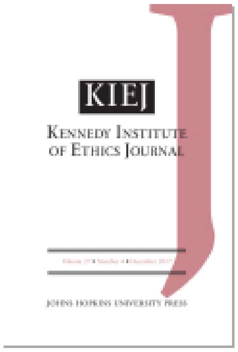
Can the Brain-Dead Be Harmed or Wronged?: On the Moral Status of Brain Death and its Implications for Organ Transplantation



Michael Nair-Collins

Kennedy Institute of Ethics Journal, Volume 27, Number 4, December 2017, pp. 525-559 (Article)

Published by Johns Hopkins University Press
DOI: *https://doi.org/10.1353/ken.2017.0041*

Michael Nair-Collins

Can the Brain-Dead Be Harmed or Wronged? On the Moral Status of Brain Death and its Implications for Organ Transplantation

ABSTRACT. The majority of transplantable human organs are retrieved from patients declared dead by neurological criteria, or "brain-dead." Since brain death is considered to be sufficient for death, the procurement of vital organs is not considered to harm or wrong such patients. In this essay I argue that this is not the case. After distinguishing welfare, experiential, and investment interests, and defining precedent autonomy and surviving interests, I argue that brain-dead patients can be, and many are, harmed and wronged by organ procurement as currently practiced. Indeed, with respect to precedent autonomy and surviving investment interests, the brain-dead are morally equivalent to patients with severe dementia, and thus can be harmed and wronged if and only if, and to the extent that, patients with severe dementia can. The "bright line" that separates brain death from all other conditions for clinical and legal purposes is not justified by any morally relevant distinctions.

INTRODUCTION

The dead donor rule, which requires that organ donors not be killed by the process of organ procurement, is thought to protect vulnerable patients from exploitation and from being harmed through organ procurement (Robertson 1999; President's Council 2008). In current practice, the majority of transplantable organs are retrieved from patients who are declared dead by neurological criteria, or "brain-dead." Because brain death is considered to be sufficient for death (President's Commission 1981; President's Council 2008), it is thought that brain-dead donors are neither harmed nor wronged by organ removal.

In this essay I argue that this is not the case. Brain-dead donors can be, and many are, harmed and wronged by organ procurement as currently

Kennedy Institute of Ethics Journal Vol. 27, No. 4, 525–559 © 2017 by Johns Hopkins University Press

practiced. Indeed, with respect to precedent autonomy and surviving investment interests, the brain-dead are morally equivalent to patients with severe dementia, and thus can be harmed and wronged if and only if, and to the extent that, patients with severe dementia can. The "bright line" that separates brain death from all other conditions for clinical and legal purposes (Gostin 2014; Magnus et al. 2014) is not justified by any morally relevant distinctions.

BRAIN DEATH AND ORGAN TRANSPLANTATION

Brain death is a state characterized by unresponsiveness to pain or other stimuli (with the exception of spinally mediated reflexes), brainstem areflexia, and apnea, in the presence of a known cause of coma and the absence of confounds such as hypothermia or intoxication. The majority of transplanted organs are retrieved from patients who have been declared dead by neurological criteria, and organs are removed while the patient remains on the ventilator and with a spontaneously beating heart. However, because the brain-dead donor is considered to be dead at the time of removal, the procurement of vital organs is not considered to cause death.

The prevailing view that takes brain death to be sufficient for death in a biological sense (Bernat 2006; President's Council 2008; Shemie et al. 2014), however, is scientifically inaccurate. Biologically, death is the irreversible cessation of the integrative functioning of the organism as a whole in its capacity to maintain homeostasis and thereby resist entropy with its external environment (Bernat et al. 1981; Shewmon 2001; Miller and Truog 2012; Nair-Collins 2015). From this perspective, the evidence refuting the prevailing view is conclusive, overwhelming, and obvious. If supported with common medical treatments such as mechanical ventilation, patients who meet diagnostic standards for brain death can engage in a wealth of integrative, homeostasis-maintaining functions, including wound healing, nutrition, cellular respiration, gas exchange, febrile responses to infection, tachycardic and hypertensive responses to incision, neuroendocrine control of electrolyte concentrations, growth and sexual maturation in children, and fetus gestation in pregnant women (Truog and Fackler 1992; Halevy and Brody 1993; Shewmon 2001; Miller and Truog 2012).

Furthermore, some can be discharged from an acute care hospital and cared for at home with mechanical ventilation and only basic nursing care, and can be maintained in this state for many years (Shewmon 1998; 1999;

Miller and Truog 2012). Though unconscious, such patients may exhibit *more* integrated physiologic stability and maintenance of homeostasis than some other dying but still-living patients found in our ICUs (Shewmon 2010). This evidence similarly demonstrates that brain-dead patients can engage in the same "vital work of an organism" (President's Council 2008) as other still-living individuals (Miller and Truog 2009; Joffe 2010; Shah et al. 2011; Nair-Collins 2013). When approached from a purely biological perspective, the state known as "brain death" is best understood as a type of irreversible apneic coma (Zamparetti et al. 2004), and such patients are biologically alive though (almost certainly) irreversibly unconscious.[1]

Given this empirical evidence, it follows that brain-dead patients remain biologically alive at the time of organ procurement, and the removal of vital organs is the proximate cause of death. This information is surely relevant to making informed and intelligent choices about end-of-life care and whether to donate organs, but it is not disclosed when donors or their surrogates agree to organ donation. Currently in the US, one may authorize organ donation by checking a box on one's driver's license or verbally stating to the clerk at the Department of Motor Vehicles that one wishes to donate organs "after death." Alternatively, one can enroll as an organ donor on Internet enrollment forms hosted by the local organ procurement organization. Finally, a family member may give surrogate authorization on a patient's behalf, after brain death has been established. Yet, the information that the brain-dead individual remains biologically alive at the time of organ procurement and that vital organ removal causes the donor's biological death is not routinely disclosed in any of these contexts. It follows that lethal heart-beating organ procurement often occurs in the absence of valid consent, whether in the form of prospective first-person informed consent, or informed surrogate consent (Woien et al. 2006; Rady et al. 2012; Nair-Collins 2013; Iltis 2015).

Since vital organ procurement from brain-dead donors causes death, yet most organ donors or their surrogates do not validly consent to lethal organ removal, this might seem *prima faci*e morally concerning. But on what grounds? Once the patient satisfies diagnostic standards for brain death, she is irreversibly unconscious, and so it may seem that she has no further interests of any kind, thus, she cannot be harmed. Furthermore, recent survey data suggest that most people in the US who would agree to donate organs after death would also agree to lethal organ procurement in irreversible coma (Nair-Collins et al. 2015). Thus, one might assume that if the donor is irreversibly unconscious, and if most people would probably agree to donate anyway, then there is no harm in lethal organ procurement from brain-dead donors. However, the relationship between sentience and harm is not so straightforward that the lack of sentience implies the lack of possibility of harm, and furthermore, harms are distinguishable from wrongs. More careful analysis is needed.

