning.

I conclude, then, that public reasoners cannot rely on science to explain how our death is to be conceived, or which criterion of death ought to be adopted in medical practice and policy. Scientists clearly have views about these matters. But they are not acting only in their capacity as scientists when they expound them.

6. Patient interests

Rawlsian deliberators, we have seen, cannot reason their way to a legal criterion of death either by consulting their shared democratic conception of the person, or by referring the matter to biological science. A third alternative, however, may seem more promising. A political community's concerns, a consensus liberal might next emphasise, do not lie in the conceptual analysis of death for its own sake. Rather, as citizens and lawmakers, our interest in death is practical: we need to determine under what conditions the law should allow us to *treat* a person as having died. And what is centrally relevant to this moral question, the Rawlsian might add, is not whether the patients whom we propose to treat as dead are *truly* so, but rather whether we would thereby cause them *harm*. Accordingly, this new proposal goes, we should reframe the debate over death as a question of what is required by respect for the interests or wellbeing of patients whose metaphysical status is in dispute. Call this 'the moralised approach' to reasoning about death. In adopting it, it may seem that we would shift the focus from a philosophical problem that public reason has no authority to consider, onto matters of justice that fall squarely within its competence.

The moralised approach is a familiar perspective in the bioethical debate on death.[61] Some advocate it in part because they think that, until the heart and lungs stop working, and the body begins to disintegrate, there is no fact of the matter about whether a person has died. All of its proponents emphasise that even if someone is alive, it does not follow that they have a stake in their life being continued, or that their wellbeing can be affected by anything we might do to them.[62] The latter point is generally illustrated with reference to organ donation. Current social attitudes and medico-legal practice both endorse the so-called 'dead donor rule', whereby vital organs may be removed from a patient only once dead. And it is standardly assumed that, to determine whether the dead donor rule is satisfied, what matters is whether someone is *truly* dead. For proponents of the moralised approach, however, this is a mistake. Instead, as James Rachels puts it, the relevant question is: 'At what point does the donor no longer have

59. Or, as Gilmore would have it (see the previous note), for it to have the capacity to live.

60. This was recognised, for instance, by the President's Council on Bioethics (2008: 49).

61. See, e.g., Rachels 1986: 42–43 or Veatch 1993.

62. The moralised approach also derives support from Derek Parfit's famous thesis (in Parfit 1987: chs. 12 and 13) that personal identity, or the truth about our survival, is not 'what matters' for the purpose of determining when it would be rational to show prudential concern about what will happen in the future. I do not discuss the Parfitian idea of 'what matters' in the text. I take it for granted that, if public reason must maintain neutrality on personal identity, and if (as I go on to argue in this section) it also cannot resolve the question of when life ceases to be worth living, then it cannot speak to the question of the conditions under which prudential or first-personal concern about the future is justified either.

any use for the organs?'[63] Parallel questions are to be asked with respect to other death-related conduct, such as disconnection from life support, redistribution of the person's estate, and so on.

Most proponents of the moralised approach appear to believe the law should continue to identify a criterion of death, and withhold legal permission to engage in activities such as burial and organ removal until after it is satisfied. Their suggestion is that our judgements about when these activities are ethically acceptable should determine the criterion of death, rather than the other way around. It is, however, worth highlighting the possibility that, if citizens were to engage in moral reasoning about the permissibility of these acts on a case-by-case basis, they could be drawn to a more radical conclusion: that death ought to be effectively abolished as a legal concept. For there is no guarantee that citizens' reasoning would lead them to think that there must be a single point in the decline of the functioning of a human brain and body to which all hitherto death-related activities need be tied.[64] Thus, deliberators might regard the search for a legal criterion of death to have been entirely superseded by a series of discrete questions about when, given the requirements of respect for patients' interests, organ retrieval and so on are to take place. For convenience, in what follows I will continue to speak as though the political question for which public reason requires an answer is: 'When should the law say that a person is to be pronounced dead?' Readers can, however, mentally add the caveat that the relevant question could instead be rendered as something like: 'When should the law allow us to carry out the set of activities which current conventions link to the occurrence of death?' The assessment I will give of public reason's ability to answer the former question also applies, mutatis mutandis, to its ability to answer the latter.

The view that the law on determining death should be formulated on the basis of patient interests has been criticised by those who believe that the metaphysics of death has at least some moral significance in its own right.[65] For present purposes, this debate is irrelevant: what matters is whether the moralised approach is open to Rawlsian deliberators, and would enable them to reach determinate policy conclusions. For two key reasons, the answer is 'no'.

The first reason is that, peculiarly enough, owing to certain complexities in the structure and content of public reason, the moralised approach does not enable citizens to successfully bypass the prohibited question of the metaphysics of death as intended. The explanation for this lies in the fact that, as I have argued in greater detail elsewhere, the moral considerations that count as eligible grounds in Rawlsian public reason for the imposition of a law pertain exclusively to the moral status, entitlements, and interests of *persons*, as defined under the PCP.[66] This creates a problem in the present context, because the question of whether the beings whose interests are centrally at issue when we are trying to decide whether some death-related activity is to be legally permitted ought still to be accounted political persons turns on their personal identity.

These claims require some unpacking. Consider first the claim that, when engaging in public reason, the moral considerations that may be factored into the justification of the use of state power relate only to what is due to political persons. This follows from the requirement that public justifications rest only on political values that reasonable citizens share. As we have seen, the values which reasonable citizens share are limited to freedom, equality, fair cooperation, and public justification (plus, we might add, their various necessary entailments). All of these values, however, on their Rawlsian characterisations, concern *interpersonal* rights and relations. That is, they identify, according to Rawls, forms of treatment that are appropriate to individuals in virtue of their possession (at least during the appropriate periods of

63. Rachels 1986: 42.

64. For an argument to that effect, see Halevy and Brody 1993.

65. See, e.g., DeGrazia 2005: 139–42. For a nuanced perspective, see McMahan 2002: 443–50.

66. See Williams 2015, especially at 30–33.

normal development) of the cognitive capacities required for citizenship.[67] That the shared moral horizons of the reasonable are limited in this way is a consequence of 'reasonableness' having been defined in terms of acceptance of the basic holdings of the democratic tradition. For democracy is (as Rawls himself construes it) simply an approach to conducting the political relationship — that is, the relationship of persons within the basic structure, whereby they exercise power over one another.[68] It does not, then, involve any characteristic stance on our ethical obligations to the planet, or to living beings in general — not even to human beings in general. Thus, to offer a moral justification for political action that is acceptable to all reasonable citizens is to defend that action in wholly person-affecting terms.

Now consider the claim that, for public reasoners to determine whether the individuals whose interests are primarily at stake in the choice of a criterion of death should be understood as persons, they must invoke considerations of personal identity. As we saw in section 4, according to the (inclusive) PCP, a person is not necessarily someone who *now* has the cognitive powers needed for citizenship, but someone who has them over the course of a complete life. Of course, none of the moral patients who might be declared dead under the reasonable conceptions of death canvassed in this paper still possess such powers. This means, however, that to assert that they are persons, whose interests count in public reason, one must identify them as late stages of individuals who earlier possessed those powers — that is, as numerically identical with such earlier individuals (as opposed to, say, beings that previously existed in association with persons and outlasted them, or beings that came into existence only when those persons died). If this is correct, then the moralised approach does not offer an alternative, for Rawlsian citizens, to reasoning about death in metaphysical terms. For to isolate the pool of interests that, from the shared public perspective, are relevant to political decision-making, they must settle the question of personal identity *first*.

Suppose, however, that one rejects my claim that political personhood is a necessary condition for a being's interests to be eligible to be tallied into the public justification of a political decision. Suppose, rather, that one takes the view that a being's interest in continued biological life would, whether they are a person or not, be recognised as a legitimate basis for imposing a law, at the bar of reasonable citizens' shared political values. There is still a second problem. Under the moralised approach, Rawlsian deliberators need to come to a judgement about whether further life would indeed be in patients' interests — a judgement, that is, about whether the future still holds any good in prospect for them. Yet, to distinguish between understandings of the conditions under which life remains worthwhile, or to affirm any one of them as the rationale for choosing between legal criteria of death, would be a paradigmatic violation of neutrality between reasonable conceptions of the good. So the moralised approach merely directs decision-makers to swap one prohibited philosophical controversy for another.

To elaborate: bioethicists who defend the moralised approach typically contend that the point at which life ceases to hold prudential value, and death-related activities may safely be carried out, is the point at which the capacity for consciousness is lost. But while the view that life without the possibility of interaction with the world is of no further benefit is clearly reasonable and widely shared, so too is its denial: many reasonable people believe, on religious or non-religious grounds, that life in a non-conscious state, though sadly diminished, remains a precious gift until one breathes one's last. Moreover, reasonable disagreement over what makes human life worth continuing is not confined to the question whether life beyond consciousness remains a good: it also ranges over the issue of whether and under what conditions life may cease to be a benefit for *conscious* beings. Many individuals, for instance, have come to the conclusion, when contemplating a future in a severely demented condition, that there

67. For the claim that these values apply to persons due to their possession of these capacities, see Rawls 2005 at, e.g., 29–35, 79, 16, and 213.

68. See Rawls 2005 at, e.g., 216–17.

would be no point in going on after the unravelling of the faculties of rationality and self-awareness that make them (in the Lockean rather than Rawlsian sense) persons. Some think, indeed, that it would be intrinsically demeaning to go on in this way. And some may take these claims to be true not only of themselves, but of everyone. That these perspectives on the good are reasonable can once again be confirmed from the fact that none violates the basic political commitments which define the constituency of public justification. To abandon neutrality with respect to them, then, would be to transgress the limits of public reason. Yet this is precisely what the moralised approach requires.

This latest impediment to determinacy arises, note, because of what is unavoidably involved in our making judgements about the limits of the interest in continued life. Public reason requires that citizens appeal only to those aspects of the good that any of their reasonable peers can recognise as such, and that they abstain, conversely, from affirming any position that is prejudicial to the latter's complete understandings of the features or determinants of a life worth living. It is impossible, however, to advance a perspective on whether and to what extent the life of an individual retains prudential value while upholding that kind of neutrality. For to pronounce on that question is necessarily to engage in an accounting of the sources and varieties of goodness that will be available or foreclosed to the patient if her life is indeed extended. It is, then, necessarily to take a stand on whether the things which rival conceptions of the good variously identify as contributors to a worthwhile existence are indeed so.

The latter point bears emphasising, because it helps to show that the way is barred to what might otherwise seem a natural Rawlsian response to the problem currently at hand.[69] This response begins by acknowledging the existence of reasonable disagreement over the conditions under which extending biological life can constitute a benefit. But instead of concluding that Rawlsian deliberators are accordingly powerless to specify those conditions, it instead proposes that, from the perspective of public reason, the point at which it is appropriate, ceteris paribus, to treat a human being as dead is the point at which it becomes possible for reasonable people to diverge on the question of whether further life is capable of serving that being's interests. To identify when this stage is reached, we must consult the beliefs that reasonable people are stipulated to share on the subject of the good. And in essence, the relevant beliefs are that persons have three basic or 'higher-order' interests: one interest in developing and exercising each of their two moral powers to the degrees required by liberal citizenship, and a third in rationally pursuing their determinate conceptions of the good.[70] These beliefs imply that, for at least as long as the possibility of realising these interests exists, our futures hold the possibility of further good. But, the anticipated Rawlsian response now suggests, once an individual's capacities for moral and rational agency have been irreversibly lost, the higher-order interests are no longer engaged by the decision whether to extend her life, and reasonableness therefore does not require citizens to accept that doing so would be worthwhile. Thus, the point at which life can no longer be publicly acknowledged as prudentially valuable is the point at which one no longer possesses the native endowments required for satisfaction of the higher-order interests. Insofar as this understanding of the benefits of existence is derived from the content of the reasonable, the response concludes, we can appropriately think of it as a *political conception of a worthwhile life*.

Although this understanding of the scope of the interest in continued life is clothed in Rawlsian language, I believe that it does not respect the limits of public reason. Before arguing for this claim, however, it is worth noting that the attractions of the envisaged solution to public reason's indeterminacy problem are likely to evaporate for most Rawlsians once we clarify what more precisely it implies. For it commits us not merely to the view that the irretrievably comatose

69. I am indebted to an anonymous reviewer who suggested this response, along with its idea of a 'political conception of a worthwhile existence', to which I turn momentarily.

70. See Rawls 2005 at, e.g., 74.

have no publicly recognisable interest in continuing to live, but also to the view that those who remain conscious (or indeed self-conscious), though in a state of dementia or cognitive impairment sufficiently severe to preclude active citizenship and rational project pursuit, can likewise be subsumed into the category of the dead (at least other things being equal). Yet, while it would be reasonable, in the specialist Rawlsian sense, for one to think that this is so, it is difficult to overstate just how radical — as well as, for all but a few, how deeply unpalatable — this conclusion is. And accordingly, if this conclusion is indeed one that citizens must acquiesce in when adopting the perspective of public reason, then while consensus liberalism will have evaded the incompleteness objection, it will instead be significantly damaged by the fact of its conspicuous breach with prevailing considered moral judgements.

To be sure, a defender of the proposal under examination may want to insist that judgements that conflict with the determinations of public reason — whether reached by the citizens of a consensus liberal polity or by political philosophers — are simply to be disregarded. But this will not do. It is true, of course, that citizens who prove willing to use their political power to resist the policy positions yielded by public reason thereby render themselves unreasonable. But insofar as consensus liberalism seeks to explain how a liberal constitutional regime can achieve 'stability for the right reasons' — stability, that is, based on willing endorsement of the primacy of public reason, as opposed to a mere balance of political forces — it cannot remain indifferent to whether otherwise reasonable citizens find, in significant number, that the implications of public reason, when teased out, are intolerable enough for them to have cause to abandon their duties of civility. On the contrary, as Rawls himself writes, consensus liberalism must 'hope' that the answers to political questions reached by public reason turn out to be within the 'leeway' that reasonable citizens' convictions allow them to accept, 'even if reluctantly'.[71] Moreover, even consensus liberals who would sever their theory's connections with the notion of stability for the right reasons must be sensitive to whether *you and I, here and now*, as Rawls would put it, find that the practical implications of public reason fall foul of our considered judgements in reflective equilibrium.[72] For if consensus liberalism fails this philosophical test, there is no higher court of appeal, as it were, at which it can be vindicated.

In any event, I also believe, to reiterate, that what was earlier referred to as the political conception of a worthwhile life cannot be put forward within the strictures of public reason. Indeed, the phrase 'political conception of a worthwhile life' is, it seems to me, a contradiction in terms. The fundamental problem with the proposal that public reasoners take up this conception is, I submit, as follows. To say that, as far as the public point of view is concerned, there are no grounds for prolonging life after the loss of the moral powers is to say, by implication, that, from that same point of view, hedonic pleasure, for example, or the satisfaction of whatever preferences individuals without moral and rational agency may still be capable of forming, are not intrinsic contributors to a worthwhile existence — contributors, that is, independently of the fact of having been chosen by a person as an end. It is therefore to say, by further implication, that citizens who follow reasonable conceptions of the good that do regard pleasure, or preference satisfaction (or what have you) as intrinsic goods, and direct that they be promoted accordingly (within the limits of justice), are wasting their time. These claims will seem implausible to many of us. But more pertinently, there would be nothing meaningful left of neutrality over the good if consensus liberalism were to permit them to be made.

71. Rawls 2005: 246.

72. This second consideration remains relevant, then, to, e.g., Jonathan Quong. For while Quong defends an 'internal conception' of consensus liberalism, whereby political arrangements need be acceptable only to a hypothetical constituency of reasonable citizens whose commitment to upholding the outcomes of public reasoning never wavers, he nonetheless accepts, if I understand him correctly, that consensus liberalism must be justified *to us*, from the perspective of the philosopher, in reflective equilibrium. See Quong 2011: chs. 5–6 (on the internal conception), and 155–56 (on the role of reflective equilibrium in justifying consensus liberalism).

Indeed, it is precisely on these grounds, I take it, that Rawls specifically cautions us that public reason must abjure evaluations of people's overall quality of life or level of wellbeing.[73] He argues that, for political purposes, assessments of how well-off people are (or would or will be) should be conducted instead in terms of their shares of primary social goods — despite the fact that 'primary goods are clearly not anyone's idea of the basic values of human life and must not be so understood'.[74] Rawls seems not to have anticipated, then, that, for resolving certain political questions, quality-of-life assessments may be indispensable, and the metric of primary goods not an acceptable substitute. In what is, as far as I can tell, his sole explicit reference to the interest in continued life, and its relevance to political decision-making, Rawls says only that 'any workable political conception of justice that is to serve as a public basis of justification … must count human life … as in general good'.[75] The words 'in general' here mask public reason's difficulties (as, for that matter, does the word 'human').