HARMS, WRONGS, AND A TAXONOMY OF INTERESTS

Interests, harms, and wrongs can take many forms, and they bear complex relationships to each other. Furthermore, there are a variety of terms of art used to describe these concepts in the philosophical literature. In what follows I review a few prominent conceptions of each, propose a taxonomy of interests, and explain how harms and wrongs are related to interests.

Feinberg conceives of a harm as "the thwarting, setting back, or defeating of an interest" (1986, 33), where an interest is anything that a person "has a stake in" (1986, 33). Interests are distinguishable components of a person's well-being, and the test of whether an interest was set back is whether, had the invasion not occurred, the condition of the interest or the well-being of the person would have been better off (1986, 34). Regan (1983/2004, 94–99) recognizes two kinds of harms: inflictions and deprivations. For Regan, the concept of harm as infliction revolves centrally around suffering, both physical and psychological, and suffering is understood in terms of prolonged pain of sufficient intensity and duration. By contrast, Regan motivates the concept of harm as deprivation by the idea that not all harms hurt or cause suffering, such as when one is denied the opportunity for an education or a career, thus diminishing one's life possibilities and opportunities, even though this loss of opportunity may not directly cause suffering, and indeed, even if the subject of the deprivation-harm does not conceive of the loss as a harm.

Most writers on this topic distinguish harms from wrongs. Harms can be caused by natural disasters, accidents, and bad luck, but these are not wrongs. Only a moral agent, a person responsible for her behavior, can wrong another. Thus a wrong involves morally culpable harm caused by a moral agent, the violation of a right, the failure to uphold an obligation, or unjust or unfair treatment. Most wrongs are also harms—an unprovoked physical attack constitutes both a wrong and a harm to the person who is attacked—but (depending on one's theory of harms), there can also be examples of wrongs that do not, on balance, harm the one who has been wronged. For example, a minor trespass of another's property, without damaging anything and without the property-owner becoming aware of the trespass, is a plausible example of a wrong that is not, on balance, a harm.

Interests may come in many forms. Regan distinguishes those things that are in our interest from those things that interest us (1983/2004, 87). The former interests are *welfare-interests*, and the latter are *preference-interests*. To say that some X is in A's welfare-interest is to say that having or doing X would benefit A, or would make a positive contribution to A's well-being (88). This includes such things as adequate nourishment, water, shelter, and rest. Furthermore, A need not have any particular preference for these things, nor conceive of them as benefits, in order for them to be in A's (welfare-) interest. Preference-interests, by contrast, are simply "those things that an individual is *interested in*, those things he likes, desires, wants or, in a word, prefers having" (Regan 1983/2004, 87).

Feinberg distinguishes welfare interests from what he calls *ulterior interests* (1986, 37). Ulterior interests constitute "a person's more ultimate goals and aspirations . . . [such as] producing good novels or works of art, solving a crucial scientific problem, . . . successfully raising a family, [. . .]" (1986, 37). To count as ulterior interests, however, these aims require not only strong desires for their achievement, but also long-term stability of the desire as well personal investment in them, that is, time and effort actually spent towards their achievement. Welfare interests, Feinberg explains, are in some sense less lofty, but in another sense are more important, because welfare interests constitute the means by which ulterior interests can be achieved. Although not sufficient, they are necessary for the pursuit and advancement of our ulterior interests. In this category Feinberg includes things such as physical and emotional health, financial stability, and the absence of excessive pain and suffering (1986, 37). Like Regan, Feinberg argues (1986, 42) that a person need not actually desire such things in order for them to nonetheless be in her interests. In a similar fashion, Davis (2007) draws on Regan's work and distinguishes welfare interests from what he calls *investment interests*, where investment interests include those things that a person is "invested in," which include one's religious commitments or understanding of personal dignity.

Finally, Dworkin (1994) offers a different kind of distinction, contrasting what he calls *critical interests* with *experiential interests*. "We all do things because we like the experience of doing them," Dworkin writes (1994, 201), and the value of these experiences is derived from the pleasure we get in having them. These experiential interests are contrasted with critical interests, which are the things that make a life go well or ill, other than the experiences that we have. They are the more important things that give a life meaning; they are the interests that people want to satisfy so that they do not feel that they are wasting their lives. As Dworkin writes, we have "critical interests in the character and value of our lives as a whole. These critical interests are connected . . . to our convictions about the intrinsic value—the sanctity or inviolability—of our own lives" (1994, 235). This emphasis on the character and value of a life taken as a whole, and Dworkin's emphasis on the importance of living (and ending) one's life "in character" (1994, 213), are usefully understood by way of their connection to MacIntyre's (1984) conception of life as a specific kind of a narrative (viz., a quest), which has elements of a story such as a plotline, characters, and importantly, a final chapter: The conclusion of the narrative that influences the quality and value of the entire narrative taken as a whole.

The above views capture overlapping but distinct and equally important ideas. The following taxonomy captures the general tenor (but perhaps not the specifics) of all of them. I consider harms to be setbacks to interests. These setbacks can involve both inflictions and deprivations, depending on the interest that is set back. Benefits are those things that advance our interests. Wrongs are morally culpable harms caused by a moral agent, violations of rights, failures to discharge duties, or unjust or unfair treatment. Hence, not all harms are wrongs, and I leave it open as to whether some wrongs are not harmful (i.e., do not on balance set back interests).

I propose three general and distinct categories of interests: welfare, experiential, and investment interests. Welfare interests constitute the necessary but not sufficient conditions of living well; they are the basic needs that are necessary for the advancement of our other interests. These include adequate nourishment and hydration, shelter and security, rest, financial stability, and so on. These things are in our interests whether we recognize them as such or not, and hence they do not require any preferences or affirmation in order to be in our interests.

Experiential interests, on the other hand, require sentience. They include pleasurable experiences of all sorts as well as the avoidance of unpleasant or painful experiences. However, unlike welfare interests, experiential interests can be roughly ordered along a continuum of complexity, from low-level, purely sensory experiences of bodily pain or pleasure without any accompanying conceptualization of the experience, to the profound or transcendent joy associated with certain aesthetic experiences, or the deep emotional–intellectual gratification involved in coming to some new understanding of a difficult intellectual question or problem. This continuum can be seen as analogous to Mill's famous distinction between lower and higher pleasures (1861). Thus we can see that the infliction of intense physical or psychological pain—Regan's infliction-harm—is understood as harm in this taxonomy in terms of setting back experiential interests.