I conclude, then, that the moralised approach is a dead end for public reason. I have argued that the political values shared by Rawlsian citizens enjoin respect only for the interests of political persons, and that public reason is therefore hamstrung by its inability to confirm, without recourse to metaphysics, whether human beings at the end of life, whose cognitive endowments have decayed, remain persons in the relevant sense. If I am wrong about that, however, and a being's interests are to be factored into the public justification of political arrangements irrespective of whether they belong to persons, public reason will still be unable to identify a point at which a patient's interest in further life runs out without violating neutrality over the good. (Finally, if I am wrong about *that too*, and what public reason instead requires is that citizens acquiesce in the conclusion that there are no grounds for extending life once the three publicly recognised higher-order interests of persons are no longer engaged, then we have also seen that consensus liberalism will still not be saved, but rather exposed to new objections which seem at least as grave as the original incompleteness objection.)

There is a certain irony to these findings. Ethics has recently witnessed considerable movement towards the view that the metaphysical truth about our identity and survival is of much less practical significance, prudentially and morally, than has conventionally been assumed.[76] It appears, however, that consensus liberalism, under which public justification must be, in Rawls's famous slogan, 'political, not metaphysical', cannot derive the expected benefit from these developments.

7. Third-party interests

I have now argued that public reasoners cannot decide between reasonable criteria for determining when a person has died by (a) metaphysical reasoning; (b) reasoning about the implications of their shared conception of the person; (c) consulting the resources of biological science; or (d) reasoning about the way in which the relevant policy options would impact upon patients' interests. Since these options seem to exhaust the viable possibilities, I submit that we are justified in concluding at this point that public reason does indeed have an indeterminacy problem with respect to the political question of death.[77]

73. See Rawls 2005: 188.

74. Rawls 2005: 188.

75. Rawls 2005: 177.

76. A powerful recent articulation of that position appears in the work of David Shoemaker. See, e.g., Shoemaker 2016.

77. If an earlier argument of mine was correct, this conclusion contrasts interestingly with the way public reason handles the primary moral problem arising at the *beginning* of life: abortion. While public reason is unable to specify, I have here maintained, whether a range of patients with eroded psychological, neurological, and physiological functioning remain living persons under the PCP or have (publicly recognisable) interests that tell against treating them as dead, we can be certain — or so I have contended elsewhere — that fetuses are *not* political persons at any stage of pregnancy, and hence are at no point eligible for the protection that that status confers. For even under the inclusive PCP, the political relation between persons within the basic structure is taken to extend only between birth and death. By that token, while public reason

This verdict comes with a caveat. In asking what conclusions Rawlsian deliberators would be warranted in drawing about death, I have implicitly assumed that, to make a decision, they require access to reasons pertaining to the intrinsic properties of the patients who stand to be pronounced dead under the various criteria at issue. Someone might think, however, that even if reasons of the latter sort are unavailable, citizens may nonetheless be able to make headway by reasoning instead about the effects that implementing the candidate criteria may have on the publicly relevant interests of third parties, or society at large. I accept that appeal to third-party-focused reasons might *in some societal contexts* enable public reasoners at least to narrow the field of public policy options — conceivably even to the point of resolving the policy question altogether. Yet, it would be a mistake to think that the availability of these reasons adequately alleviates the indeterminacy problem that I have developed thus far.

Any attempt by citizens to reason their way to conclusions about when patients should be declared dead on grounds of considerations of the foregoing sort would, I take it, have to proceed in a particular way. One would have to say that, while public reason cannot offer any answer to the question of whether those patients are living persons who retain lives of value, the interests of third parties are sufficient to carry the day no matter what the answer might be assumed arguendo to be. To be sure, it seems that nobody could reasonably argue, in that vein, that whether or not a patient, P, is assumed to be a person with a life worth living at time t1, we should nonetheless go ahead and treat him as dead at t1, as a means of securing the benefits that would thereby accrue to, say, people on organ waiting lists. For so to argue would express a readiness to engage in the instrumentalisation of persons that is incompatible with the special priority of basic rights over the general welfare — a priority which, according to Rawls, any sufficiently liberal understanding of justice will endorse.[78] It does, however, seem possible for citizens to argue that, even if P really were dead and beyond harm at t1, he still ought not to be declared so, on grounds that adopting the relevant criterion of death would cause too much third-party harm. Citizens might be able to make a compelling case, for instance, that given prevailing social attitudes, pronouncing patients dead at that stage would attract too much public hostility, or unduly damage trust in doctors or state officials.

It seems to me that it would be implausible to try to defend consensus liberalism's handling of the problem of death on grounds that, although it prevents citizens from asking morally pertinent questions about patients themselves, it at least allows them to whittle down their policy options with reference to their third-party effects. For it would be far more natural to conclude that, if public reason forces this degree of reliance on third-party interests, its rules are unduly burdensome. Irrespective of its surface plausibility, however, reasons pertaining to third-party effects are too contingent on variable social circumstances

fails to deliver a verdict regarding when we are to be considered dead, its verdict regarding abortion seems both determinate and radically permissive (indeed disturbingly so, as most would think). See Williams 2015.

Matthew Kramer has newly disputed my earlier position. While he is likewise a critic of public reason's management of the abortion controversy, he believes that 'the Rawlsian conception of persons does not *in abstracto* entail or exclude the personhood of foetuses' (2017: 152–55). This view requires that Kramer discount various statements by Rawls to contrary effect — as when, e.g., Rawls claims (2005: 41) that, before entering society by birth, 'we have no prior identity'. Yet while I think Kramer's understanding of the PCP does not square with the Rawlsian account, I cannot make that case here. Notice, however, that even if his interpretation of the PCP were right, it would not follow, as he contends, that resolving the problem of abortion via public reasoning is 'impossible' (Kramer 2017: 92). Kramer assumes too readily, in particular, that appeals to women's prerogatives to prioritise themselves over their fetuses (irrespective of the latter's moral status) can justify abortion under only rare circumstances, when it seems at least reasonable for a citizen to argue that, given the burdens of pregnancy and childbirth, such appeals justify abortion frequently, or indeed always. Nor does Kramer anticipate that, where public reasoning runs out, consensus liberalism might call for a procedural resolution to the problem at hand. And *that* means, I believe, that like others he misses the ultimate practical and moral significance of indeterminacy in public reason, as I develop it below.

78. See Rawls 2005 at, e.g., 450.

for the envisaged defence of public reason to be relied upon as generally applicable. Although such reasons *might* be an aid to deliberators, given a particular confluence of social attitudes, institutions, practices, and so on, we cannot be expected to grant that they will always come to the rescue. It is appropriate that public reason be judged in part on the basis of its consequences for cases in which, given the social facts, there happens to be no reasonable criterion of death of which citizens could claim, with appropriate warrant, that it would cause significant third-party harm, or in which it is clear that any such harms would be more than compensated by benefits.

On these grounds, in what follows I propose to set third-party-focused reasons aside as the source of a potential solution to the indeterminacy problem I have identified. We can simply stipulate, without extravagance, that we are considering public reason's performance under societal conditions in which these reasons would not provide a catalyst to decision-making.

What would follow from this fact? Proponents of the incompleteness objection have generally assumed that, where public reason proves unable to answer some question for which its use is mandated under the duty of civility, this suffices to show the permissibility of appealing to non-public reasons, and hence the falsity of the claim that doing so, within the relevant class of political decisions, is morally wrong. As Andrew Williams and Micah Schwartzman have argued, however, that conclusion does not follow.[79] For selection of a legislative course of action by non-public reason may not be the only remaining alternative. And if it is indeed objectionably sectarian to govern free and equal persons in accordance with non-public reasons, as consensus liberalism claims, then these other possibilities must first be explored.

79. See Williams 2000: 209–11 and Schwartzman 2004: 209–14.

8. Coping with indeterminacy: five unsuccessful strategies

Schwartzman has identified no less than five distinct strategies that citizens might employ to cope with incompleteness in public reason without reaching for their comprehensive doctrines.[80] They are: (1) 'intrapersonal delegation', or deferral of a decision until later, when further public reasons may have come to light; (2) deference to others who claim to have succeeded where one has failed to answer the relevant question by public reason alone; (3) moral accommodation between opposing perspectives; (4) calling time on deliberation, and proceeding to a majority vote; (5) random adjudication, by a procedure such as a coin flip. The efficacy and moral appropriateness of these strategies has so far not received much attention, and our investigation provides a good opportunity to do so. In this penultimate section I argue that, in the case of indeterminacy over the definition of death, the ideal of public reason requires (5), and that, insofar as it does so, the ideal is objectionable.[81]

It is not entirely clear how many of Schwartzman's strategies he takes to be applicable to cases of indeterminacy in public reason, as opposed to the different problem of inconclusiveness. The only approach which he rules out explicitly — calling it 'useless in the face of indeterminacy' — is (4).[82] It is worth pausing to clarify why. Recall from earlier that public reason is indeterminate when it provides, as in the present case, insufficient reasons to justify one's venturing to choose in any way from among the relevant policy options, and inconclusive when citizens find that they have sufficient reasons to adopt their various competing policy preferences, but public reason cannot bring them into agreement regarding which is best justified by vindicating any option beyond reasonable doubt. Democratic selection

80. Schwartzman 2004: 209–14.

81. Some parts of this argument refine and expand upon parallel claims which I have defended elsewhere about the utility of Schwartzman's proposals in the different context of *global* public reason. See Williams 2016: 18ff.

82. Schwartzman 2004: 211. Quong (2013), on the other hand, appears to believe that at least (1), (3), and (5) are relevant to indeterminacy.

of an inconclusively justified policy appears fully compatible with the ideal of public reason. For the policy imposed is indeed justified, so those who propose it can sincerely attest, by a reasonable balance of public reasons, even if many do not consider it optimal, or most reasonable. That claim cannot be made, however, where public reason is indeterminate. If no policy is supported or ruled out by public reason, then citizens who are nonetheless able to reach a judgement must have done so on the strength of their comprehensive doctrines. And enforcement of those judgements by a democratic majority would be a straightforward violation of the Rawlsian ideal.

With (4) eliminated, then, let us consider Schwartzman's other proposals. I assume that (1) and (2) are also irrelevant here. For if it is correct that public reason supplies insufficient grounds to make a decision about the definition of death because it prohibits appeal to the necessary philosophical considerations, then deferring the decision until later, or looking to someone else, will not help.

At first sight, proposal (3), for moral accommodation, might seem no more promising. The only form of accommodation that Schwartzman mentions explicitly is compromise-brokering. And it may be difficult to imagine what compromise between proponents of opposing definitions of death would even look like, let alone to envisage the prospects for obtaining one being any more than extremely remote. After all, compromise on this issue would generally mean, for one side, acceding to some people's lives being ended prematurely, and for the other agreeing to the pointless squandering of organs and other scarce resources. Depending on the factions involved, however, and their particular concerns, compromise may sometimes be conceivable. But even if it were, it is ruled out in cases of indeterminacy, for reasons that run parallel to those ruling out resolution by democratic voting. Suppose that public reason is indeterminate between policies P_1, P_2, and P_3, and that the public is split between advocates of P_1 and P_3. As before, since no policy is supported by public reason, if citizens are nonetheless able to reason their way into a preference, it must be by reference to their non-public doctrines. Compromise on P_2, in this context, means agreeing to govern by striking a balance between those doctrines. And that approach to political decision-making is condemned by the Rawlsian view as 'political in the wrong way'.[83]

This conclusion also applies to another, somewhat different potential strategy for reaching moral accommodation between differing perspectives on death.[84] Here we select a criterion of death on the basis that all reasonable perspectives can at least agree that its satisfaction is sufficient for the death of a person. Assuming a society in which all reasonable views are represented, this would presumably yield a criterion of death as the irreversible breakdown of biological functioning to the point where both the cardiopulmonary and whole-brain standards are met. Moral accommodation in this form is not naturally described as compromise, because, while it aims at a policy that is in one respect acceptable to all parties (it ensures that nobody will think that the law declares people dead prematurely), it is not an attempt to split the difference between existing policy proposals, or to find a settlement that the opposing camps themselves deem equally satisfactory. Indeed, under current technological constraints, the policy obtained under this approach aligns almost exactly with the cardiopulmonary criterion, at the heavy expense of all neurological standards, since absent head transplants irreversible cardiopulmonary failure makes total brain death unavoidable, thereby satisfying the joint criterion, while psychological disintegration and brain death are compatible, as we have seen, with long-term maintenance of cardiopulmonary function. In common with compromise-brokering, however, this proposal is 'political in the wrong way'. For in the absence of a public justification of any particular criterion of death, it again views the political task at hand as one of seeking an accord between citizens, addressed in their capacities as holders of rival comprehensive doctrines.

There is, however, yet a further form of moral accommodation remaining that might fare better. This is what we might call the strategy

83. Rawls 2005 at, e.g., xlv.

84. I am grateful to Paul Billingham and Jeff McMahan for suggesting that I consider this possibility.

of *privatisation* — that is, of ceding a political matter to individuals to resolve in their own cases, rather than insisting on a unitary, community-wide response for all.[85] In the current context, privatisation means allowing persons to decide what criterion of death will be applied to them. This possibility is of particular interest for two reasons. First, some jurisdictions already grant their citizens a degree of this sort of discretion. Japan and the American state of New Jersey, for instance, have both legislated to allow individuals to exempt themselves from neurological criteria of death, out of a concern to accommodate religious beliefs to the effect that (earthly) death occurs only once the traditional cardiopulmonary standard is met. Second, some bioethicists who are sympathetic to Rawlsian liberalism have advocated privatising the decision over the definition of death, precisely as a means of accommodating reasonable pluralism over human survival and the value of life.[86] As a solution to indeterminacy, however, the strategy of privatisation fails.

One reason for this is that it can at most obviate the need for the enforcement of a collectively made decision in cases where the wishes of a previously competent person are known. Many cases, however, will obviously not be like this. And we cannot avoid this problem just by requiring that everyone records a prior personal decision, or by implementing a system of presumed consent, whereby the state communicates that it will infer that everyone accepts some default criterion of death if they do not opt out. For that still leaves the issue of what to do with individuals who, like children, lack the mental capacity to make their own medical decisions. It does not appear that the community would be justified in granting family members, as the designated legal agents of incompetent patients, the power to decide when the latter should be declared dead. Rawls stresses that any reasonable political conception of justice will perforce accept that citizens lack untrammelled authority over their children or other dependent persons in their care, and that the state is entitled to intervene in the home to prevent abuse and neglect.[87] Yet, unless public reasoners can determine how the options placed before a fiduciary agent stand to affect the interests of the patient on whose behalf she purports to act, they will be unable to judge whether the amount of discretion granted falls within reasonable bounds, or constitutes a license to engage in mistreatment. And in any case, for at least some incapacitated persons, who lack loved ones, or sufficiently responsible loved ones, the appointed agent will be a state official.

Even when it comes to persons whose prior wishes are known, however, there is a further obstacle. This is that a policy of deferring to these wishes cannot itself be justified except by ruling on precisely the sort of contentious philosophical issues that privatisation aims to sidestep. To fix ideas, suppose that a hospital patient signs, in an appropriately voluntary fashion, an advance medical directive requiring that, in his case, the point at which death should be treated as having occurred is the point at which the cardiopulmonary standard is fulfilled. Later, he suffers a serious medical complication that results in total brain failure, though cardiopulmonary functioning is artificially sustained. Is the directive authoritative? That depends on whether the patient remains, at the point at which the choice arises whether or not to fulfil its terms, a source of valid claims against us, as Rawlsians would put it. But public reason is powerless to answer that question.

The grounds for the latter claim can likely by now be at least partly anticipated. First, public reason cannot take a stand on whether the individual who signed the directive still survives as a person under the PCP, to whose treatment the political values apply. And nor, second, can it take a stand on whether it matters in any way to the dead, or those whose capacity for a mental life has been annihilated, that their earlier wishes be carried out. To say that they retain an interest that we so

85. While Schwartzman does not discuss privatisation, it is at the heart of Gaus's approach to overcoming the (somewhat different) problem of indeterminacy that he regards as a danger for his version of convergence liberalism. See Gaus 2011: ch. VI.