Finally, investment interests constitute those things in which we have some preference or investment. Like experiential interests, investment interests can be roughly ordered along a continuum of importance to the person whose interests they are, from relatively trivial preferences (e.g., what to eat for dinner tonight), all the way up to our ultimate goals and aspirations, our deepest philosophical or religious commitments to the nature of a good life and a good death, the long-term projects that give our lives flavor and meaning, and the construction of our lives' narratives. Here, at the farthest end of the continuum of investment interests, we find Dworkin's critical interests and Feinberg's ulterior interests.

It is not uncommon to suppose that the capacity for sentience is a necessary condition for having interests (of any kind), whereas on the proposed taxonomy, the capacity for sentience is only necessary for experiential interests. However, the intuition that sentience is necessary for having interests may be explained by the fact that the advancement of some particular interest (such as a preference or goal) is nearly always accompanied by an awareness *that* the primary interest was satisfied, and this in turn usually constitutes the advancement of an experiential interest. For example, achieving an important investment interest, such as attaining an advanced academic degree, might also be accompanied by feelings of joy, relief, satisfaction, pride, and so on, and these feelings themselves constitute an advancement of experiential interests. Thus, in the normal case, probably the vast majority of our interest advancements (i.e., benefits) and the vast majority of our interest frustrations or deprivations (i.e., harms) include or are accompanied by experiences and hence, the advancement or frustration of experiential interests. This may account for the common belief that harms and benefits entail sentience; however, as Nozick's experience machine thought experiment suggests (1974), it is not merely the *experience of* achieving our goals and creating our lives that matters—it's the goal-achievement and life-creation themselves. See Figure 1 for a graphical overview of this proposed taxonomy of interests.

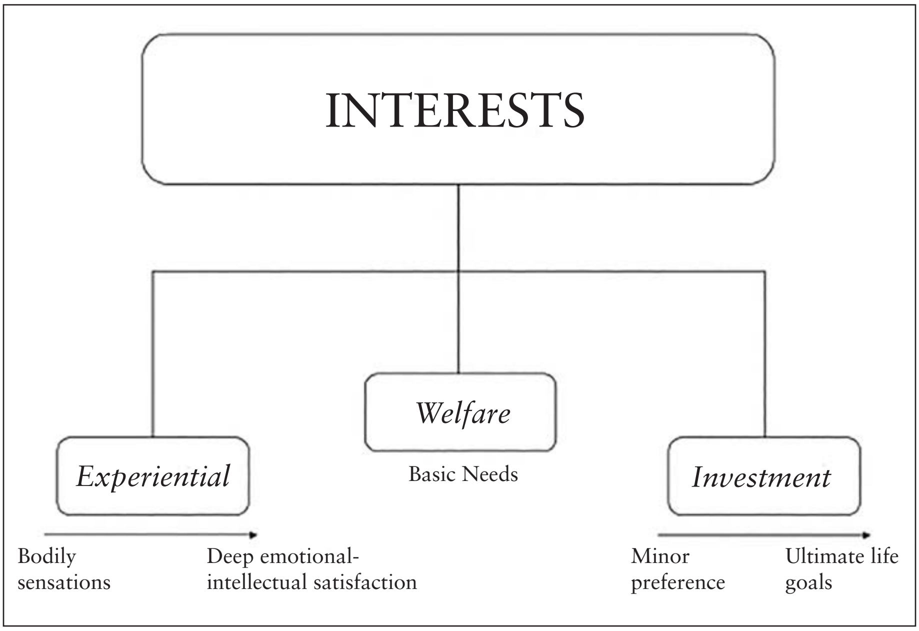


Figure 1. Taxonomy of interests

PRECEDENT AUTONOMY, AND SURVIVING INTERESTS

Roughly, autonomy is the capacity for self-governance, which typically involves the ability to pursue one's interests in light of reasons. (The precise meaning of "autonomy" is of course contested, but nothing substantial in this essay hangs on a more precise analysis.) The principle of respect for autonomy demands that we respect each autonomous person's moral right to guide the course of her own life and to make choices in light of her own reasons and values. This principle posits a sphere of authority over one's interests, akin to the sphere of authority that a nation has over its territory (Mill 1859). Of course, after a person has become severely demented or permanently unconscious, she no longer has the capacity to express her own values and commitments through the choices that she makes, nor to act in accordance with reason. Here, the concept of *precedent autonomy* (Dworkin 1994, 226) can be used to make sense of the idea that a person's past decisions about how she wished to be treated after incompetence ought to be respected, in spite of the fact that she can no longer reaffirm those preferences. This view, which Davis (2007) describes as the *Extension View*, characterizes the moral authority of advance directives as having the same kind of moral authority "as autonomous decisions in general—just extended into the future" (2007, 354).[2]

This standard understanding of the moral authority of advance directives sees them as not merely evidence for what might be good for the later incompetent individual, but rather as acts of self-determination demanding the same kind of respect as any other act of self-determination (President's Commission 1983; Buchanan and Brock 1990). Furthermore, just as the right of autonomy grants a sphere of authority over one's interests, the right of precedent or future-oriented autonomy (Rhoden 1990) grants a sphere of authority over one's *surviving interests*.

Feinberg suggests that "we can think of some of a person's interests as surviving his death, just as some of the debts and claims of his estate do" (1986, 83). Any interest that can still be thwarted or promoted after a person's death can be considered a surviving interest (hence, welfare interests in nourishment and hydration, for example, cannot constitute death-surviving interests). For example:

> the desire to maintain a good reputation, like the desire that some social or political cause triumph, or the desire that one's loved ones flourish, can be the basis of interests that survive their owner's death . . . and can be promoted or harmed by events subsequent to that death. Fulfillment and thwarting of interest, after all, may still be possible, even when it is too late for satisfaction or disappointment [both of which require conscious awareness, as Feinberg uses these terms here]. (1986, 86)

The idea of an interest that can survive death can be readily adapted to the case of an interest surviving incompetence but not yet death. *Incompetence-surviving interests* are the typical subjects of advance directives, which often implicate investment interests in being treated in accordance with a person's understanding of personal dignity, or in maintaining religious commitments. As Dworkin might describe it, incompetence-surviving interests include one's critical interests in the character and value of one's life as a whole, especially the final chapter, since "none of us wants to end our lives out of character" (1994, 213).

The most straightforward means by which a person's authority over her incompetence-surviving interests can be respected is when the incompetent individual had previously made an explicit advance directive. However, even when no explicit advance directive is made, the patient's surviving investment interests may not be simply discarded or ignored. Rather, in the absence of an advance directive, precedent autonomy should still be respected by having a surrogate decision-maker use the substituted judgment standard, attempting to understand what the patient would most likely have chosen, were she able. In this context, the patient's surrogate "stands in for" the patient, and thus is entitled to the same level of relevant information as the patient.