86. See Zeiler 2009 and DeGrazia 2005: 138.

87. See Rawls 2005: 466–74.

act would be to violate neutrality between reasonable conceptions of the good, by contradicting the controversial *experience requirement*, on which an individual's interests are affected only by things that make a difference to her experience. But to say that we ought to respect the determinations of a person's autonomous will irrespective of whether our doing so would benefit her would likewise be to venture beyond public reason's remit. For the public justification of political arrangements, according to Rawls, must not rely on any reasonably rejectable understanding of the ethical significance of autonomy.[88] Rawls cites the doctrines of Kant and Mill in this regard. And I take it that, in giving the examples of those particular thinkers, he meant to suggest that it is verboten to appeal to unshared conceptions not only of the ways in which respecting people's autonomous choices may contribute to their good, but also of the ways in which doing so may serve values independent of their good — as derived, say, from a philosophical account of the nature and demands of human dignity, or of the intrinsic or impersonal value of states of the world. Certainly, the rules of public reason would seem utterly arbitrary if these species of view were not treated even-handedly. Yet, while reasonable citizens necessarily accept that persons have autonomy rights grounded in the higher-order interests, there is nothing unreasonable in their taking the view that our reasons to respect people's choices are exhausted once their good is (as those citizens see it) no longer at stake.[89]

So much, then, for moral accommodation. At this point, the only one of Schwartzman's coping strategies left standing is (5) — random adjudication. None of the other proposals, as we have seen, enables decision-makers to select a policy without reliance on non-public reasons. Random adjudication does so — though admittedly only at the cost of abandoning the ambition of justifying laws on the basis of a positive balance of public reasons. Yet, insofar as the ideal of public reason not only directs citizens to aim at public justification, but also requires them to abstain from foisting their non-public doctrines on others, it appears, as Schwartzman and Williams claim, that they should avoid doing the latter even when they are unable to achieve the former.[90] If this is right, then given that random adjudication is the only available way to proceed while maintaining independence from comprehensive justification, this is what consensus liberalism requires.

Schwartzman acknowledges that the suggestion that we resolve important political problems randomly when public reason proves indeterminate, rather than by inquiring after the best available non-public reasons, is likely to strike us as 'highly implausible, if not altogether absurd'.[91] In defence of random adjudication, however, he provides an example in which it seems like the right thing to do. This is the case of a hospital board charged with deciding which of two patients should receive an organ. As Schwartzman constructs the example, both individuals are equally suitable from the point of view of public reason: the board are able to confirm that they are in equal need, would derive an equal benefit from the organ, have spent the same length of time on waiting lists, are equally non-responsible for their plight, and so forth. Schwartzman then asks whether, given that public reason does not identify a preferred candidate, it would be appropriate for board members to break the tie by discriminating on the basis of religious affiliation or sinfulness. He concludes — and I assume everyone would agree — that deciding the matter on that particular basis would be wrong, and that random adjudication is morally required.

Schwartzman's example, however, is not entirely apposite to a defence of public reason. For the selection of organ beneficiaries on the basis of religious devotion or purity does not only fall foul of the

88. Rawls 2005 at, e.g., 78 or 400.

89. Notice that these considerations suggest that public reason has a problem justifying legal recognition not only of advance medical directives, but also of people's wills. To be sure, it is commonly accepted in liberal societies that wills ought to be upheld, precisely on grounds of respect for the autonomy of the dead. But it is by no means a requirement of reasonableness that one should accept this, and many philosophers of course do not.

90. See Schwartzman 2004: 213 and Williams 2000: 210.

91. Schwartzman 2004: 212.

ideal of public reason — it also constitutes a violation of church-state separation, and of basic religious freedom, to a degree that would be condemned by public reason liberals and their critics alike.[92] Given the details of Schwartzman's case, it is not only public reasons, but reasons of justice generally, that have been exhausted in the comparison between the two patients. So comprehensive or ethical liberals can agree that a coin toss, say, is the appropriate solution. The question here is not whether random adjudication is ever called for, but whether public reason forces citizens to rely on it excessively. To test *that*, we need to concentrate on political questions for which Rawlsians would have to turn to randomisation before ordinary moral deliberation has run its course.

If the argument of this paper is correct, the legal determination of death is (in at least some societal circumstances) just such a question. This example does not work to the advantage of consensus liberalism, since the claim that a political community should randomly decide the provisions of its laws in this area is intuitively and reflectively unacceptable. To resolve this particular problem arbitrarily would mean abstaining from asking whether the policy adopted will prematurely end lives that are worth continuing, or, conversely, extend biological life past the point of ethical justification, to the detriment of (for instance) the supply of life-saving organs. This is to play Russian Roulette with people's lives and wellbeing. And the stakes are all the higher given that the range of reasonable understandings of when death should be taken to have occurred seems rather broad. As we have seen, reasonable answers to the question of when there is no longer any patient-centred objection to pronouncing that death has occurred range from the final stages of dementia, when one's distinctive psychological attributes, or the higher cognitive powers associated with personhood, have been lost, through higher and whole brain failure, to the point at which the heart and lungs finally stop working. This suggests that the menu of policy options between which random adjudication might be called for will (except insofar as contingent third-party-focused reasons intervene) in turn be wide. This conclusion is at least damaging to consensus liberalism. It would be far from implausible, I think, to regard it as a reductio of it.

92. For an ethical liberal defence of church-state separation, see Arneson 2014.

9. Conclusion

Rawlsian consensus liberalism requires that the justification of coercive laws (or at least the most fundamental laws) be formulated without reliance on reasonably rejectable claims about the basic nature and value of human survival. I have argued, as critics of consensus liberalism have often suspected, that fundamental political problems are not always susceptible to resolution by public deliberation conducted within these constraints. The determination of death provides an example of a political dispute which does not merely depend upon but essentially *is* a dispute about the nature and value of life. One conclusion to be drawn from our investigation, then, is that to forbid democratic engagement between rival comprehensive doctrines is in some cases equivalent to forbidding citizens to resolve fundamental problems of justice — at least through the use of reason.

Another conclusion to be drawn is that the procedural mechanisms which Rawlsians have proposed for coping with incompleteness in public reason are not only of insufficient help, but in at least some cases exacerbate public reason's difficulties. The claim that indeterminacy should be resolved by random adjudication takes on the objection that public reason is sometimes unable to decide what to say about a policy problem, and transforms it into an objection that seems more dramatic: that public reason can require picking political arrangements in an intolerably arbitrary way. This finding alters the cast of the incompleteness objection, by closing the gap between it and what I have elsewhere called the *ethical objection* — the objection, that

is, that public reasoning can in some cases generate (or be at undue risk of generating) determinate but morally unacceptable decisions.[93]

If my assessment of the implications of public reason for the determination of death is correct, Rawlsians face a difficult choice how to respond. It would be tempting to suppose that problems of incompleteness like this one can be satisfactorily addressed with a bit of theoretical tinkering — and more specifically by amending the content of public reason to allow extra reasons in and facilitate better decision-making. But this would seem at odds with the fundamental commitments of consensus liberalism. Under the consensus model, the justificatory reasons that citizens may invoke depend on the reasons their peers can accept. Thus, additions to the content of public reason require corresponding amendments to the *constituency* of public reason. The sine qua non of the Rawlsian view, however, is, as we have seen, that the constituency of public reason should be open to all who subscribe to (the Rawlsian interpretation of) the basic insights of the democratic tradition. To discriminate further among persons who are fully reasonable by this standard, then, seems an abandonment of this Rawlsian commitment.

Indeed, doing so would appear to produce a slide from the idea of public justification to the so-called *correctness-based* standard of justification to which ethical liberals typically subscribe, under which political decisions are permissibly implemented when justified by *valid* — as opposed to public — reasons.[94] The proposal we are now considering for restriction of the justificatory constituency is designed to meet an objection to the effect that, in its current form, public reason resolves the political question of death in a morally unacceptable (because arbitrary) fashion. Given that motivation, however, the discrimination called for among the reasonable would presumably have to be on the basis that particular beliefs about life and death are needed to facilitate morally better outcomes. Yet, once the principle is conceded that it is appropriate to discriminate among the reasonable for that purpose, it is doubtful that there could be any principled objection to doing so again for other political questions. The circle of justification would then quickly close to those with morally sound beliefs on every issue. And, as Lister has noted, the view that political justification need only be acceptable to those with sound beliefs is equivalent, precisely, to the view that justification should be correctness-based.[95]

It might be replied that there is one way in which the content of public reason could be revised to circumvent the incompleteness objection while maintaining the required distance between consensus and ethical liberalism. This would be to abandon public reason's use of what I earlier called the inclusive PCP, in favour of the PCP's exclusionary reading. You will recall from section 4 that, whereas the inclusive PCP counts us as persons for the duration of our complete lives in society, without specifying when that life ends, we remain persons under the exclusive PCP only for as long as we possess the moral powers needed for citizenship. Substituting the inclusive for the exclusive PCP, a proponent of this move might argue, achieves determinacy without importing alien philosophical content from the realm of comprehensive doctrines into the political domain. But it is unclear to me, nonetheless, that doing so would be any more consistent with consensus liberalism's founding values. For *if* Rawls is correct that it is the inclusive rather than exclusionary PCP that is presupposed by the democratic tradition, then the proposal at hand still discriminates against paid-up democrats whom we had previously been told are entitled to be counted among the constituency of public justification. It is dubious that the fact that the presence of these citizens within the relevant constituency is an obstacle to determinacy on some political questions would be a good enough reason, by consensus liberal lights, for casting them out.

Leaving aside the question of whether it would be *coherent* for consensus liberals to endorse adoption of the exclusionary PCP, however,

93. See, e.g., Williams 2015: 49.

94. The term 'correctness-based justification' comes from Wall 2002: 386.

95. Lister 2013: 40.

I believe that it would not be *prudent* for them to do so. As with the suggestion we encountered in section 6, to the effect that public reason should identify the irretrievable loss of the moral powers as the point beyond which further life can no longer be considered in someone's interests, making the mooted change would only succeed in exchanging a problem of indeterminacy for a problem of public reason's yielding conclusions that are dramatically out of step with mainstream moral judgements. As I have argued, public reason recognises only the rights and interests of persons as legitimate grounds for political action. Thus, for public reason to withhold the status of person from someone is, as Rawls himself puts it in the context of slavery, for it to deem them 'socially dead'.[96] To pronounce us socially dead, if not literally so, immediately upon the loss of the moral powers needed for citizenship would certainly avoid the incompleteness objection as I have developed it. But since, as we have seen, the absence of these powers is consistent with the presence of (self-)consciousness — and may thus allow for various forms of enjoyment, recognition of and affection towards others, and so forth — adoption of the exclusionary PCP would not only not help consensus liberals: it would make things considerably worse for them, in two respects. First, it multiplies the fronts on which they are exposed, by generating a range of additional implausible implications about the status and permissible treatment of individuals who have yet to develop the moral powers, in addition to those who have lost them.[97] And second, even where the end of life alone is at issue, it seems at least somewhat less bad for consensus liberalism to require random selection from among the various reasonable criteria of death available, as it does when public reason is indeterminate, than for it to directly dictate adoption of a policy that, while reasonable, the vast majority of us would consider the most repugnant of the options on offer.[98]

If, on the foregoing grounds, consensus liberals cannot respond to the problem posed in this paper by altering the terms in which public justifications must be offered, their only alternative seems to be to bite the bullet, and accept that public policy on the determination of death should (absent a fortuitous balance of third-party-centred reasons) be determined randomly. The sustainability of the bullet-biting response depends, however, on how many other important political questions the Rawlsian model may fail to resolve satisfactorily. I have argued elsewhere that there are indeed other such questions, and I believe that there are yet more to be discovered. If so, it will become increasingly implausible to suggest that the sort of counter-intuitive consequences I have here described are a bearable cost to be priced in when adopting the consensus liberal view.

96. Rawls 2005: 33.

97. A consensus liberal might counter that to say that a being is not a person is not to say that they may be treated in any way we like. It is merely to say that their treatment is not a constitutional essential or matter of basic justice. And since this means, on some Rawlsian views, that it is a question that can be resolved in accordance with non-public reason, withholding the status of person can be compatible, the interjection goes, with granting an individual stringent legal protections. Two points in response. First, this defence is obviously not available to those who think that public justifications must be offered for fundamental and non-fundamental political decisions alike. And second, even if the use of public reason is required only in tackling fundamental questions, the treatment of non-persons will still count as such a question insofar as the basic interests and rights of persons are simultaneously at stake. Yet it appears that this will frequently, if not always, be so. I take it, for instance, that insofar as organs are analogous to other scarce resources, it is a question of basic justice how we harvest and distribute them. By that token, however, when the exclusionary PCP is in place there will be no meaningful room to consider the ethical treatment of humans who have lost the moral powers outside the confines of public reason, even if it is assumed that public justifications must be produced only when fundamental questions are at issue.

98. I am grateful to an anonymous reviewer for prompting me to consider the option of amending the PCP in the name of achieving determinacy. Technically, I should add, the reviewer's proposal was for a halfway house between the inclusive and exclusionary PCPs, whereby members of society count as persons prior to developing the moral powers, but not after losing them. With the exception of the penultimate sentence of the paragraph to which this note is appended, I believe that the considerations adduced in the text against adoption of the exclusionary PCP also carry over to this suggestion.

Acknowledgements

Earlier versions of this paper were presented at a meeting of the Oxford Moral Philosophy Seminar in March 2015, and a workshop on 'Personal Identity and Public Policy'—also held in Oxford—in October 2016. I am grateful to the participants for their comments on those occasions, and for a wealth of written feedback provided by Paul Billingham, Tim Campbell, Jeff McMahan, Anthony Taylor, and two referees for this journal. I also record my thanks to the Leverhulme Trust for the Early Career Fellowship during which the research for this article was carried out.

References

Arneson, Richard. 2014. Political Liberalism, Religious Liberty, and Religious Establishment. In H. Dagan, S. Lifshitz, and Y. Z. Stern, eds., *Religion and Human Rights Discourse*. Jerusalem: Israel Democracy Institute. 117–44.

Bernat, James L. 2006. The Whole-Brain Concept of Death Remains Optimum Public Policy. *Journal of Law, Medicine & Ethics*, 34: 35–43.

Brierley, Joe. 2015. UK Court Accepts Neurological Determination of Death. *The Lancet*, 385: 2254.

DeGrazia, David. 2005. *Human Identity and Bioethics*. Cambridge: Cambridge University Press.

Feldman, Fred. 2000. The Termination Thesis. *Midwest Studies in Philosophy*, 24: 98–115.

Gaus, Gerald. 1996. *Justificatory Liberalism*. New York: Oxford University Press.

——. 2011. *The Order of Public Reason*. Cambridge: Cambridge University Press.

Gilmore, Cody. 2012. When Do Things Die? In B. Bradley, F. Feldman, and J. Johansson, eds., *The Oxford Handbook of Philosophy of Death*. Oxford: Oxford University Press. 5–59.

Green, Michael B., & Daniel Wikler. 1980. Brain Death and Personal Identity. *Philosophy & Public Affairs*, 9: 105–33.

Greenawalt, Kent. 1988. *Religious Convictions and Political Choice*. New York: Oxford University Press.

Halevy, Amir, and Baruch Brody. 1993. Brain Death: Reconciling Definitions, Criteria, and Tests. *Archives of Internal Medicine*, 119: 519–25.

Hartley, C., and Lori Watson. 2009. Feminism, Religion, and Shared Reasons: A Defense of Exclusive Public Reason. *Law and Philosophy*, 28: 493–536.

Kramer, Matthew. 2017. *Liberalism with Excellence*. Oxford: Oxford University Press.

Lister, Andrew. 2013. *Public Reason and Political Community*. London: Bloomsbury.

Luper, Steven. 2016. Death. In E. N. Zalta, ed., *The Stanford Encyclopedia of Philosophy*, Summer 2016 Edition. URL = https://plato.stanford.edu/entries/death/.

McKinnon, Catriona. 2012. *Climate Change and Future Justice*. Oxford: Routledge.

McMahan, Jeff. 2002. *The Ethics of Killing: Problems at the Margins of Life*. Oxford: Oxford University Press.

Olson, Eric. 1997. *The Human Animal: Personal Identity Without Psychology*. Oxford: Oxford University Press.

——. 2010. Ethics and the Generous Ontology. *Theoretical Medicine and Bioethics*, 31: 259–70.