HARMS AND WRONGS AFTER AUTONOMY: A COMPARATIVE ANALYSIS

Consider three individuals who will represent three kinds of cases: Daniel, Veronica, and Christine. Daniel is in a state of end-stage dementia. He is no longer communicative and does not respond in purposeful ways to speech or to the presence of his caregivers; however, he does clearly react to noxious stimuli and shows signs of pain, such as grimacing, moaning, and restlessness. Daniel also appears soothed, and to experience some comfort, when his caregivers give him a gentle back rub or apply lotion to his hands or feet. Daniel is sentient, but lacks any self-awareness; his experiences are fractured and unconnected to each other through memory or a concept of self.[3] Veronica is in a reliably diagnosed permanent vegetative state. Thus, for the purpose of this exercise, assume that Veronica is entirely unaware of self and surroundings; Veronica is biologically alive, but lacks sentience.[4] Finally, Christine is in a state of irreversible apneic coma satisfying diagnostic standards for brain death. She is biologically alive, and is dependent on a mechanical ventilator. She is behaviorally unresponsive to stimuli except for spinally mediated reflexes, and she has no brainstem reflexes. Christine, like Veronica, is biologically alive but irreversibly unconscious.

The Transitivity Argument

One morally relevant difference between Daniel and the other two patients is that Daniel is sentient while the others are not. Therefore, Daniel has incompetence-surviving *experiential* interests, while Christine and Veronica have no further experiential interests that can be either frustrated or satisfied. Daniel has experiential interests in avoiding pain and discomfort, and in experiencing pleasurable bodily sensations, for example, from his back rubs. These interests are at one end of the spectrum of experiential interests, farthest away from the richer kinds of aesthetic, intellectual, and emotional experiences that all three of our patients were once able to enjoy, but can do so no longer. Thus, Daniel's caregivers have beneficent obligations of pain control and the minimization of discomfort towards Daniel, which could not be coherently attributed as obligations to Veronica or Christine. In this regard, Daniel can be harmed by the frustration of experiential interests in ways that the other two cannot. If Daniel experiences pain, then Daniel is harmed. Assuming Daniel's caregivers are duty-bound to minimize his pain and discomfort, then to the extent that they fail in that duty, Daniel is also wronged. Since neither Veronica nor Christine can experience pain or pleasure, they have no surviving experiential interests, and thus can be neither harmed nor wronged from the frustration of experiential interests.

When we examine *investment* interests, on the other hand, we see a very different situation. Other than the potential for frustration or satisfaction of experiential interests due to Daniel's sentience, there are no morally relevant features distinguishing Daniel from Veronica, such that Veronica may be treated differently than Daniel with respect to precedent autonomy or incompetence-surviving investment interests. To the extent that Daniel retains incompetence-surviving investment interests and a right of precedent autonomy over those interests, Veronica also retains them. For example, insofar as it would be morally wrong to continue treating Daniel if he had previously expressed a clear refusal of treatment when he can no longer recognize friends and family based on his understanding of personal dignity, it would be equally wrong to continue treating Veronica had she made the same advance directive. The retention of incompetence-surviving investment interests, along with others' moral obligations to respect acts of precedent autonomy, *has nothing to do with bare sentience*. Rather, these obligations are premised on the assumption that some investment interests, particularly those surrounding "writing the final chapter" of one's life, one's sense of personal dignity, or one's ultimate religious commitments, can survive incompetence. And further, acts of precedent autonomy are genuine acts of self-determination on the same order as other self-determining acts, and therefore, these acts demand respect, even after the capacity to reaffirm preferences is lost.

Just as there are no morally relevant differences between Daniel and Veronica with respect to precedent autonomy and investment interests, there are also no morally relevant differences between Veronica and Christine. Both lack the capacities for sentience, autonomous decision-making, and the reaffirmation of preferences. Thus, other than obligations of pain control due to the preservation of experiential interests, Daniel's moral standing is equivalent to Veronica's, and Veronica's moral standing is equivalent to Christine's. The conclusion therefore follows that *there are no morally relevant differences with respect to precedent autonomy or surviving investment interests between the severely demented and the brain-dead*. This establishes what I will henceforth refer to as the *Equivalence Thesis:* ***Brain-dead patients can be harmed and wronged with respect to***

precedent autonomy and surviving investment interests, if and only if, and to the extent that, patients with severe dementia can.

The Harms and Wrongs of Organ Removal without Valid Consent

In light of the above analysis, I argue that brain-dead patients can be, and some are, both harmed and wronged by lethal organ removal in the absence of *valid* consent. The standards for a valid consent to lethal organ removal are the same for a brain-dead patient as they would be for a patient with severe dementia. The patient (in making an advance directive) or the surrogate (in using the substituted judgment standard) should have a basic understanding of a sufficient amount of relevant information that a reasonable person would want in order to make an informed choice. This would surely include the fact that organ removal would take place while the donor was biologically alive, and removal of vital organs would be the direct and proximate cause of biological death. I argue that in the absence of valid consent, lethal organ removal wrongs the donor as a form of Kantian exploitation and is a violation of bodily integrity. Furthermore, lethal organ removal can potentially harm some donors by setting back surviving investment interests, which I delineate below. Importantly, I do not argue that organ procurement harms and wrongs brain-dead organ donors solely because it causes death. Rather, organ procurement can harm and wrong brain-dead donors because, *as currently practiced*, donors or their surrogates do not give a valid consent to lethal organ procurement.

Kantian exploitation involves treating an individual as a means only to another's ends, without respecting that individual's own ends. It involves treating a person as a thing that can be used to achieve others' ends (regardless of whether they are good ends), while failing to show due regard for the dignity of a choosing, autonomous agent whose own values, aims, and commitments are morally significant. But what are a person's "ends" if not her investment interests? Our ends are our aims, goals, and commitments, and our investment interests are those things that we prefer, value, or in which we are invested. Our critical investment interests implicate our deepest moral, philosophical, or religious commitments to a good life and a good death, and indeed, our critical investment interests are our most important ends—they are our *self-defining* ends. If investment interests can survive incompetence, it follows that ends can survive incompetence. Furthermore, if acts of precedent autonomy demand respect after incompetence, then respect for precedent autonomy demands that the incompetent individual cannot be used as a means only toward others' ends, without taking into account her incompetence-surviving interests (i.e., her ends). This could be satisfied by taking into account explicit acts of precedent autonomy, such as an advance directive. When that is not possible, then respect for the incompetent patient can be satisfied by engaging a surrogate decision-maker to act in place of the patient, deciding as the surrogate believes the patient would have decided, attempting to understand and, insofar as feasible and otherwise morally permitted, attempting to advance the incompetent patient's final surviving investment interests (such as by respecting the patient's understanding of personal dignity, or respecting the patient's religious commitments).

In the absence of valid consent (either first-person prospective consent or surrogate consent), lethal organ removal exploits the donor. The donor may have had strong commitments to helping others as her last act, and would have wanted to donate organs *even though* it caused her biological death. On the other hand, she may have had strong commitments to a religious interpretation of the sanctity of life and would have wanted to donate organs *after* her death, but would have been strongly opposed to being killed by the process of organ removal. From a deontic perspective, it is irrelevant as to what she would have chosen were she (or a surrogate speaking on her behalf) given the opportunity to make an informed choice. Rather, in the absence of valid consent, she (or her surrogate) was simply not given the opportunity to engage in autonomous deliberation in choosing her own ends and acting in accordance with them. Her surviving investment interests—*whatever they may be*—are simply ignored, and in that sense, she is treated as a means only without regard for her ends, hence, this is Kantian exploitation. Notice that this argument has nothing to do with sentience but is grounded entirely in precedent autonomy and surviving investment interests. Since patients who are severely demented and patients who are brain-dead are morally equivalent in this regard, it follows that lethal organ removal in the absence of valid consent wrongs brain-dead patients, with respect to precedent autonomy, to the same extent that it would wrong severely demented patients.

Patients with dementia, no matter how severe, retain certain basic moral rights, such as the right to bodily integrity. This entails that violating bodily integrity by removing a patient's organs may be done only if it is for the patient's benefit or at least with valid consent, either prospective or via surrogate. This widely accepted deontic constraint is not grounded in surviving experiential interests and, with the exception of experiential interests, the severely demented and the irreversibly comatose are of equivalent moral standing. It therefore follows that the brain-dead retain the right to bodily integrity if and only if, and to the extent that, the severely demented do. Therefore lethal organ removal in the absence of valid consent is an unjustified violation of bodily integrity, and is a form of medical battery and theft of organs. As above, this wrongs the irreversibly comatose patient to the same extent that it would wrong the demented patient.

Lethal organ removal also has the potential to harm donors by setting back surviving investment interests. Consider Consuela, who is in an irreversible coma satisfying criteria for brain death. Consuela strongly valued independence, self-control, and self-determination. These traits formed the core of Consuela's narrative self-understanding. She was someone who "made her own destiny," and she strongly valued her autonomy; this included a desire to have some say about the end of her own life. Consuela has a surviving investment interest, very simply, in *having a say* over what happens to her at the end of her life's narrative, in writing her final chapter. By causing Consuela's death by organ removal without valid consent (either prospective or surrogate), she is harmed by setting back her surviving critical investment interest in having some control at the end of life. Failing to disclose relevant facts about organ removal that any reasonable person would want to know not only wrongs Consuela; it also harms her by frustrating her critical interest in ending her life's narrative "in character."

Charles also satisfies criteria for brain death. Charles had a lifelong, self-defining commitment to a particular religious worldview, which included an understanding of human dignity as a commitment to the sanctity of human life. In line with this worldview, Charles wanted to donate his organs to help save others, but only after "natural death," by which he meant that withholding excessively burdensome or "extraordinary" treatment is permissible, but that taking any action that directly causes death is a serious moral and religious violation. Charles would not have agreed to lethal organ removal if he had been given accurate information, particularly that organ removal would cause his biological death. Like Consuela, Charles is both wronged and harmed by lethal organ removal in the absence of valid consent, because it sets back his critical investment interest in ending his life in accordance with his religious commitments. Dworkin's memorable phrase—"making someone die in a way that others approve, but he believes a horrifying contradiction of his life, is a devastating, odious form of tyranny"—would ring particularly true with Charles (1994, 217).

As mentioned above, recent survey data suggest that most (though not all) individuals in the US who would support organ donation after death would also support lethal organ procurement in a scenario of irreversible coma that violates the dead donor rule (Nair-Collins et al. 2015). Consider then Chanda, who satisfies brain death criteria. Chanda signed up as an organ donor when he got his driver's license, but never understood that organs would be removed while he remained biologically alive. However, organ donation after brain death is consistent with his values and preferences and, had he had better information, Chanda would have agreed to lethal organ procurement. In this case, Chanda has no surviving investment interests in not being killed by organ procurement. Rather, he has a surviving investment interest in being an organ donor. Hence, by advancing this surviving interest, Chanda is benefited by lethal organ procurement.[5] However, he is still *wronged*.

To see how Chanda is wronged, consider this analogy. Assume that, if asked, I would gladly accede to a request from a stranger for a cutting of my fig tree. Assume further that I was going to prune that part of the tree anyway. However, if that same person were to walk into my yard without asking permission and remove a small cutting of my fig tree, we must agree that the fig tree cutting was stolen from my yard, and that it was my fig tree cutting to keep or give away, not the stranger's fig tree cutting to remove without permission.[6] This point is not rendered moot by the fact that, *had I been asked*, I would have surely given a cutting. In a similar fashion, like many and perhaps even most brain-dead organ donors, Chanda would have agreed to donate even if he had been given accurate information, however, this does not obviate the point that he was never really asked.

One might respond that, unlike my fig tree example, Chanda was asked, and he did authorize organ procurement, even though it may not have been a completely informed decision. In response, consider a second analogy. Imagine that I ask if you wouldn't mind pet-sitting my cat for the weekend. Being a cat enthusiast, you agree, so I give you my key. When you arrive at my home, you find my 70-pound, slobbery, exuberant Boxer named Mac, who has a penchant for giving doggie kisses, sitting on your lap, and sleeping in your bed. When we speak next, you are annoyed, stating that you had agreed to pet-sit a cat, not a dog. You would have agreed to pet-sit my dog if I had asked, but you are (justifiably) irritated that I did not ask. I respond, “You see in the context of pet-sitting, I use the word ‘cat’ to mean *dog*. Although in most other contexts ‘cat’ still means what it usually does, in this context it has changed somewhat. You did agree to pet-sit my cat (which really means *dog*, I add *sotto voce*), so this isn’t the same as if you stayed at my home without any agreement to pet-sit at all. You authorized your pet-sitting services.” Plainly, this response demonstrates sophistry at best, and deceitful manipulation at worst. Most certainly, you did *not* agree to pet-sit a dog, because I asked if you would pet-sit a cat. The fact that you would have agreed to pet-sit a dog had I asked does not change the fact that I never actually asked.

A parallel situation occurs in the organ donation context, where autonomous moral agents with a right of precedent autonomy over their incompetence-surviving investment interests, like Chanda, are encouraged to donate their organs “after death.” In this context, “dead” means something other than what it usually means outside of the organ procurement context (Nair-Collins 2013). Here, “dead” means *not biologically dead, but in an irreversible apneic coma where organ removal will cause biological death*. Like many if not most organ donors, Chanda would have agreed had he been asked. However, as in the pet-sitting analogy, he was never really asked. An agreement to donate organs *after death* does not constitute valid permission to take organs *prior to death*, any more than an agreement to pet-sit a cat is a valid agreement to pet-sit a dog. We are led back to the fig tree analogy: As with the stranger taking a cutting of my tree without valid permission, vital organs are often removed without *valid* permission. Like the fig tree cutting, only in a far more serious way, Chanda’s organs are his to give, but not ours to take without valid permission. In light of the Equivalence Thesis, it is worth a reminder that the organs of the brain-dead are no more ours to take without valid permission than are the organs of a patient with severe dementia.