Parfit, Derek. 1987. *Reasons and Persons*. Oxford: Oxford University Press.

President's Council on Bioethics. 2008. *Controversies in the Determination of Death*. Washington, DC: President's Council on Bioethics.

Quong, Jonathan. 2011. *Liberalism Without Perfection*. Oxford: Oxford University Press.

——. 2013. Public Reason. In E. N. Zalta, ed., *The Stanford Encyclopedia of Philosophy*, Summer 2013 Edition. URL = https://plato.stanford.edu/entries/public-reason/.

Rachels, James. 1986. *The End of Life: Euthanasia and Morality*. Oxford: Oxford University Press.

Rawls, John. 2005. *Political Liberalism*. Expanded Edition. New York: Columbia University Press.

Schwartzman, Micah. 2004. The Completeness of Public Reason. *Politics, Philosophy, & Economics*, 3: 191–220.

Shoemaker, David. 2010. The Insignificance of Personal Identity for Bioethics. *Bioethics*, 24: 481–89.

——. 2016. The Stony Metaphysical Heart of Animalism. In S. Blatti and P. F. Snowdon, eds., *Animalism: New Essays on Persons, Animals, and Identity*. Oxford: Oxford University Press. 303–28.

Vallier, Kevin. 2014. *Liberal Politics and Public Faith*. New York: Routledge.

Veatch, Robert. 1993. The Impending Collapse of the Whole-Brain Definition of Death. *Hastings Center Report*, 23: 18–24.

Wall, Steven. 2002. Is Public Justification Self-Defeating? *American Philosophical Quarterly*, 39: 385–94.

Williams, Andrew. 2000. The Alleged Incompleteness of Public Reason. *Res Publica*, 6: 199–211.

Williams, Jeremy. 2015. Public Reason and Prenatal Moral Status. *Journal of Ethics*, 19: 23–52.

——. 2016. War and Global Public Reason. *Utilitas*. DOI: 10.1017/S0953820816000376.

Zeiler, Kristin. 2009. Deadly Pluralism? Why Death-Concept, Death-Definition, Death-Criterion and Death-Test Pluralism Should be Allowed, Even Though It Creates Some Problems. *Bioethics*, 23: 450–59.

Theoretical Medicine and Bioethics (2022) 43:127–154
https://doi.org/10.1007/s11017-022-09583-2

How many ways can you die? Multiple biological deaths as a consequence of the multiple concepts of an organism

Piotr Grzegorz Nowak[1] · **Adrian Stencel**[1]

Accepted: 3 May 2022 / Published online: 20 July 2022
© The Author(s) 2022, corrected publication 2022

Abstract

According to the mainstream position in the bioethical definition of death debate, death is to be equated with the cessation of an organism. Given such a perspective, some bioethicists uphold the position that brain-dead patients are dead, while others claim that they are alive. Regardless of the specific opinion on the status of brain-dead patients, the mere bioethical concept of death, according to many bioethicists, has the merit of being unanimous and univocal, as well as grounded in biology. In the present article, we challenge such a thesis. We provide evidence that theoretical biology operates with a plurality of equally valid organismic concepts, which imply different conclusions regarding the organismal status of a brain-dead patient. Moreover, the theoretical biology concepts of an organism are very distant from the view on an organism that appears by way of bioethicists theorizing on death. We conclude that if death is to be understood as the cessation of an organism, there is no single correct answer to the question of whether a brain-dead patient is alive or dead.

Keywords Brain death · Organismal pluralism · Developmental concepts of an organism · Physiological concepts of an organism · Evolutionary concepts of an organism · Soul

✉ Piotr Grzegorz Nowak
piotr.grzegorz.nowak@uj.edu.pl

Adrian Stencel
adrian.stencel@uj.edu.pl

[1] Institute of Philosophy, Jagiellonian University, Grodzka 52, 31-044 Kraków, Poland

Introduction

Since the 1980s, it became a mainstream position in regulatory bioethics to define death "biologically," meaning "… the permanent cessation of functioning of the organism as a whole" [1] or by a President's Commission[1] as "that moment at which the body's physiological system ceases to constitute an integrated whole" [2, p. 33]. Both we and many bioethicists understand these definitions as simply equating death with the cessation of an organism. For example, Melissa Moschella, in reference to Bernat et al.'s definition, comments that "the early defenders…of neurological criteria for human death take biological integration to imply ontological wholeness (unity) and thus persistence of the human organism" [3]. According to this interpretation, "persistence of an organism" is synonymous with the persistence of "a living organism," and strictly speaking a dead organism is no longer an organism but rather a former organism. See Table 1 for more citations subscribing to such a view on the debate.

Besides defining death utilizing the notion of an organism, other proposals have also been present from the very beginning of the debate. For example, Robert Veatch put forward a moral idea, defining the "word death as the name applied to the category of beings who no longer have full moral standing as members of the human community."[8] (cf. [9]), while Michael Green and Daniel Wikler proposed identifying death with the cessation of personal identity [10]. However, these alternatives were evaluated as too vague to constitute the primary basis for the Uniform Determination of Death Act (UDDA) that was proposed in 1981. President's Commission, the author of UDDA, noticed that the concepts of moral standing and that of personal identity vary between different people, societies, and cultures [2, p. 39]. Therefore, one can "rely on them only as confirmatory of other views (i.e., biological views) in formulating a definition of death" [2, p. 39], while Bernat et al. states that the concept of "person" "is inherently vague. Death is a biological concept. Thus in a literal sense, death can be applied directly only to biological organisms and not to persons" [1]. This was later echoed in 2008 by the President's Council on Bioethics statement, according to which there are serious difficulties with John Lizza's [11, pp. 51–59] ideas, which resemble those of Veatch or Green and Wikler. The members of the Council stated that "one such difficulty is that there is no way to know that the 'specifically human powers' are irreversibly gone…" [12, p. 51].

It seems that many bioethicists believe that there is a single univocal and agreed-upon concept of an organism that corresponds to reality and the associated concept of biological death. For example, the conservative President's Council on Bioethics stated that "death is a single phenomenon marking the end of the life of a biological organism. Death is the definitive end of life and is something more complete and final than the mere loss of 'personhood'" [12, p. 52]. Meanwhile, the liberal thinker Peter Singer once asked a question that he intended to be a rhetorical one:

[1] In the United States, Presidential Commission is a special task force ordained by the President to complete specific, special investigation or research. They are often quasi-judicial in nature; that is, they include public or in-camera hearings.

Table 1 A sampling quotes identifying death with cessation of an organism

Author	Quote
D. Allan Shewmon	Even if (hypothetically) degree of integration *could* be meaningfully measured, there would be no point along that continuum that could reasonably nonarbitrarily constitute the dividing line between extremely sick, dying organisms, and *just-dead (non-)organisms* [29] (emphasis added)
James L. Bernat	In this article, I offer a refined account of the organism as a whole to more convincingly explain how its cessation spells death [5]
James L. Bernat, Charles M. Culver, Bernard Gert	We define death as the permanent cessation of functioning of the organism as a whole. We do not mean the whole organism, for example, the sum of its tissue and organ parts, but rather the highly complex interaction of its organ subsystems. The organism need not be whole or complete, it may have lost a limb or an organ (such as the spleen), but it still remains an organism [1]
Adam Omelianchuk	Bernat…asserted the loss of the organism itself is what matters. This assertion is deeply metaphysical because human death is linked to human organisms, not some special property of those organisms…. Nor does it permit there to be such things as dead organisms, or at least a dead organism as a whole. It also raises a pressing question: What apart from an organism's activity indicates that an organism as a whole exists? [6]
Maureen L. Condic	Of course, this [lack of rationality and global, self-integrated organismal function] in itself does not prove that a brain dead body is not a living human organism. More argumentation would be needed in order to show that (1) the capacity for global, self-integrated organismal function is necessary for the persistence of an organism [7]

"Isn't the distinction between life and death so basic that what counts as dead for a human being also counts as dead for a dog, a parrot, a prawn, an oyster, an oak, or a cabbage?" [13, p. 20]. Quite recently, Andrew Huang and James Bernat, stated that "*death is biologically univocal*" [30] (cf. [2, p. 31– 40; 3; 7; 13, p. 20; 14, pp. 59–85; 15, pp. 1–109; 16–19]), presupposing that all living beings are organisms in the same unified meaning and cease to be organisms in the same sense. These bioethicists and many more have all contributed to the view that we will subsequently call the "biological-bioethical" view on the nature of an organism and on the nature of death. This view will be contrasted with the theories of an organism and the associated concepts of death developed by theoretical biologists and philosophers of biology.

We emphasize in our investigation that, given the plurality of organismic concepts in theoretical biology [20–28], there is no such thing as a univocal biological or easily accessible sense of organism as a concept. Instead, there is a plurality of biological concepts of an organism, which implies that the cessation of an organism is not an idea that can be defined objectively but will rather depend on the concept used in a given situation. Indeed, if we can define an organism in many ways, then the state of "being dead" might vary between them.

Our first aim is to present the "bioethical-biological" concept of death and its implicit presuppositions on the theory of an organism as this concept has been elaborated by different bioethicists engaged in the definition of death debate. We then present the organismal pluralism within theoretical biology and the philosophy of

biology. Finally, we reach a conclusion regarding the plurality of biological deaths. We show that there are plenty of biological and "biological-bioethical" concepts of an organism which give different results on the status of brain-dead patients.

Different variants of "bioethical-biological" view on death and their implicit presuppositions on the theory of an organism

All "bioethical-biological" concepts that are of interest here are elaborated for the sake of determining the status of brain-dead patients. The most classical version proposed by Bernat et al. and the President's Commission does not imply anything more about an organism than that it is a whole that can exist if, and only if, its subsystems are functionally integrated. Regarding brain-dead patients, the authors of this concept simply stated that they are "merely a group of artificially maintained subsystems since the organism as a whole has ceased to function" [1]. The vagueness of the notions of "integrative functions", "integrative unity", and other synonyms were later utilized by Alan Shewmon to argue for the opposite thesis, namely that brain-dead patients are living organisms. Shewmon was the first to operationalize the vague "bioethical-biological" definition of an organism for the sake of resolving the brain-death controversy. According to him, one might be counted as belonging to the class of organisms if one poses a sufficient level of integrative unity, operationalized by two criteria:

> CRITERION 1. "Integrative unity" is possessed by a putative organism (i.e., it really is an organism) if the latter possesses at least one emergent, holistic-level property. A property of a composite is defined as "emergent" if it derives from the mutual interaction of the parts,…and as "holistic" if it is not predicable of any part or subset of parts but only of the entire composite.
>
> CRITERION 2. Any body requiring less technological assistance to maintain its vital functions than some other similar body that is nevertheless a living whole must possess at least as much robustness of integrative unity and hence also be a living whole. [4]

In Shewmon's view, brain-dead patients perform a "litany" of functions that fulfill criterion 1, such as maintaining homeostasis [4, 29]. Moreover, some of them, i.e., those who survive the acute period of spinal shock, fulfill criterion 2, requiring less technical assistance than some patients with high spinal cord transection. Patients in an acute phase of high spinal cord transection suffer from spinal shock and need much more artificial support than stabilized brain-dead patients. For example, they need medication to manage their bradycardia, while some stable brain-dead patients do not require such assistance. Since no one questions that conscious patients with high spinal cord transection are living organisms, the conclusion follows, according to Shewmon, that brain-dead patients are living organisms as well.

Suppose Shewmon's proposal is to be evaluated as a universal biological definition of an organism. In that case, it has an obvious drawback: due to criterion 2, it presupposes at the outset that we know that some groups of patients count as living

organisms [9]. Despite this potential drawback, Shewmon made a significant contribution to the bioethical debate since his analysis compelled bioethicists to say something more about organisms than that they are entities that are functionally integrated. Recently, bioethicists have pursued numerous attempts to clarify the concept of an organism associated with the definition of death debate. For example, Bernat stated that intuitions have an essential role in distinguishing the class of organisms from non-organisms, such as organisms' parts:

> People intuitively grasp that while many parts of the technologically supported brain-dead patient remain alive, the patient has died. The essence of this intuition is the recognition of the fundamental distinction between the life status of an organism's parts and of its whole. As dramatically shown by the examples of ex vivo cell cultures and tissue and organ transplantation, parts of the human organism can be kept alive for prolonged periods by technology after the organism has died [5].

According to Bernat, intuition plays a crucial role in distinguishing organisms from non-organisms, not only in lay people but also among the scientists engaged in the debate. It is impossible, according to him, to provide a uniform definition of an organism as a whole that is neither too strict nor too broad. He confesses that "the inescapable conclusion is that all members of the immense diversity of life forms cannot be neatly separated into distinct categories delineated by specific criteria that correctly and comprehensively classify them into either living or nonliving categories" [5].

In another recent work coauthored with Andrew Huang, Huang and Bernat notice that there are two incompatible intuitions about human death: on the one hand, we intuit that a human dies in the same sense as other living organisms, but on the other hand, we believe that there is something peculiar in the human way of ceasing to exist, that is, we intuit that patients without any residual consciousness are gone, even though they might be capable of performing many physiological functions, such as spontaneous breathing for example [30].

Huang and Bernat distinguish between a concept and the conception of death. The conception is general and based on a vague notion of an organism as a whole as an integrated, complex entity, possessing some emergent functions, being capable of combatting entropy, and possessing a common ontogenetic. Such a general concept might become more precise in delineating different organisms' lives and deaths when we identify the "most macroscopic unifying and integrating emergent functions" [30] of a given type of organism as a whole. In this way, it is possible to obtain a conception of an organism as a whole. In humans, Huang and Bernat state that the crucial function is to be identified with neurological control over consciousness and breathing. A supposed upshot is that a brain-dead patient, a patient with a compromised ability to provide neurological control over respiration and consciousness, has ceased to be a human organism, or at least, has ceased to be an organism as a whole.

Huang and Bernat's concept seems to be best suited to the organismal pluralism of all "bioethical-biological" views. However, even these authors do not acknowledge or discuss their ideas in an organismal pluralism context–this is the general

problem with the view. The specific problem with Huang and Bernat's model is that they seem to simply ignore other "macroscopic unifying and integrating emergent functions" such as the capacity to fight infections or an ability to digest and assimilate resources without which breathing and consciousness fade away, which are for us as much intuitive aspect of human organisms like the one mentioned by them. To be more convincing, Huang and Bernat should analyze intuitions much more closely and justify them as reliable instruments in accessing biological reality. Yet, this seems implausible since it is the culture that mainly shapes the intuitions that Huang and Bernat refer to [31, 32]. In particular, intuitions about the status of brain-dead patients are shaped somewhat by the most famous bioethical texts, including the one authored by Bernat in the 1980s [1]. For this reason, Huang and Bernat's appeal to intuition is a matter of begging the question: through their recent work utilizing intuitions about organisms, they are trying to defend the concept of death as formulated by Bernat et al. in the 1980s, which itself shaped our intuitions on organisms.

Another idea which helps to pinpoint the "bioethical-biological" concept of an organism comes from Melissa Moschella. She is one of the scholars who have tried to explain the notion of an organism through the Aristotelian-Thomistic theory of the soul:

> A putative organism really is an organism if it possesses the *root capacity for self-integration*. Possession of the root capacity for self-integration (of which the soul is the principle) is evidenced [in humans and other sentient animals] by (1) possession of the material basis of the capacity for self-integration—i.e., the capacity for control of respiration and circulation—or (2) possession of the material basis of the capacity for sentience. [18]

Moschella argues that integrated functioning manifested, for example, by the maintenance of homeostasis of the brain-dead body, results from artificial support and does not count as self-integration caused by the soul. She believes that so-called "root capacities" for consciousness and spontaneous breathing are dependent on the brain in adult human beings, so a brain-dead body on artificial support is a dead organism. However, it is hard to understand why (1) and (2) are to be treated as the only proofs of the root capacity for self-integration. Why is the joint function of organs, such as the kidneys, lungs, hypothalamus, posterior pituitary, pancreas, adrenal glands, parathyroid glands, bone, liver, intestines, the bicarbonate buffer system within the extracellular fluid, and the hemoglobin buffer in red blood cells in maintaining homeostasis not perceived as self-integration? [33, 34] Clearly, brain-dead patients would not survive artificial support being turned off, even with all the organs mentioned above undamaged. Still, neither could a patient with a functioning brain and dysfunctional kidneys survive without dialysis or kidney transplantation.