OBJECTIONS

Brain Death Marks a Moral or Social Category

One might object that my arguments are misguided because the relevant concept of death marks a moral or social category, not a purely biological one. In other words, the concept of death in this context is “morally thick,” where normativity and views of moral status are built-in, and it need not refer to a strictly biological category delimited by homeostatic capacity. On this morally thick concept of death, patients in irreversible apneic coma *really are* dead, whereas patients with severe dementia are not. Therefore, it is justified to treat the brain-dead as dead for clinical, legal, and policy purposes.

There are two versions of this idea. First, we might claim that the concept of death straightforwardly embodies a claim about moral status and thus normativity is analytically built-in to the concept. For example, Veatch argues that in the organ donation context, the word "dead" has come to have a social, moral, and legal meaning that is divorced from its usual biological meaning:

> A normative, policy use of the term can define 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. . . . One first identifies . . . who is no longer part of the community in the full sense, that is, those not protected by laws against homicide . . . etc., and then calls that group *dead* by definition. (2015, 297)

Thus, to call brain-dead patients "dead" in this context *just is* to assert that they no longer have moral standing; thus the moral difference between the brain-dead (who are not in the moral community, on this view) and the severely demented (who are) is established.

The second version of this view holds that *death itself* (as opposed to the meaning of the word "death") is in some important sense a matter of social choice, or it is socially constructed, or biology does not completely determine the boundaries of the concept of death. Because of this important ambiguity, social and normative concerns play an appropriate role in choosing a point at which we will call patients dead for social and legal purposes. Khushf (2010), for example, argues that biological considerations underdetermine the truth of statements about death, and thus normative considerations play a legitimate role in choosing a precise point within vague boundaries that we will call patients dead. In a similar fashion, Brody (1999; Halevy and Brody 1993) argues that life and death are fuzzy sets, and therefore it is appropriate to choose different points in the dying process as answers to different questions, such as when life support may be unilaterally withdrawn, when organs may be removed, and when burial or cremation is appropriate. As on Veatch's view, these views of death are morally thick because normativity is built in, thus, the moral difference between the severely demented (who are not dead) and the brain-dead (who are dead, on these views), is established. Importantly, these views do not merely state that our laws and policies surrounding death are a matter of social choice (which they surely are); rather, this family of views holds that *death itself* is partially a function of human choice.

In the dialectic of this paper, the mere assertion that normativity is analytically built-in to the concept of death, *without giving further reasons for why the brain-dead lack moral status*, is question-begging. *Whether* brain-dead patients can be harmed or wronged, and *whether* they have significant moral standing, are precisely what is at issue. If this objection were to consist of no more than simply stating that "death" now has a normative meaning, or that death itself is a reflection of a normative social choice, and further, the brain-dead are dead (on this normative meaning of "dead"), then it would not be a cogent objection on grounds of question-begging. Instead, we must directly confront the *reasons* for believing that brain-dead patients have less or different moral status. According to my arguments above, with respect to precedent autonomy and surviving investment interests, patients who satisfy diagnostic criteria for brain death are morally equivalent to patients with severe dementia.

To be clear, I am not suggesting that the authors cited above give no reasons for their views; those reasons will be addressed in the following sections. Rather, I am asserting that the objection that death is normative and socially constructed, or that "death" has a normative meaning, is insufficient on its own: The *reasons why* brain-dead patients are thought to lack moral status, rights, or interests must be evaluated on their own terms.

Personal Identity

Theories of personal identity seek to identify our persistence conditions: What (if anything) makes it the case that I am the same person that I was yesterday, or even 20 years ago, and what would make it the case that I will be the same person 20 years from now, in spite of the numerous physical and psychological changes associated with the passage of time? Identity across time is closely related to the question of what we essentially are: If we have some essential nature or set of essential properties, then a change in that nature or loss of those essential properties would entail that we have ceased to exist. Although some authors distinguish personal identity from *human* identity, a theory of the persistence conditions for human beings (as opposed to *persons*), I will not make this distinction here. Following McMahan (2002, 5–6), I will use "personal identity" to refer to the persistence conditions for beings-like-us, whatever kinds of beings we are. This terminology does not presuppose that we are essentially persons in a Lockean sense, nor embodied minds, nor human organisms, but is neutral on what kinds of beings we are.

Even though brain-dead patients are not biologically dead, one might object that there is nonetheless an important ontological difference with respect to personal identity, and this ontological difference—not mere biology—morally justifies treating brain-dead patients differently than other patients, including those with severe dementia. In this section, I review three leading theories in this domain—psychological continuity, bodily continuity, and narrative identity—and argue that, for each of these views, each of the three patients discussed above all fall into the same ontological category with respect to personal identity. Hence, we cannot draw moral differences between them on this basis either. I will discuss the important view that we are essentially embodied minds in the following section.

Locke held that psychological continuity is necessary for personal identity across time (1690; cf. Uzgalis 2014), and Parfit extended and clarified this philosophically influential view to include a distinction between continuity and connectedness (1984). On Parfit's view, a given entity at time T_2 is the same person as some given person at T_1 only if there is either a minimum degree of shared memories and other psychological properties between the two (connectedness), or, if there is some minimum degree of overlapping chains of these shared memories and psychological properties (continuity). Whatever minimum degree of continuity or connectedness is required, if that continuity or connectedness fails to obtain between two entities across times, then the entity at T_2 is not the same person as the entity at T_1. In this instance, either the entity at T_2 is a person but is not the same person as at T_1, or, the entity at T_2 simply fails to be a person altogether. The view of personhood (that is, what makes an entity a person vs. not a person) that is most consonant with this understanding of personal identity is also historically associated with Locke. This view identifies certain cognitive capacities as necessary conditions on personhood, particularly self-consciousness or the ability to be conscious of oneself *as a self* with both a past and a future; the view also includes capacities such as the ability to appreciate reasons for acting and the ability to act purposefully in purposive sequences of actions (Buchanan 1988, 284).

By contrast, animalist theories, associated with bodily continuity views, hold that we are essentially human organisms (DeGrazia 2005), and further that two entities are numerically identical across time only if there is a minimum degree of physical, bodily connectedness. On this view, identity can survive complete disruption of psychological continuity, so long as bodily continuity is intact and the organism remains biologically alive.

Finally, narrative views see persons as their stories, not just as psychological states or bodily sensations (Kuszewski 1999, 33). On this family of views, the concepts of person and personal identity are relational concepts that involve not just the individual's self-construction of her own life narrative or life plan, but also the social and interpersonal relations that an individual bears to others, who form key components of the story. Furthermore, one's death is an integral part of one's story (Kuszewski 1999, 33), and hence, on narrative views—in similar fashion to bodily continuity views—personal identity can survive considerable, and even complete disruption of psychological continuity.[7]

Now let us reconsider Daniel, Veronica, and Christine. According to the psychological continuity view, each of these three individuals is non-identical to the person who was previously embodied in, realized by, or instantiated by the body that we now refer to as "Daniel," "Veronica," or "Christine." For Veronica and Christine, they have no remaining psychological states, and so it is clear that those individuals (now, after the illness or injury that left them unconscious) are non-identical to the persons they previously were, and furthermore, they are not *different* persons now; rather, they are not persons at all, on this view. However, the same conclusions follow with respect to Daniel. Although Daniel retains some bare sentience and thus the capacity to experience bodily pain and pleasure, he cannot recognize friends, family, or caregivers. He does not share memories, goals, life plans, or personality characteristics with the person that once was instantiated in this body. The disruptions of Daniel's psychological continuity and connectedness are so severe and so complete that this individual is not identical to the previous person, according to this view of personal identity. Furthermore, because of the irreversible neurological pathology, Daniel (now, after severe dementia) is no longer capable of a concept of self-consciousness, or an awareness of himself as existing with both a past and a future; nor is he able to appreciate reasons for acting; nor is he able to act purposively, as he is dependent on caregivers for all needs, including his most basic needs. It follows that on this view, Daniel is not a person at all, just as both Veronica and Christine are not persons.

On the bodily continuity and narrative views, Daniel, Veronica, and Christine are each numerically identical to their individual selves prior to their illness or injury. Each of them certainly shares enough physical continuity with his or her prior self to satisfy the bodily continuity view. Furthermore, each of their stories is continuing, and will continue until death. The family members, friends, and caregivers who played critical roles as characters in their narratives continue to do so; and their life stories, although now in their final chapters, are nonetheless not yet complete.

My aim here is not to endorse or challenge any view of personal identity. Rather, I simply want to show that, on several leading theoretical contenders, the severely demented and the brain-dead *are in the same ontological category* with respect to personhood or personal identity, and therefore, moral differences cannot be drawn between them on these grounds. This provides further support for the Equivalence Thesis.

The Moral and Ontological Significance of Sentience

The crux of my argument is that there are no morally relevant differences with respect to precedent autonomy or incompetence-surviving investment interests between the severely demented and the brain-dead. Because of this, the different ways that these patients are treated is morally arbitrary. To justify the differences in practice, one needs to demonstrate that morally relevant distinctions underlie the differential treatment.

Biology will not do this, because patients in irreversible apneic coma meeting criteria for brain death are biologically alive. Personal identity will not do this, at least if any of above three views are correct, because on these theories of personal identity, the severely demented, vegetative, and irreversibly comatose all fall into the same ontological category. A morally thick concept of death will not do this, because *whether* there are relevant moral differences between the patients is precisely what is at issue; hence simply stipulating by definition that there are such differences is question-begging. However, sentience distinguishes severely demented patients from the other two classes of patients. This is a morally relevant difference with respect to experiential interests, as I've explained above, so patients with dementia can be harmed and wronged with respect to experiential interests while brain-dead patients cannot. Does sentience go further, and ground differences with regard to investment interests and precedent autonomy in a way that would justify the different treatment?

McMahan (2002) argues that we (or beings-like-us) are essentially embodied minds, distinct from our organisms. On this view, the neurologically instantiated capacity for sentience is a necessary and sufficient condition for our existence. McMahan also endorses a Lockean conception of personhood, requiring self-consciousness and a complex mental life, but argues that while we are *essentially* embodied minds, we are only *contingently* persons. This means that personhood is a phase of the embodied mind but is not essential for its existence, so the very same embodied mind (or “self”) could cease to be a person while continuing to exist so long as the capacity for consciousness is preserved. Finally, because we are distinct from our organisms, we need two conceptions of death: one for the organism, and one for the embodied mind. McMahan endorses the standard understanding of the biological death of an organism in terms of the integrated functioning of the organism’s subsystems, and thus acknowledges that brain-dead patients remain biologically living organisms. The embodied mind or the self, however, ceases to exist at the irreversible loss of the capacity for consciousness: “when [an embodied mind] ceases to exist by losing the capacity for consciousness, he dies. If this is a technical notion of death, so be it” (McMahan 2002, 425).

On this “higher brain view of death,” Veronica and Christine have died because, in virtue of lacking the capacity for consciousness, both embodied minds no longer exist, although they have left behind living organisms. By contrast, Daniel remains alive by virtue of his preserved capacity for sentience. Although Daniel is not a person, he is numerically the same individual he was prior to the onset of dementia, when he was still a person. In a similar fashion, Veatch argues that “proponents of [the ‘social meaning of death’ view] would consider total and irreversible unconsciousness the basis for loss of full moral status as a member of the human community” (2015, 298). Since Veatch argues that “death” now means something like *lacking full moral status as a member of the human community*, he would agree that Veronica and Christine have died, even though their organisms remain biologically alive, whereas Daniel remains alive, even though Daniel is no longer a person.

A proponent of the higher brain view of death might object that, surely, there *is* a morally relevant difference between Daniel and the other two patients: Daniel is alive and continues to exist; the other patients have died and no longer exist. Furthermore, one might continue, although in theory we would consider a vegetative-state patient like Veronica to be dead, in practice the diagnosis of unconsciousness is fraught with problems. To avoid false positives, it is appropriate to use physiologically conservative diagnostic criteria such as those currently associated with “brain death.” The philosophical justification for this practice would rely on the higher brain concept of death paired with conservative diagnostic criteria to eliminate false positives. Finally, the concern that donors or their surrogates do not provide valid consent for organ donation is vitiated by the higher-brain theory: Although the organism has not died in brain death, what matters is the self, the being-like-us, the thing that we essentially are, and that has indeed ceased to exist, or died, in brain death.

I believe that this is the most plausible objection against the view that I'm defending; however, it is not successful. For clarity of exposition, in this section I'll henceforth use "hb-death" and its cognates to refer to higher brain death, or the ceasing to exist of the self by virtue of loss of the capacity for consciousness. I'll use "bio-death" and its cognates to refer to the biological death of the organism.