Moreover, even though brain-dead patients cannot regain spontaneous breathing at the current level of the development of medical technologies, they might still have the "root capacity" for spontaneous breathing in Moschella's sense and possess a human soul. They might be like spinal cord intersection patients where, according to Maureen Condic, "the organizing principle of the body [i.e. the soul] must persist (otherwise the individual would be dead), but the full function of this principle is blocked by an injury-induced material deficiency" [7]. Just as spinal cord

intersection patients are living human organisms with the impeded ability to breathe independently, so might be brain-dead patients. They might be capable of recovering spontaneous breath if the physiological obstacles are removed. Currently, scientists can grow mini-brain organoids from stem cells [35]. Perhaps it is physiologically possible to grow a standard size human brain in this way, together with a functioning brainstem and then transplant it into the brain-dead body. Provided that such a brain would be grown from the patient's cells, genetically it would be the patient's brain, even though it would be a *tabula rasa* with no mental content. If something like this is physiologically possible, although not yet feasible, it would mean that brain-dead patients *today* have a root capacity for spontaneous respiration, *a fortiori* they have a root capacity for self-integration and a rational human soul.

These conclusions, which are undesired by Moschella, could be avoided if we look at the other model of organism unity enshrined in another of her works. According to this view, each organism exists so long as its master part persists. Such a part is understood as "the vital, essential part that has the biological function of controlling all of the organism's parts, directly or indirectly" [3], (cf. [36]). The view according to which a master part is a sine qua non condition for the existence of each organism is reasonable, since as Hoffman and Rosenkrantz argue, all known organisms have a master part [37]. According to Moschella, it is beyond controversy that the central nervous system constitutes a master part of adult human organisms. Therefore, humans without functioning brains no longer form human organisms.

A structurally similar approach to Moschella's first idea of defining an organism in terms of the function of the "self" was adopted even earlier by the President's Council of Bioethics. It was enshrined in position no. 2 of the "Controversies in Determination of Death" report. The Council pointed out that the concept of death developed in the 1980s by Bernat et al. and the President's Commission was right in perceiving an organism as *a whole*. However, it was wrong to interpret an organism's wholeness as functional or somatic integration [12, pp. 59–60]. Instead, the President's Commission proposed to define "organisms as a whole" as entities, capable of performing "the work of self-preservation, achieved through the organism's need-driven commerce with the surrounding world" [12]. In turn, the capability to perform this vital work was interpreted as being dependent on three "fundamental capacities":

> 1. Openness to the world, that is, receptivity to stimuli and signals from the surrounding environment. 2. The ability to act upon the world to obtain selectively what it needs. 3. The basic felt need that drives the organism to act as it must, to obtain what it needs and what its openness reveals to be available. [12, p. 61]

Many commentators have already noted that this concept of an "organism as a whole" is more unclear, underspecified, and nonscientific than the view formulated in 1981 by Bernat et al. and the President's Commission [17; 34; 38; pp. 72–75; 39]. The most severe problem with this definition of an "organism as a whole" is that it is circular: it defines an organism as a whole in terms of self-preservation. However, what is at stake is indeed the self-preservation of *an organism as a whole*, so we need first to know what an "organism as a whole"

is before considering whether it can perform "the vital work" of preserving itself. However, the President's Commission does not explain their concept of an "organism as a whole" further.

Recently, Adam Omelianchuk proposed a strategy to address this gap [6]. According to him, the work of self-preservation should be interpreted as a "second-order capacity (viz. a capacity for having a capacity) for self-movement towards species-specific ends" [6]. Note that the notion of "second-order capacity" is quite similar to Moschella's idea of "root capacity."

In our view, such a defense of the President's Council idea is problematic due to several reasons. First, the idea that there are species-specific ends is based on an Aristotelian metaphysics which, with its final causes, soul, entelechies, and so on, is foreign to contemporary natural science, [cf. 34]. Second, even if we agree for the sake of argument that there are final causes, and so are species-specific ends, moreover, if we agree that the distinctively human end is that of rational thought and action, it is hard to understand why Omelianchuk upholds that anencephalic newborns and persistent vegetative patients (PVS) are alive while brain-dead patients are dead. What is the difference between these groups of patients regarding the second-order capacity for rational thought and action? To justify this distinction, Omelianchuk states that anencephaly and PVS are only disorders that impede the capacity for rational thought and action. At the same time, brain death is more than an impediment. It is the destruction of the second-order capacity for the achievement of human ends.

Yet, just as there is currently no therapy that could help the brain-dead regain consciousness, there is no treatment that might help anencephalic newborns develop consciousness either. Thus, it is hard to understand why anencephaly impedes the second-order capacity for rationality while brain death destroys it. Moreover, we can recall the case of brains grown from stem cells discussed in the context of Moschella's view. If it is physiologically possible to grow brains from a brain-dead patient's steam cells, perhaps these patients have not lost the second-order capacity for rational thought and action. If this is the case, brain death is only an impediment to the second-order capacity of being conscious.

In terms of its scientific background, the most promising "bioethical-biological" concept of an organism utilizes the modern scientific notions of homeostasis and entropy. Julius Korein was the first in the context of the definition of death debate to define organisms as open systems that tend to minimalize their own entropy and maximize their negentropy at the cost of increasing entropy in the environment [40]. Much more recently, Michael Nair-Collins defined organisms in the following manner:

> Living organisms are localized pockets of anti-entropy, achieved by mutually interdependent functional structures jointly maintaining internal equilibrium, or homeostasis of the extracellular fluid, a necessary condition for all organismic function, while resisting chemical and thermal equilibrium with the external environment. Second, living organisms are a social collective, consisting of trillions of cells working together to actively maintain their environment within conditions suitable for their continued functioning and existence. [34]

Given such a definition of an organism, Nair-Collins noticed that entropy and homeostasis are inversely related and identified death with "the irreversible cessation of the organismic capacity to maintain homeostasis of the extracellular fluid and thereby resist entropy" [34]. While Nair-Collins believes that brain death does not constitute the death of a human organism, Korein holds the adverse opinion. According to him, it was impossible to maintain the functioning of a mature brain-dead human body for a period longer than a week. We know today that this statement has proven to be false [41].

The entropic-homeostatic concept of an organism has certain merits over the classic "biological-bioethical" view given its precise nature. "Entropy" is not a nebulous concept like "integrated functioning," and its change is formally operationalized by scientists as the measurement of the dispersion of energy at a stated temperature [42]. According to Nair-Collins, the concept has the essential merit of being a "part of a coherent, unified story of the world and our place in it, drawing on a well unified ontology within a mechanistic explanatory framework" [34]. Perhaps it is the part of the "unified story of the world," yet we are afraid that it might be neither a story of "*our* place" in the world nor even the story of organisms, but the more general story of living matter.

The entropic-homeostatic concept tells us only that if we want to know whether x (still) constitutes a living organism, we should consider whether x might be capable of maintaining homeostasis and thereby resisting entropy. Yet, it gives us no clues as to how we might solve the crucial issue in the philosophy of biology of recognizing the borders between different organisms, and it provides no instruments for differentiating organisms from their parts [43]. Given this theory, we have no advice on where to find the borders of some person's organism. Suppose we consider the borders of Adam's organism. Are these borders coextensive with the commonsense notion of Adam's body? Or perhaps the boundaries coincide with the commonsense idea of Adam's body plus his microbiota? A different option is to identify an individual organism as constituted by the objects we term the commonsense notion of Adam and Bill's bodies. Yet another idea is that the individual organism might be formed solely by Adam's kidneys. Biological concepts, such as the immunological, or zygotic concepts, can provide answers to such questions (see the next section), while the entropic-homeostatic view is incapable of serving this purpose [cf. 43]. This concept is perhaps a fragment of the true story of the world. However, it is not a story about individual organisms but rather one about the difference between living and inanimate material. Since it cannot facilitate the distinction between an organism and its parts, it cannot answer the question of whether a brain-dead body constitutes an individual organism or perhaps only a part of an organism, albeit a large one [cf. 3, 18, 19]. The concept can only confirm the obvious truth, namely that brain dead patients on artificial support are part of the living world as opposed to inanimate material such as rocks.

It seems that while looking for a scientific theory of the death of an organism, followers of the homeostatic-entropic concept have focused on the wrong part of science. To see that this is the case, let us notice that the idea of engaging the concept of entropy in theoretical biology investigations originally comes from Erwin

Schrödinger's book "What is life?". The question settled in the title of the book was answered by Schrödinger in the following way:

> When is a piece of matter said to be alive? When it goes on 'doing something', moving, exchanging material with its environment, and so forth, and that for a much longer period than we would expect an inanimate piece of matter to 'keep going' under similar circumstances. When a system that is not alive is isolated or placed in a uniform environment, all motion usually comes to a standstill very soon as a result of various kinds of friction; differences of electric or chemical potential are equalized, substances which tend to form a chemical compound do so, temperature becomes uniform by heat conduction. [44, p. 69]

Although Schrödinger writes in his book that organisms are alive, meaning that they avoid decay through exchanging material with their environment, minimalizing their inner entropy through metabolism, we believe that organisms are only particular examples of living systems. In addition, parts of organisms such as the above-mentioned kidneys or even cells are alive in Schrödinger's sense. Moreover, the whole ecosystem of Earth, as opposed to Mars, might be perceived as alive, taking into account the quotation from Schrödinger. Schrödinger aimed to grasp the general nature of life instead of providing a comprehensive theory of an organism [45].

The above discussion on the entropic criterion reinforces our point that if the discussion in bioethics on a definition of death is to be interpreted as the discussion on the requirements an individual organism must fulfill to go out of existence, it needs more reference to contemporary work in the philosophy of biology. This is because it conflates two questions: (i) how to define organisms [20–25, 27, 28, 46]; and (ii) how to define life [47–49]. The two often seem to be confused. From the point of view of the problem discussed in this paper, only the first question seems to be relevant because brain death obviously does not transform organisms into non-living matter. No one would argue that cells and the human body always become non-living matter after brain death–certainly not in the time frame that bioethicists are interested in. The question is to realize whether the link between elements of the human organism is broken to the extent that it ceases to exist. Thus, to answer that type of question we have to understand what an organism is, rather than what is life. In other words, living matter might still be part of something that used to be an organism. Thus, figuring out whether something is alive does not set up the debate as to whether it is an organism. Ant societies are undoubtedly alive, but this does not help us to realize whether they are individuals or a group of individuals [50, 51] and that is why the two questions are separate in the philosophy of biology.

The plurality of concepts of an organism in theoretical biology

The status and meaning of the concept of an organism is one of the greatest issues to have been raised in the philosophy of biology in recent years. Despite the fact that a bioethical consensus on the definition of death was established in the 1980s, the emphasis in theoretical biology throughout the 20th century was placed on genes

and other sub-elements of cells rather than on the concept of an organism [52]. This was partially caused by the fact that during this time, biologists had unraveled the mystery of DNA and learned a lot about the molecular mechanisms of many traits. Thus, the very concept of an organism in theoretical biology was somehow put aside. However, this has changed recently as an increased focus has been placed on understanding "organisms as a whole," a concept which is also crucial for those engaged in the definition of death debate [cf. e.g., 1; 2, pp. 1–75; 5; 6; 12, p. 1–121; 30]. So, what is an organism from the philosophy of biology point of view? What conditions does something have to fulfill to be considered an organism? Can we unequivocally put forward conditions to call something an organism? Many researchers have tried to tackle this issue and alternative approaches to the concept of organisms have been proposed over the years. For instance, Ellen Clarke [28] counted at least thirteen concepts of an organism in use in 2010. Given the explosion of interest in the topic in recent years, with many papers published [23, 25, 51, 53–55], special issues edited, and conferences organized on this subject, one might expect that the number of concepts has at least doubled.

Let us present a few of the most popular concepts found in philosophy and biology that will fuel our further discussion. We selected a number of concepts from different fields where scientists pursue different goals (e.g., development, physiology, evolutionary biology), in order to show the diversity of concepts that exist in biology.

The most classic concept of the organism in theoretical biology is called the developmental concept of an organism. The concept has been around for about 170 years and was put forward even before the publication of Darwin's famous work. It was T.H. Huxley who wrote that: "the individual animal is the sum of the phenomena presented by a single life: in other words, it is, all those animal forms which proceed from a single egg taken together" [56]. The developmental concept of an organism is an enduring one and one of the most popular ways of defining the organism by people working on development. For instance, Gilbert et al. defined it in the following way: "the individual animal proposed here is understood to be that which proceeds from ovum to ovum" [22]. Similarly, Moore et al. stated that "human development begins at fertilization when an oocyte (ovum) from a female is fertilized by a sperm (spermatozoon) from a male. Development involves many changes that transform a single cell, the zygote, into a multicellular human being" [57, p. 1].

Thus, this concept emphasized development as the process that marks the difference between two individuals. Here, organisms come from a fertilized egg and consist of all the cells that make up its body, like muscle cells, nerve cells, or the cells that build the digestive tract. All those cells coming from the fertilized egg constitute a developmental organism until the next process of fertilization, which marks the emergence of another individual. The concept seems to be quite correct: you and your friends are different individuals because you all developed from different fertilized eggs. However, in other cases, it generates quite strange, non-intuitive individualization. For instance, Janzen [58] argued in a famous paper that if an organism reproduced from an unfertilized egg, as is the case with many species (dandelions, fungus, aphids), then this should not be considered as a process of generating a

new organism, but only growth because there is a lack of a sexual event of fertilization that marks the difference between two individuals. Thus, all aphids that grow in a given meadow from an unfertilized egg should be considered one organism, albeit one that is physically disconnected across the meadow. This would mean that twins from a zygote that at some point has undergone mitotic division would also be considered a single individual because there was no sexual event of fertilization to mark the differences between them. Note that bioethicists engaged in the brain death debate sometimes take it for granted that twins cannot be perceived as a single organism [3], which seems to be a proper approach for carrying on bioethical considerations, as twins are so disintegrated in many dimensions that it is justified to treat them as separate units for bioethical investigations.

If we try to look at this concept from the perspective of the bioethical debate, then we realize that it implies a certain way of thinking about the status of brain-dead patients. This is mainly because, since all the elements of such patients are derived from a fertilized egg, they should be considered developmental organisms. Therefore, we should consider it as an organism, even if it starts to decay. Furthermore, even if the body of an organism is turned into dust, it can be still alive if its twin is around–since they constitute a single organism coming from a fertilized egg.

The peculiar consequences of this concept that link such different, physically disconnected elements into a single organism have led scholars to propose an alternative that emphasizes functional integration, one that we can call the functional developmental concept of organisms [25]. The main conceptual problem of the classical developmental concept of an organism was the presupposition that all life forms coming from the fertilized eggs are functionally integrated. It might not always be the case, however. If we explicitly assume that there must be some sort of functional integration among elements that come from a fertilized egg, then this excludes cases like twins–they are rarely functionally integrated in any way. Furthermore, it excludes cases of considering the decaying body as an organism as well, as such an aggregation of cells is not functionally integrated. At the same time, this would deliver a different verdict on the status of brain-dead patients. Suppose they are functionally integrated due to some medical equipment provided by doctors, as is often the case with intensive care units. In that case, they should no longer be considered organisms–because those elements do not come from fertilized eggs, while this kind of origin is a sine qua non condition for the existence of an individual organism, given all variants of the developmental concept of an organism. Therefore, conscious patients who are dependent on pacemakers or transplanted organs in their integrated functioning are also not organisms, but rather something akin to cyborgs.

The functional developmental concept of an organism operates with the notion of functional integration, one which is familiar in the bioethical definition of death discourse. For example, Bernat et al. define an organism functioning as a whole to be an entity whose "spontaneous and innate activities" are carried out by "the integration of all or most of its subsystems," and is capable of "at least limited response to the environment" [1]. Yet, despite several attempts to operationalize it, the notion of "integrated functioning" (see the previous section for discussion), remains "undefined and vague in the views of those who attempt to

define death" [59]. The functional developmental concept of an organism is not different here in this matter–it understands the term intuitively and without elaborating it.

Is the notion of functional integration always as unspecified within theoretical biology as in the case of the developmental concept of an organism? To answer this question, we will discuss another concept that fundamentally relies on functional relations—namely the physiological concept of an organism [e.g., 22, 23, 27, 60]. According to the physiological conception of individuality, if a group of entities engages in a significant amount of physiological interactions with one another, then the group of entities will be considered a physiological individual. In contrast to the developmental concept of an organism, this view does not require that elements making up an organism go through a certain type of development. The origin of the elements of the organism is not very important, but what matters is the existence of certain functional relations. Indeed, the physiological concept of the organism focuses rather on the certain mechanisms of cohesion that make a group of elements a single unit, rather than a group of single units.

As good as it sounds, this idea is approached by scholars in different ways [e.g., 23, 60, 61], and one approach that seems to provide good criteria for the fuzzy term of "functional integrations" is the immunological concept of an organism [23, 61, 62 pp. 239–269]. This approach focuses on immunological properties as the main drivers responsible for setting the boundaries of physiological individuality. Traditionally, the immunological conception of individuality has assumed that the immunological system acts as a gatekeeper that determines the boundaries between the self and the non-self by triggering an immune response in order to eliminate any possible intruders. Self-elements are those that do not trigger an immune response while non-self-ones are those that do. Furthermore, the elements that belong to the self are generally considered those that come from "inside" (i.e., from the zygote) [23]. Thus, the distinction between the self and non-self is quite obvious.

More recently, the immunological view of individuality has emphasized that immune responses are more diverse, and the boundaries set by the immunological view are considerably more dynamic [see 23, 61]. Firstly, constituents that come from the zygote can trigger an immune response even in healthy people [63, 64]. Secondly, there are elements that do not come from the zygote but are tolerated by the immune system, such as symbiotic microbes [65]. This leads to the idea that immunological individuality should be conceived in a more dynamic fashion [23, 61]. This implies that in the context of this concept a given element of the organism (like a nervous cell) might one day be considered part of an organism because it is tolerated by the immune system. At the same time, a few months later, it might be excluded as an element of the organism if it is not tolerated by the immune system anymore. This might transpire because, for instance, a disease such as cancer [66] changes the immunological properties of the immunological system as it does not tolerate some cells anymore. The immunological system defines the boundaries of the organism, and as long as there is an immunological system, as long as an organism exists, its constituents might change dynamically. In other words, whether some elements are part of a given individual should not be based on their origin (from the zygote vs outside), but rather the emphasis

should be placed on their tolerance. If they are tolerated by the immune system, then they are part of the organism. This was summarized nicely by Pradeu:

> Immunological criterion suggests that any entity which interacts regularly with the immune system and is not eliminated by it is part of the physiological individual. In other words, the physiological individual, immunologically, is the unit made of the association of a host and many microbes (those that are tolerated by the immune system. [23]

If we consider the bioethical definition of death debate from the point of view of the immunological concept of an organism there are several interesting takeaways. First, it seems that participants in the bioethical definition of death debate do not account for the immunological concept in their investigations. Their conclusions quite often clash with the conclusions of proponents of the immunological view of an organism. For example, in contrast with the above quotation from Pradeu, Melissa Moschella writes that:

> Both termite and protozoans live within an enclosed membrane in a complex symbiotic relationship, dependent upon each other for survival, functioning in a coordinated manner in the service of a larger whole. Yet they do not constitute a single organism. The protozoans are not parts of the termite. Rather, each protozoan is itself an organism, distinct from the termite. Any plausible account of organismal unity must be fine-grained enough to explain cases like this one. [19 cf. 3, 36]

A second important thing to note is that, given the immunological view, brain-dead patients are living organisms. That is because they are capable of "fighting of infections and foreign bodies through interactions among the immune system, lymphatics, bone marrow, and microvasculature" and the "development of a febrile response to infection" [4]. Finally, the third factor which might have some influence in bioethical contexts other than in the definition of death debate, is the fact that transplanted organs are not genuine body parts of the recipient since they are not recognized as such by their immunological systems [67]. It is an interesting upshot since it raises questions regarding the content of the moral right to bodily integrity. Do we always use the phrase "human body" in a manner synonymous with the biological meaning of a "human organism"? For the moment, we place this fascinating question to one side.

We can now move to the evolutionary concept of an organism–one which is supposed to capture what constitutes an organism from the evolutionary perspective. In other words, it is supposed to state when a given individual is a unit that undergoes evolution by natural selection (ENS). Classically, as indicated by Lewontin [68], for a population to undergo ENS its members must be characterized by variance, fitness differences, and heritability. This classic formula has recently been elaborated in detail [69, 70]. According to this elaborated view, evolutionary individuals (Darwinian individuals) are units that are capable of reproduction. In other words, an evolutionary individual is every unit that is capable of producing offspring. This is an important factor since reproducers

are causally responsible for parent–offspring similarity (fulfilling the heritability criterion mentioned above). Thus, if you have a group of reproducers that vary in some traits, those traits influence their fitness (number of offspring), and those traits are heritable, then you can expect that the population will undergo natural selection because some reproducers will produce more offspring than others and, as a result, their frequency will change in a population.

Very diverse types of reproducers exist in nature [71, pp. 87–109, 72] and three paradigm cases of reproducers can be distinguished. The first is a scaffolded reproducer, and they are characterized by the fact that their reproduction is entirely dependent on external machinery. For instance, viruses belong to this category because they use cells to reproduce. The second category consists of simple reproducers. Simple reproducers only need external resources to initiate reproduction, (e.g., a bacterial cell). The third category constitutes collective reproducers which are built of simple reproducers. In other words, a collective reproducer is an entity that can reproduce itself, but which is also built of elements that can reproduce themselves. An example would be multicellular individuals built of eukaryotic cells.

The above concept of individuality based on the theory of evolution is not the only concept present in the literature. There are other concepts that use the theory of evolution to single out organisms from the environment to a greater or lesser extent. One can list here the replicator-vehicle/interactor framework [73, 74], the concept of Organismality [72], "It's a Song Not a Singer" [75], the concept of Stability of Traits [26] or the the concept of Unity of Purpose [76, pp.43–72]. The latter frame organisms in terms of agency and will be discussed here. Agency is a very important metaphor in considerations of the theory of evolution. It treats organisms as agents similar to human beings. Mainly, as units that have traits that help them pursue some goals. For instance, it is uncontroversial to say that a peacock's tail has evolved to attract mates. Indeed, we can easily ascribe functions to the majority of biological traits and show that the function of those traits is to increase the fitness of their bearers. This led Okasha [76, pp. 4372] to define organisms along those lines. According to his view, a given unit is mainly an evolutionary agent–an organism from the evolutionary perspective–if it possesses the "unity-of-purpose." In other words, the different traits of an organism have evolved because of their contribution to the same goal–enhancing bearer fitness.

Even though it might seem that evolutionary agents and Darwinian individuals are quite the same, once we delve into the details it is not so simple. A good case to illustrate this would be the part of noncoding-DNA that people sometimes term "junk DNA". This is a part of DNA that does not perform any function–although some junk DNA might turn out to be functional in future studies–however, it is replicated with the rest of the genome during cell divisions [77]. If one considers these to be Darwinian individuals, then one might come to the conclusion that junk DNA is part of its bearer's individuality–because it is replicated together with the rest of the genome, it constitutes part of the genome. For instance, when humans reproduce, we transfer that DNA with all our other genes. At the same time, junk DNA does not perform any function that benefits the fitness of the bearer; thus, it would be hard to say that it contributes to the unity-of-purpose of a human being as it does not benefit the bearer's fitness. Indeed, it seems the junk DNA has not evolved to

enhance the fitness of the organism and thus it would be uncontroversial to say that it does not constitute part of a human organism if the organism is understood as an evolutionary agent.

Two approaches based on evolutionary considerations will differ as well in their verdict about the status of brain-dead patients [78]. If we follow the Godfrey-Smith approach pointed out above, after brain death, the nature of the reproducer changes. The collective reproducer (i.e., a human being) becomes a scaffolded reproducer, as its reproduction becomes dependent on medical equipment and other reproducers (i.e., doctors in charge of the equipment)–so the brain-dead patients eventually resembles cellular organelles like mitochondria or chloroplast that used to be free-living entities, but for now can only reproduce with the assistance of the cellular machinery of eukaryotic cells. Alternatively, someone might take a different route and argue that being unable to reproduce does not rule out the status of "evolutionary individual." One can zoom out and argue that individuals that cannot reproduce are still evolutionary individuals if they are part of an evolving population. After all, nature is full of individuals that are unable to reproduce for one reason or another, like mules. So as long as the unit is part of an evolving population, it is an evolutionary individual. This point was elaborated in detail by Chodasewicz [79].

The agency view of the organism would lead us to a different conclusion. This view says that a given unit is an evolutionary agent if its different traits have evolved because of their contribution to the same goal–enhancing its bearer's fitness. At first glance, it might seem that this property is still present in brain-dead patients because several capacities which promote an organism's fitness are present in brain-dead patients. Besides the capability of generating an immune response to infections, they are also capable of: maintaining homeostasis; the elimination, detoxication, and recycling of cellular wastes; maintaining energy balance and body temperature; wound healing; cardiovascular and hormonal stress responses to unanesthetized incision; sexual maturation; and proportional growth [4]. Moreover, the undamaged spinal cord in brain-dead patients is capable of performing some integrative functions and even of primitive sensorimotor learning, which might manifest itself in an extreme form by means of the Lazarus sign [4]. Of course, all those works only contribute to enhancing fitness if they are assisted with medical equipment, which triggers another issue. However, unity-of-purpose is obtained only if *different traits have evolved* to obtain the same goal. The problem is that medical equipment is not a property that has evolved to enhance fitness; it is a tool made by humans, helpful for enhancing fitness, as is the case with many other tools, but not a phenotype trait that has evolved over thousands of generations in the same way that the eye or the brain have done. Therefore, if the organism must rely on it, it seems it should not be considered an evolutionary agent, because it cannot obtain its goal solely on the basis of the traits that have evolved to do so. However, it seems the same conclusion should follow for patients with functioning brains that rely on their functioning and reproduction on any kind of artificial technology.

This, of course, is only the consequence if we assume that absolute unity-of-purpose is required. In fact, the-unity-of purpose quite often breaks down due to the fact that conflicts within organisms emerge. As a result, Okasha [76, pp. 43–72] argued that we should accept that sometimes unity-of-purpose breaks down because

some traits evolve in a neutral manner or to benefit some element below the organism level, like genes for example. Furthermore, these biases from the unity-of-purpose seems to be widespread in nature [80, pp. 1–18]. So, if unity-of-purpose breaks down at times and does not exclude something as an evolutionary agent, it might sometimes be the case that the evolutionary agent is sustained by "traits" that have not evolved for this reason. Indeed, perhaps we can say that sometimes some traits did not evolve to sustain the unity of purpose, but they do so currently, like artificial technology. We think it is a justified position and interesting issue to be explored, but of course, this requires being discussed in the framework of Okasha, and we do not have time for this here. Yet if this is so, then people that are supported by medical equipment, including those with destroyed brains, would be considered evolutionary agents.

To sum up, in biology and philosophy of biology we do not have a universal concept of an organism. The concept of an organism is equivocal and unclear–rather it may be more accurate to say that multiple concepts coexist. The same is true for the concept of death, which might be understood differently depending on the concept of an organism adopted. Furthermore, it is sometimes not even clear how we should understand the concept of death if we stick to a single concept, which comes from the lack of proper analysis of the problem. We have presented how death should be understood in the context of different concepts, but note that as the topic is unexplored, we might be wrong in many places. Our aim was not to provide a definitive answer, but rather to show that neither the notion of an organism nor death is clear. All this shows that philosophy of biology and biology, just like bioethics, cannot currently deliver either a universal concept of an organism or one of death. However, this does not necessarily mean that the concept of an organism is waiting to be developed. Perhaps pluralism is to be expected? This issue and other related ones will be discussed in the next section.

Four remedies for the gap between bioethics and theoretical biology

As evidenced in the prior section, there are multiple concepts of an organism in the philosophy of biology that give a different verdict about death. There are also numerous versions of the "bioethical-biological" concept of death and an organism (see above). Besides the notion of functional integration and its synonyms, the biological and "bioethical-bioethical" theorizing on organisms differ quite substantially. Furthermore, theoretical biologists are relatively aware of the plurality of the biological concepts of an organism and the fact that it might be hard to deliver a universal concept of the organism. Meanwhile, bioethicists engaged in the definition of death debate seem to avoid taking this into account while developing their arguments. An upshot of this is that there is a considerable gap between the investigations of theoretical biology on the nature of organisms and the bioethical understanding of the issue.

In our view, bioethicists engaged within the definition of death debate in the present situation might consider four options: (1) they might try to show that pluralism is simply wrong and that there is a universal concept of the organism and perhaps even

one of the concepts developed by them can serve such a role; (2) they might accept pluralism about the concept of the organism and subscribe to one of the actual biological views as described herein, and accept their conclusions on the status of brain-dead patients; (3) they might take organismal pluralism at face value and defend their concept of an organism as one of many views of an organism which is valid within some subdiscipline of natural sciences, namely within some subcategory of medical sciences; finally (4) they might claim that the death of an organism is a phenomenon distinct from the cessation of an organism. The rest of this section discusses these four options. Whatever bioethicists choose, they can no longer continue with a strategy of "isolation" from contemporary ideas about organisms. We think as well that biologists and philosophers of biology can gain much from the debate about organisms in bioethics, as some of the concepts found here, such as the one developed by Moschella [3], differ substantially from biology and so might provide interesting insights. This cross-fertilization can certainly benefit both, but we will not discuss the potential influence of bioethical concepts on the philosophy of biology as it is not the aim of the paper.

The option (1) of the four mentioned above seems the one which is the most implausible. Its adoption would require the refutation of the dominant position of pluralism in the philosophy of biology [20, 22–25, 27, 28, 46]. Furthermore, it would require developing a concept of an organism that would be universal or showing that one of the concepts found in bioethics or biology can serve such a role. While this might be possible, it would require a lot of effort, and thus if one wants to claim the superiority of a given concept of an organism, one would have to show that organismal pluralism can be replaced by monism. This is a difficult and ambitious project to pursue and to date there have been very few attempts. One well-known example is the work of Ellen Clarke [53] who tried to provide such an account by arguing that all concepts can be reduced to a single one. In such a view, there are a "set of conditions," meaning, in her view, the existence of "policing" and "demarcation" mechanisms, that, if fulfilled, single out organisms. However, those conditions are realized by different mechanisms in different lineages due to the different evolutionary history of species. Thus, multiple concepts of an organism exist, because people confuse "conditions" with the mechanisms through which they are realized. Once the distinction is made, we can turn pluralism into monism. There exist universal conditions for distinguishing organisms that are realized by various biological mechanisms.

Technically, her idea seems very convincing. However, philosophically speaking, the idea is very problematic. How does one single out conditions? What makes one feel that a given set of conditions is the most appropriate? Clarke [53] believes that evolutionary theory unifies biology, and therefore conditions should be based on it. Sadly, there are two problems with this position. Firstly, it is unclear whether bioethicists would be happy with the idea that their discussion of death should be rooted in the theory of evolution since this theory also does not provide a unanimous verdict on the status of brain-dead patients (see above section). If they are not, they would have to show that some of the conditions they prefer are superior to evolutionary ones. Secondly, many authors have suggested that different conditions are equal to each other [23, 25, 81], so one would have to provide a philosophical argument against pluralism. Of course, one can try to build such an argument, for instance following Clarke [53] and argue that for epistemological reasons (like

counting organisms) such a concept is necessary. Furthermore, one might try to refer to more ontological arguments, for instance, to Ockham's Razor, which requires that we do not multiply entities without necessity–a reference to this principle is very popular in the philosophy of biology [82, pp. 153–239]. However, thus far such a concept and the philosophical arguments to back it up have not been put forward in such a compelling way to convince the majority. As a result, pluralism is the dominant position [23, 25, 51, 53–55] toward which we are sympathetic as well. To sum up, if bioethics wants to defend the monistic view of organisms, then this is a legitimate position, but numerous philosophical investigations are needed to develop and defend it. Indeed, monism about the concept of an organism cannot be considered as something a priori, as it is commonly done in consideration of bioethics, as we have shown in the second section.

In the last paragraphs we have shown that monism about the concept of an organism requires a lot of work to be defended and not everyone would be satisfied with the solution, even if this were possible. Of course, bioethicists can also accept pluralism about the concept of the organism and follow one of two paths. They might choose option (2) or (3) from these listed above. Before outlining these paths, a word is required here concerning the notion of pluralism itself. Pluralism about the concept of an organism is something that seems to be a weak position. There are multiple concepts of an organism, so we have to accept them. However, some scholars have tried to provide ontological justification for pluralism, developing it into a mature philosophical framework. There are many approaches, but two are quite popular and will be outlined below. Both rest on the idea that the concepts of an organism depend on the research aim, with the first explaining it in terms of pragmatism [21, 25, 81], the other in terms of process ontology [83, 84].

Let us start from the pragmatic justification for pluralism concerning the concept of an organism. The idea here is that the concept of an organism has a special role in biology. It is a tool that is supposed to help solve scientific problems. Mainly, if scientists have a problem to solve, a concept of an organism is developed to fit the needs of scientists that pursue the research tasks. The way they define organisms reflects the goal they want to achieve. In other words, the concept is goal-oriented and emerges within the context of given research; it is rooted within the theoretical and empirical basis of a given scientific discipline [21, 25, 81]. Mainly, different concepts of an organism exist because researchers in different fields of the biological sciences are interested in solving different problems, and thus they use different concepts of an organism to pursue those different desires. This was nicely summarized by Wolfe, who wrote that, "the organism is neither a *discovery* like the circulation of the blood or the glycogenic function of the liver, nor a particular biological *theory* like epigenesis or preformationism. It is rather a *concept* which plays a series of roles–sometimes overt, sometimes masked–throughout the history of biology, and frequently in very normative ways, also shifting between the biological and the social" [81].

The other approach depends on the ontology of the process [83, 84]. The idea here is that organisms are not some sort of "things" or philosophically speaking, "substances," but rather some sort of processes that undergo constant changes. They integrate elements of the environment, their parts (like organs) wear out and are replaced by other elements (cell with another cell) and some of them even undergo drastic

changes during development. Just think about the tremendous change human zygote or insect eggs undergo when they are transformed into adults. They constitute the flow of living matter, and we only capture instances of this flow, which is temporarily stable and creates an illusion that organisms are things. Can we speak about the organism at all in such a framework? To show that it is possible we can compare it to a river. It flows and seems to have no clear boundaries, but this does not mean there are no boundaries. There exist multiple ways to set boundaries to a river, ranging from the geographic to the geological and so on. Each one is correct but merely captures a different part of the process–the same is true for organisms. Organisms are processes, and we can put different types of boundaries on them by putting forward different concepts of an organism. Each of these concepts equal, just capturing a slightly different part of this flow of living matter. The sort of boundaries that should be chosen for research will vary from one scientific project to another [83, 85, pp. 69–143].

The above paragraphs show that the pluralistic approach to the concept of an organism can be defended by means of philosophical arguments. Therefore, bioethicists do not have to struggle to develop the universal concept of an organism or to defend monism–of course, if they want to do so, they can. However, it is unnecessary as they can accept pluralism about the concept of the organism with good justification. If they do so, they might choose either option (2) or (3). Given option (2) they might base their investigations on one of the concepts of the organism from biology. Going down this road bioethicists should acknowledge that the merger between all views of organisms, that is the notion of "functional integration," is nebulous and cannot provide a basis to unequivocally answer whether brain-dead patients are dead or not. Therefore, when they want to base their arguments on the concept of an organism, bioethicists should review the existing concepts and select the one that would be the most appropriate for their arguments. By doing so they assure that their argument will be based on a concept that is rooted in contemporary biological research, rather than conceptualizing their concept which might be less capable of staying in touch with current knowledge and even overlook some important biological aspects.

The option (2) would not be very satisfying for the majority of bioethicists. We think that opening doors for biological ideas can help bioethics by providing the flow of new ideas–and the same is true for biologists. However, it is very likely that a lot of concepts from biology and the philosophy of biology cannot simply be transferred and accepted as they have been developed to serve different research purposes that vary across biological sciences. For instance, people working on development want to understand how species develop from single cells, so they might define organisms along those lines. Evolutionary biologists, in turn, want to understand how populations undergo natural selection, so their concepts might be defined in such a way as to be relevant for this goal. This is far removed from the goals that bioethics pursue. We have presented these concepts to show that even if we stick to the concepts of the organism from biological theories–which should be empirically informed–they are still different about determining death.

The option (3) is an interesting one and seems to be the most appropriate. Going down this road, bioethicists should, just as with the previous option, realize that the merger of "functional integration," that is common for all understandings of an

organism, is nebulous and prone to different yet equally valid operationalizations within the life sciences. Scientists utilize such operationalizations for pragmatic reasons–mainly to accomplish their research goals. Bioethicists engaged in the brain death debate might argue that they also have such goals. In essence, they are theorizing about the biological status of a brain-dead patient. It seems that there are some concepts of an organism, including "bioethical-biological" ones, which implies that the brain-dead patient is alive (e.g., Shewmon's view), and there are some different perspectives (e.g., Moschella's master part view) which indicate the opposite–that the brain-dead patient is dead.

Since the strategy for the bioethicists that we are discussing takes organismal pluralism at face value, those who adopt it should not avoid the conclusion that brain-dead patients can be alive given some concepts of an organism, and dead according to equally valid views. This is just the same as the fact that biologists believing in organismal pluralism accept the conclusion that, for instance, protozoans are part of a termite organism (according to the immunological view) and simultaneously are not parts of a termite's organism (from the zygotic functional perspective). While the acceptance of biological pluralism is not a problem for biologists, with biologists both aware and even calling for it [20], it seems to be a much graver issue for bioethicists. This is because all bioethicists want to investigate the status of brain-dead patients for the sake of providing some objective moral guidance, or at least that which would be invariant within a given society. Most of them believe that it is objectively wrong to cause death. Therefore, they investigate the status of a brain-dead patient for the sake of reaching a conclusion of whether the patient is an entity of the sort that might die at some point [3; 18; 19; 86; 87, pp. 1–520; 88; 89, pp. 89–114]. Others, even though they do not believe that it is always wrong to cause death, uphold that it is objectively wrong not to disclose that a patient has the status of an organism when a decision about the patient's fate is to be made [38, pp. 1–174; 39; 78; 90].

However, if there is organismal pluralism, as we believe is the case, there is no definitive answer as to whether the brain-dead patient is alive or dead in a biological sense. Therefore, strictly speaking, the information disclosed should be the information about the patient's organismal status assessed from all biologically valid concepts of an organism. Moreover, bioethicists that believe that causing the biological death of an organism is wrong should accept that there is no definite answer as to whether conduct such as organ removal causes the death of the brain-dead patient. It would be wrong to remove organs from brain-dead patients given some concepts of an organism such as the immunological or Shewmon's view. It would be morally neutral given some other concepts, such as Moschella's master part view or the developmental functional view. This second upshot of organismal pluralism is especially problematic from the perspective of ethics, since ethicists and policymakers would like to have a single correct answer when considering whether someone has caused the death of some other human. After all, we cannot both blame and reward someone for a given act.

Another risk coming from the organismal pluralism and from the fact that we conceptualize organisms to pursue some goals is that bioethicists can have different goals in mind. Some of them work on the elaboration of the idea that brain-dead patients are alive and construct a number of arguments for this. As a result, they conceptualize the concept of an organism in a way that fits their philosophical agenda. Others

conceptualize organisms differently, leading to a situation in which their concepts imply that brain-dead patients are dead. We do not suggest that they might do so deliberately, but rather this might be done as a by-product of their way of reasoning. Bioethicists rarely conceptualize organisms first and then start thinking about the status of brain-dead patients. Rather, they first develop some sort of intuition (for instance, those driven by their religious views) and then try to develop the concept of the organism as a part of their argumentation. This poses the danger that their views will impinge on how they conceptualize organisms, making the concept biased towards the conclusion they want to develop. This is because the concept of an organism is very elastic and, as we have shown throughout this paper, can be defined in many ways depending on the goals.

The option (4), which might work for some bioethics, is to claim that death cannot be conflated with the cessation of an organism. Perhaps the whole talk about the 'bioethical-biological' *concept of an organism* presented in this article springs from some misunderstanding. One could stress that Bernat et al. identified death precisely speaking with "the permanent cessation *of functioning* of the organism *as a whole*" (emphasis added), while the most recent of Bernat's and Huang's ideas identify death with cessation of an organism *as a whole*. A similar strategy might also work for the President's Commission view related to the integrated functioning of the body, or the idea of Nair-Collins that we should identify organism death with the "irreversible cessation of the organismic capacity to maintain homeostasis of the extracellular fluid and thereby resist entropy." Along these lines, it might be stated that a definition of death cannot be deduced from the sine qua non conditions for the existence of an organism as they appear in theoretical biology. To prove that the death of an organism and its cessation might be distinct phenomena, one could look to paleontological research and argue that permanently non-functioning organisms are dead, yet still exist as organisms, or that the paleontologists' object of research does not constitute an organism *as a whole,* while at the same time it is still an existent organism. Such a strategy, however, is problematic for several reasons. First, as noted in the introductory section, many bioethicists explicitly or implicitly interpret the debate about the definition of death as precisely a debate about the sine qua non conditions for the existence of an organism. Arguably, for many authors, terms like "permanently disintegrated organism," an organism that permanently ceased to be a whole, or the "permanently non-functioning organism" strictly mean nothing more than "the former organism" or the entity that used to be an organism but now ceases to be such. Therefore, according to such an approach, a paleontologist would strictly conduct her research on the remains of an organism, not on an organism itself.

Second, although the concept of an organism is quite nebulous and prone to different interpretations, one thing about this notion is more straightforward, namely that organisms are entities that by definition are functionally integrated and constitute wholes that differ from the sum of their parts. This becomes clear if we look at the historical-philosophical analysis of the notion in question provided by Daniel Nicholson:

> The concept of organism is grounded in two key organizational relations: (a) the parts–whole relation, according to which an organism is construed as a structurally and functionally differentiated whole, and (b) the inside-outside relation,

> according to which an organism is construed as an autonomous system capable of maintaining itself in the face of changes to its environment…. The intricate relation between parts and whole was first recognized by Kant in his Critique of the Power of Judgment … in which he observed that living beings are self-organizing systems in the sense that their parts reciprocally produce one another in accordance to the organization of the whole. [52; cf. 91; 92, p. 245/374]

Therefore, again, if one speaks of a permanently functionally disintegrated organism or an organism that no longer constitutes a whole, it is best understood as the "former organism," the entity that used to be an organism but ceased to be such. So, if death is understood as the functional disintegration of an organism or as a moment when an organism ceases to be a whole, then death is nothing more than the cessation of an organism. Not only could Nicholson's analysis be recalled here but also the words of Thomas Pradeu, according to whom "asking what a biological individual is [and in particular asking what individual organism is–we would add] means asking what constitutes a countable, relatively well-delineated, and cohesive unit in the living world" [23].

Of course, one could speak about death also in a different biological sense than the cessation of an organism, referring to the opposition between dead and living material. This is one of the most debated issues in biology and the philosophy of biology–one that is especially interesting when scholars try to find out whether viruses are alive [93] or when they try to define life for astrobiological investigations [94]. For both cases, figuring out whether something is living or dead is undoubtedly an important issue to drive research. As Koonin and Starokadomskyy put it, "the 'dead-alive' dichotomy in the classification of biological entities seems to present unsolvable conandra whereby the borders of life cannot be clearly defined" [93]. However, in such a sense, it is beyond all controversy that brain-dead patients on artificial support are living since there are living organs, tissues, and cells in their bodies. They belong to the "life" domain of the world. Therefore, in our opinion, the whole discussion about the status of brain-dead patients within the "biological-bioethical" paradigm makes sense if it is a controversy about the existence of some cohesive biological individual such as an organism.

The last problem with the strategy is associated with the ordinary meaning of the phrase "to die." In everyday talk, it means the cessation of the existence of some living entity. When a person dies, "the person goes out of existence; subsequently, there is no such thing as that person" [95]; (cf. [96; 97, p. 287]), even though the remains of the person might still be present, and even though we might refer to, for example, Socrates by the phrase "dead person," meaning a person that used to exist at some point. We dare to claim that when we talk about a dead organism, we mean the entity that has gone out of existence as an organism.

Conclusions

This investigation might lead us to the conclusion that we should accept that there is no single correct answer as to whether a type of human action causes death or not. Perhaps we should assume that there are no definitive answers about moral blame or

the prize of performing organ transplantation. However, such a conclusion would be premature. Note that people believed that causing the death of their peers, and especially killing them, was wrong long before the notion of an organism appeared in modern science. They believed this in ancient times even though, strictly speaking, there was no theory of an organism in ancient philosophy, the Bible, the Koran, or the sutras [99]. Perhaps by causing death or killing their peers, people always had in mind something different from causing the cessation of an organism, something of non-reducible moral phenomenology [cf. 8–10, 99].

Acknowledgements This research article was funded by National Science Centre (Poland) under Sonata 16 research Grant (No. 2020/39/D/HS1/02907) awarded to Piotr Grzegorz Nowak, and under Opus 14 research Grant (No. 2017/27/B/HS1/00290) awarded to Adrian Stencel. We want to thank two anonymous reviewers and the associate editor of *Theoretical Medicine and Bioethics* for their helpful comments on the earlier version of this article. We would also like to thank Julianna Kober for her assistance during the copyediting of this article.

Author contributions Both authors contributed equally to the present study.

Funding The research was funded by National Science Centre, Poland Sonata research grant (Grant No. 2020/39/D/HS1/02907) to Piotr Grzegorz Nowak, and by National Science Centre, Poland Opus research grant (Grant No. 2017/27/B/HS1/00290) to Adrian Stencel. This publication has been funded from the Faculty of Philosophy budget under the Strategic Programme Excellence Initiative at Jagiellonian University.

Data availability Not applicable.

Code availability Not applicable.

Declarations

Conflict of interest Both authors have no conflict of interest to disclose.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

1. Bernat, James L., Charles M. Culver, and Bernard Gert. 1981. On the definition and criterion of death. *Annals of Internal Medicine* 94 (3): 389–394.
2. President's Commission for the Study of Ethical Problems in Medicine. 1981. *Defining death: A report on the medical, legal and ethical issues in the determination of death*. Washington, DC: United States Government.

3. Moschella, Melissa. 2016. Integrated but not whole? Applying an ontological account of human organismal unity to the brain death debate. *Bioethics* 30 (8): 550–556. https://doi.org/10.1111/bioe.12258.
4. Shewmon, D. Alan. 2001. The brain and somatic integration: Insights into the standard biological rationale for equating "brain death" with death. *Journal of Medicine and Philosophy* 26 (5): 457–478. https://doi.org/10.1076/jmep.26.5.457.3000.
5. Bernat, James L. 2019. Refinements in the organism as a whole rationale for brain death. *The Linacre Quarterly* 86 (4): 347–358. https://doi.org/10.1177/0024363919869795.
6. Omelianchuk, Adam. 2021. Brain death as the end of a human organism as a self-moving whole. *Journal of Medicine and Philosophy* 46 (5): 530–560. https://doi.org/10.1093/jmp/jhab021.
7. Condic, Maureen L. 2016. Determination of death: A scientific perspective on biological integration. *Journal of Medicine and Philosophy* 41 (3): 257–278. https://doi.org/10.1093/jmp/jhw004.
8. Veatch, Robert M. 2015. Killing by organ procurement: brain-based death and legal fictions. *Journal of Medicine and Philosophy* 40 (3): 289–311. https://doi.org/10.1093/jmp/jhv007.
9. Nowak, Piotr Grzegorz. Forthcoming. Death as the cessation of an organism and the moral status alternative. *Journal of Medicine and Philosophy*.
10. Green, Michael B., and Daniel Wikler. 1980. Brain death and personal identity. *Philosophy & Public Affairs* 9 (2): 105–133.
11. Lizza, John P. 2004. The conceptual basis for brain death revisited. In *Brain Death and Disorders of Consciousness. Advances in Experimental Medicine and Biology*, ed. Calixto Machado and D. Alan Shewmon, 51–59. Boston, MA: Springer US.
12. The President's Council on Bioethics. 2008. *Controversies in the determination of death: A white paper of the President's Council on Bioethics*. Washington, DC: President's Council on Bioethics.
13. Singer, Peter. 1996. *Rethinking life and death: The collapse of our traditional ethics*. New York: St. Martin's Press.
14. Grisez, Germain, and Boyle Joseph M. Jr. 1979. *Life and death with liberty and justice*. Notre Dame: University of Notre Dame Press.
15. Lamb, David. 1985. *Death, brain death and ethics*. London: Routledge.
16. Marquis, Don. 2018. Death is a biological phenomenon. *Diametros* 55: 20–26.
17. Singer, Peter. 2018. The challenge of brain death for the sanctity of life ethic. *Ethics & Bioethics (in Central Europe)* 8 (3–4): 153–165.
18. Moschella, Melissa. 2016. Deconstructing the brain disconnection–brain death analogy and clarifying the rationale for the neurological criterion of death. *Journal of Medicine and Philosophy* 41 (3): 279–299. https://doi.org/10.1093/jmp/jhw006.
19. Melissa, Moschella. 2019. The human organism is not a conductorless orchestra: A defense of brain death as true biological death. *Theoretical Medicine and Bioethics* 40 (5): 437–453.
20. Pepper, John W., and Matthew D. Herron. 2008. Does biology need an organism concept? *Biological Reviews* 83 (4): 621–627. https://doi.org/10.1111/j.1469-185X.2008.00057.x.
21. Kovaka, Karen. 2015. Biological individuality and scientific practice. *Philosophy of Science* 82 (5): 1092–1103. https://doi.org/10.1086/683443.
22. Gilbert, Scott F., Jan Sapp, and Alfred I. Tauber. 2012. A symbiotic view of life: We have never been individuals. *The Quarterly Review of Biology* 87 (4): 325–341. https://doi.org/10.1086/668166.
23. Pradeu, Thomas. 2016. Organisms or biological individuals? Combining physiological and evolutionary individuality. *Biology & Philosophy* 31 (6): 797–817. https://doi.org/10.1007/s10539-016-9551-1.
24. Şencan, Sinan. 2019. A tale of two individuality accounts and integrative pluralism. *Philosophy of Science* 86 (5): 1111–1122. https://doi.org/10.1086/705523.
25. Stencel, Adrian, and Agnieszka M. Proszewska. 2018. How research on microbiomes is changing biology: A discussion on the concept of the organism. *Foundations of Science* 23 (4): 603–620. https://doi.org/10.1007/s10699-017-9543-x.
26. Suárez, Javier. 2020. The stability of traits conception of the hologenome: An evolutionary account of holobiont individuality. *History and Philosophy of the Life Sciences* 42 (1): 11. https://doi.org/10.1007/s40656-020-00305-2.
27. Suárez, Javier, and Adrian Stencel. 2020. A part-dependent account of biological individuality: Why holobionts are individuals and ecosystems simultaneously. *Biological Reviews* 95 (5): 1308–1324. https://doi.org/10.1111/brv.12610.
28. Clarke, Ellen. 2010. The problem of biological individuality. *Biological Theory* 5 (4): 312–325. https://doi.org/10.1162/BIOT_a_00068.

29. Shewmon, D. Alan. 2010. Constructing the death elephant: A synthetic paradigm shift for the definition, criteria, and tests for death. *Journal of Medicine and Philosophy* 35 (3): 256–298. https://doi.org/10.1093/jmp/jhq022.
30. Huang, Andrew P., and James L. Bernat. 2019. The organism as a whole in an analysis of death. *Journal of Medicine and Philosophy* 44 (6): 712–731. https://doi.org/10.1093/jmp/jhz025.
31. Alexander, Joshua, Ronald Mallon, and Jonathan M. Weinberg. 2010. Accentuate the negative. *Review of Philosophy and Psychology* 1 (2): 297–314.
32. Pust, Joel. 2019. Intuition. In *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta. https://plato.stanford.edu/entries/intuition/. Accessed 16 January 2022.
33. Nair-Collins, Michael. 2015. Taking science seriously in the debate on death and organ transplantation. *Hastings Center* 45 (6): 38–48. https://doi.org/10.1002/hast.459.
34. Nair-Collins, Michael. 2018. A biological theory of death: Characterization, justification, and implications. *Diametros* 55: 27–43.
35. Trujillo, Cleber A., Richard Gao, Priscilla D. Negraes, Gu. Jing, Justin Buchanan, Sebastian Preissl, Allen Wang, et al. 2019. Complex oscillatory waves emerging from cortical organoids model early human brain network development. *Cell Stem Cell* 25 (4): 558-569.e7. https://doi.org/10.1016/j.stem.2019.08.002.
36. Moschella, Melissa. 2017. Complexity of defining death: Organismal death does not mean the cessation of all biological life. *Journal of Medical Ethics* 43 (11): 754–755. https://doi.org/10.1136/medethics-2017-104363.
37. Rosenkrantz, Gary S. 2012. Animate beings: Their nature and identity. *Ratio* 25 (4): 442–462.
38. Miller, Franklin G., and Robert D. Truog. 2012. *Death, dying, and organ transplantation: Reconstructing medical ethics at the end of life*. New York: Oxford University Press.
39. Shewmon, D. Alan. 2009. Brain death: Can it be resuscitated? *Hastings Center.* 39 (2): 18–24.
40. Korein, Julius. 1978. The problem of brain death: Development and history. *Annals of the New York Academy of Sciences* 315 (1): 19–38.
41. Shewmon, D. Alan. 1998. Chronic "brain death": Meta-analysis and conceptual consequences. *Neurology* 51 (6): 1538–1545. https://doi.org/10.1212/Wnl.51.6.1538.
42. Lambert, Frank L. 2002. Disorder–A cracked crutch for supporting entropy discussions. *Journal of Chemical Education* 79 (2): 187–192. https://doi.org/10.1021/ed079p187.
43. Lizza, John P. 2018. In defense of brain death: Replies to Don Marquis, Michael Nair-Collins, Doyen Nguyen, and Laura Specker Sullivan. *Diametros* 55: 68–90.
44. Schrödinger, Erwin. 2012. *What is life? With mind and matter and autobiographical sketches*. Cambridge: Cambridge University Press.
45. Doolittle, W. Ford. 2019. Making evolutionary sense of gaia. *Trends in Ecology & Evolution* 34 (10): 889–894. https://doi.org/10.1016/j.tree.2019.05.001.
46. Suárez, Javier. 2018. The importance of symbiosis in philosophy of biology: An analysis of the current debate on biological individuality and its historical roots. *Symbiosis* 76 (2): 77–96. https://doi.org/10.1007/s13199-018-0556-1.
47. Cleland, Carol E., and Christopher F. Chyba. 2002. Defining 'life.' *Origins of life and evolution of the biosphere* 32 (4): 387–293. https://doi.org/10.1023/A:1020503324273.
48. Bich, Leonardo, and Sara Green. 2018. Is defining life pointless? Operational definitions at the frontiers of biology. *Synthese* 195 (9): 3919–3946. https://doi.org/10.1007/s11229-017-1397-9.
49. Trifonov, Eduard Nikolajevič. 2012. Definition of life: Navigation through uncertainties. *Journal of Biomolecular Structure and Dynamics* 29 (4): 647–650. https://doi.org/10.1080/073911012010525017.
50. Heikki, Helanterä. 2016. An organismal perspective on the evolution of insect societies. *Frontiers in Ecology and Evolution* 4: 6. https://doi.org/10.3389/fevo.2016.00006.
51. Canciani, Mark, Argyris Arnellos, and Alvaro Moreno. 2019. Revising the superorganism: An organizational approach to complex eusociality. *Frontiers in Psychology* 10: 2653. https://doi.org/10.3389/fpsyg.2019.02653.
52. Nicholson, Daniel J. 2014. The return of the organism as a fundamental explanatory concept in biology. *Philosophy Compass* 9 (5): 347–359. https://doi.org/10.1111/phc3.12128.
53. Clarke, Ellen. 2013. The multiple realizability of biological individuals. *The Journal of Philosophy* 110 (8): 413–435.
54. Chiu, Lynn, and Gérard. Eberl. 2016. Microorganisms as scaffolds of host individuality: An eco-immunity account of the holobiont. *Biology & Philosophy* 31 (6): 819–837. https://doi.org/10.1007/s10539-016-9552-0.

55. Ronai, Isobel, Gregor P. Greslehner, Federico Boem, Judith Carlisle, Adrian Stencel, Javier Suárez, Saliha Bayir, et al. 2020. "Microbiota, symbiosis and individuality summer school" meeting report. *Microbiome* 8 (1): 117. https://doi.org/10.1186/s40168-020-00898-7.
56. Huxley, Thomas Henry. 1852. Upon animal individuality. *Proceedings of the Royal Institute of Great Britain* 11: 184–189.
57. Moore, Keith L., T.V.N. Persaud, and Mark G. Torchia. 2016. *Before we are born: Essentials of embryology and birth defects*, 9th ed. Philadelphia: Saunders Company.
58. Janzen, Daniel H. 1977. What are dandelions and aphids? *The American Naturalist* 111 (979): 586–589. https://doi.org/10.1086/283186.
59. Lizza, John P. 2018. Defining death: beyond biology. *Diametros* 55: 1–19.
60. Dupré, John E. and Maureen A. O'Malley. 2009. Varieties of living things: Life at the intersection of lineage and metabolism. *Philosophy, Theory, and Practice in Biology* 1(3). https://doi.org/10.3998/ptb.6959004.0001.003.
61. Pradeu, Thomas. 2019. Philosophy of biology: Immunology and individuality. *eLife* 8: e47384. https://doi.org/10.7554/eLife.47384.
62. Tauber, Alfred I. 1994. *The immune self: Theory or metaphor?* Cambridge: Cambridge University Press.
63. Germain, Ronald N. 2012. Maintaining system homeostasis: The third law of Newtonian immunology. *Nature Immunology* 13 (10): 902–906. https://doi.org/10.1038/ni.2404.
64. Wynn, Thomas A., Ajay Chawla, and Jeffrey W. Pollard. 2013. Macrophage biology in development, homeostasis and disease. *Nature* 496 (7446): 445–455. https://doi.org/10.1038/nature12034.
65. Chu, Hiutung, and Sarkis K. Mazmanian. 2013. Innate immune recognition of the microbiota promotes host-microbial symbiosis. *Nature Immunology* 14 (7): 668–675. https://doi.org/10.1038/ni.2635.
66. Allen, Breanna M., Kamir J. Hiam, Cassandra E. Burnett, Anthony Venida, Rachel DeBarge, Iliana Tenvooren, Diana M. Marquez, Nam Woo Cho, Yaron Carmi, and Matthew H. Spitzer. 2020. Systemic dysfunction and plasticity of the immune macroenvironment in cancer models. *Nature Medicine* 26 (7): 1125–1134. https://doi.org/10.1038/s41591-020-0892-6.
67. Wood, Kathryn J., and Ryoichi Goto. 2012. Mechanisms of rejection: Current perspectives. *Transplantation* 93 (1): 1–10. https://doi.org/10.1097/TP.0b013e31823cab44.
68. Lewontin, Richard C. 1970. The units of selection. *Annual Review of Ecology and Systematics* 1: 1–18. https://doi.org/10.1146/annurev.es.01.110170.000245.
69. Godfrey-Smith, Peter. 2013. Darwinian individuals. In *From groups to individuals: Evolution and emerging individuality*, ed. Frédéric. Bouchard and Philippe Huneman, 17–36. Cambridge (MA): MIT Press.
70. Griesemer, James. 2001. The units of evolutionary transition. *Selection* 1 (1–3): 67–80. https://doi.org/10.1556/select.1.2000.1-3.7.
71. Godfrey-Smith, Peter. *Darwinian populations and natural selection*. Oxford: Oxford University Press.
72. Stencel, Adrian, and Dominika M. Wloch-Salamon. 2018. Some theoretical insights into the hologenome theory of evolution and the role of microbes in speciation. *Theory in Biosciences* 137 (2): 197–206. https://doi.org/10.1007/s12064-018-0268-3.
73. Sterelny, Kim, Kelly C. Smith, and Michael Dickison. 1996. The extended replicator. *Biology and Philosophy* 11 (3): 377–403. https://doi.org/10.1007/BF00128788.
74. Hull, David L. 1980. Individuality and selection. *Annual Review of Ecology and Systematics* 11: 311–332. https://doi.org/10.1146/annurev.es.11.110180.001523.
75. Doolittle, W. Ford., and Austin Booth. 2017. It's the song, not the singer: An exploration of holobiosis and evolutionary theory. *Biology & Philosophy* 32: 5–24. https://doi.org/10.1007/s10539-016-9542-2.
76. Okasha, Samir. 2018. *Agents and goals in evolution*. Oxford: Oxford University Press.
77. Palazzo, Alexander F., and T. Ryan Gregory. 2014. The case for junk DNA. *PLOS Genetics* 10 (5): e1004351. https://doi.org/10.1371/journal.pgen.1004351.
78. Truog, Robert D., and Walter M. Robinson. 2003. Role of brain death and the dead-donor rule in the ethics of organ transplantation. *Critical Care Medicine* 31 (9): 2391–2396. https://doi.org/10.1097/01.Ccm.0000090869.19410.3c.
79. Chodasewicz, Krzysztof. 2014. Evolution, reproduction and definition of life. *Theory in Biosciences* 133 (1): 39–45. https://doi.org/10.1007/s12064-013-0184-5.

80. Burt, Austin, and Robert Trivers. 2009. *Genes in conflict: The biology of selfish genetic elements*. Cambridge (MA): Harvard University Press.
81. Wolfe, Charles T. 2014. The organism as ontological go-between: Hybridity, boundaries and degrees of reality in its conceptual history. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* 48: 151–161. https://doi.org/10.1016/j.shpsc.2014.06.006.
82. Sober, Elliott. 2015. *Ockham's razors*. Cambridge: Cambridge University Press.
83. Dupré, John. 2020. Processes within processes: A dynamic account of living beings and its implications for understanding the human individual. In *Biological Identity*, ed. Anne Sophie Meincke and John Dupré, 149–166. London: Routledge.
84. Meincke, Anne Sophie. 2021. One or two? A Process View of pregnancy. *Philosophical Studies*. https://doi.org/10.1007/s11098-021-01716-y.
85. Dupré, John. 2012. *Processes of life: Essays in the philosophy of biology*. Oxford: Oxford University Press.
86. Nguyen, Doyen. 2018. A holistic understanding of death: Ontological and medical considerations. *Diametros* 55: 44–62. https://doi.org/10.13153/diam.1175.
87. Nguyen, Doyen. 2018. *The new definitions of death for organ donation: A multidisciplinary analysis from the perspective of Christian ethics*. Bern: Peter Lang.
88. Shewmon, D. Alan. 1985. The metaphysics of brain death, persistent vegetative state, and dementia. *The Thomist: A Speculative Quarterly Review* 49 (1): 24–80.
89. Shewmon, D. Alan, and Elisabeth Seitz Shewmon. 2004. The semiotics of death and its medical implications. In *Brain Death and Disorders of Consciousness. Advances in Experimental Medicine and Biology*, ed. Calixto Machado and D. Alan Shewmon, 89–114. Boston, MA: Springer US.
90. Nair-Collins, Michael. 2018. The public's right to accurate and transparent information about brain death and organ transplantation. *Hastings Center* 48 (S4): S43–S45. https://doi.org/10.1002/hast.953.
91. Ruiz-Mirazo, Kepa, Arantza Etxeberria, Alvaro Moreno, and Jesús Ibáñez. 2000. Organisms and their place in biology. *Theory in Biosciences* 119 (3): 209–233. https://doi.org/10.1007/s12064-000-0017-1.
92. Kant, Immanuel. 2000. *Critique of the Power of Judgment*. Cambridge: Cambridge University Press.
93. Koonin, Eugene, and Petro Starokadomskyy. 2016. Are viruses alive? The replicator paradigm sheds decisive light on an old but misguided question. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* 59: 125–134. https://doi.org/10.1016/j.shpsc.2016.02.016.
94. Ruiz-Mirazo, Kepa, Juli Peretó, and Alvaro Moreno. 2004. A universal definition of life: Autonomy and open-ended evolution. *Origins of life and evolution of the biosphere* 34 (3): 323–346. https://doi.org/10.1023/B:ORIG.0000016440.53346.dc.
95. Feldman, Fred. 1994. *Confrontations with the reaper: A philosophical study of the nature and value of death*. New York: Oxford University Press.
96. Luper, Steven. 2021. Death. In *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta. https://plato.stanford.edu/entries/death/.Accessed 17 January 2022.
97. Black, Henry Campbell. 1968. *Black's law dictionary, definitions of terms and phrases of american and english jurisprudence, ancient and modern, revised fourth edition*. 253: 287. St Paul (MN): West Publishing Company.
98. Cheung, Tobias. 2006. From the organism of a body to the body of an organism: Occurrence and meaning of the word 'organism' from the seventeenth to the nineteenth centuries. *The British Journal for the History of Science* 39 (3): 319–339. https://doi.org/10.1017/S0007087406007953.
99. Nowak, Piotr Grzegorz. 2018. Brain death as irreversible loss of a human's moral status. *Ethics & Bioethics (in Central Europe)* 8 (3–4): 167–178.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.