To begin, consider the following case. Desiree has end-stage dementia. She can no longer recognize the people around her, she does not share connections to her past self through memory, life plans, or personality traits, and her present sentient experiences are fractured and unconnected to each other through memory or self-consciousness. When she was competent, Desiree explained to her family and physician that she would find it a grave affront to her sense of personal dignity to be kept alive in a condition when she could no longer recognize her family and friends. She explained that once she reaches this stage of cognitive decline, she wanted to be kept comfortable by being fed by mouth and having pain medications if indicated, but she refused any other life-sustaining treatments, including antibiotics or artificial nutrition. Overnight, Desiree suffered a stroke that obliterated her brain's remaining capacity for sentience, but it did not kill her and some brainstem and autonomic regulatory functions remain intact as well. On Tuesday (before the stroke), Desiree was barely sentient but not capable of forming or reaffirming preferences, not capable of understanding what is happening around her and to her, and unable to act intentionally or purposively. On Wednesday (after the stroke), Desiree remains incapable of acting purposively or appreciating what is happening to her, but now lacks the capacity for bare sentience as well.

The higher brain theorist would argue that Desiree is alive on Tuesday but hb-dead (though bio-alive) on Wednesday. However, surely on Wednesday, just as on Tuesday, her physician and family ought to take into account her previous autonomously formed preferences, and may not simply ignore them. This widely held assumption is reflected in the practice of honoring advance directives for patients in a vegetative state. And indeed, McMahan agrees. In considering the possibility of organ donation in the vegetative state, McMahan's views are worth quoting at length:

> . . . much depends on the prior wishes of the person whose body it was. If the person had expressed a preference about how his body should be treated in a [permanent vegetative state], there is a presumption that the preference should be honored. This, then, is the second possible objection to the removal of organs from an organism in a [permanent vegetative state]: that to do so would be a failure of respect, albeit posthumously, *for the person*, if he had preferred that the organs not be removed. It is, I think, largely irrelevant what a person's reasons might have been. . . . Even if his reasons were bad ones, there is a strong presumption that they must be respected. Suppose, for example, that a person believes that he would continue to exist in a [vegetative state] and that his life would continue to have value to him in that state. Even if I am right that he is mistaken on both counts, his preference is probably decisive. (2002, 447–48, emphasis in the original)

On Wednesday, surely Desiree's family and physician still ought to act in accordance with her prior preferences (absent, let us assume, some countervailing reason), and McMahan would agree. But if this is so, then why? Certainly nothing has changed on Wednesday to provide some new justification, which didn't already exist on Tuesday, for the authority of her prior wishes. Therefore, it must be that the same justification exists on both days: Because that is what she autonomously preferred when she was able to make those determinations. To disrespect her wishes now is to set back her surviving investment interests in concluding her life's narrative as she saw fit, and in accordance with her sense of dignity. Thus, again, the right of precedent autonomy over surviving investment interests *is not grounded in bare sentience*.

This shows that the higher brain theorist should conclude that critical investment interests can survive not only incompetence but also hb-death. Furthermore, the higher brain theorist should also accept that the right of precedent autonomy extends beyond hb-death. It follows that patients can be harmed by setting back surviving interests, though the higher brain theorist would describe this as "posthumous harm," where "posthumous" means *after hb-death*, though not necessarily after bio-death. But regardless of how we describe it, the very same deontic moral concern, grounded in the authority of her previous autonomously formed preferences, applies to Desiree on Tuesday as it does on Wednesday. It follows, again, that with respect to precedent autonomy and surviving investment interests, patients with severe dementia are morally equivalent to patients in the vegetative state. And since the vegetative state is morally equivalent to irreversible coma meeting brain-death criteria, our conclusion once again follows by transitivity. The higher brain theory does not supply a morally relevant distinction (other than with respect to experiential interests) between patients in severe dementia and brain death.

The second objection under this framework was that hb-death is what matters, not bio-death, and therefore donors or their surrogates are not misled when they agree to organ donation. But this misconstrues the point of respecting autonomy. Whether beings-like-us are essentially embodied minds, ontologically distinct from our organisms, is a deep metaphysical question about which reasonable people, including professional philosophers, can and do reasonably disagree. It is a further deep normative question as to whether hb-death or bio-death are what matters. Even if, objectively, McMahan's view (or some variant) is correct, it is a complex metaphysical view, fundamentally tied to a large-scale philosophical worldview. Respect for autonomy would not demand forcing this view on other autonomous moral agents who reach different conclusions (indeed, quite the opposite; see the immediately preceding quote for McMahan's agreement here). And it is worth noting that what might be described as the "mainstream" view holds that bio-death, not hb-death, is what matters in this context (Bernat 2006; President's Council 2008).

Thus, people retain the right to make reasonably informed end-of-life decisions about organ donation, which includes the right to have, at minimum, basic information in clear, non-technical language. Even though many would presumably agree to donate organs in irreversible coma, respect for persons nonetheless demands that those persons be allowed to make that choice for themselves.

Posthumous Harm and Absurdity

Patients in an irreversible apneic coma meeting brain death criteria are irreversibly unconscious, permanently lacking the capacity for autonomous choice as well as sentience. But dead patients are also irreversibly unconscious, and permanently lack the same capacities. The only difference between brain death and death is that patients in the former category retain a homeostatic capacity of the organism as a whole, while those in the latter do not. But why should this mere biological difference make for a moral difference? This consideration may seem to threaten absurdity in two ways. First, if there are no morally relevant differences between the brain-dead and the dead, then transitivity obtains in the other direction, and we must conclude that patients with severe dementia are morally equivalent to dead patients (with the exception of experiential interests). This may seem absurd or even offensive to some. Second, if there are no morally relevant differences between the brain-dead and the dead, and if the brain-dead retain significant moral status that can create obligations on us, it may seem that this can be extended to absurdity, with all of the countless dead posing endless moral obligations on us, potentially yielding a second *reductio ad absurdum*.

I suspect that there is no deep moral divide separating the brain-dead from the dead, because mere biological categories usually do not yield morally important distinctions. However, for the moment I remain agnostic about what moral significance, if any, attaches to biological human life as such. At this point, I propose to simply assume that there are no morally relevant differences distinguishing the dead from the brain-dead, and see if the implications truly threaten my thesis.

Granting our assumption that there are no morally relevant differences between the dead and the brain-dead, it must follow that the dead are morally equivalent, with respect to precedent autonomy and surviving investment interests, to the severely demented. Some may consider this to be counterintuitive or even disrespectful, but I believe that it is neither. Patients with severe dementia are not equivalent to the dead with respect to experiential interests, since the dead have no experiential interests. Beneficence demands concern for the experiential interests of patients with severe dementia, and thus a judicious attention to their experiential well-being. This equivalence does not entail that patients with dementia may be ignored or treated as if they were dead bodies. On the other hand, most people also believe that the dead are entitled to respect, and indeed, as Feinberg (1986) argues, that some interests can survive death. If this is right, then there is no objection to acknowledging that the dead can retain surviving investment interests, just as do patients with severe dementia.

But this brings us to the second objection. If we acknowledge that the dead have a moral status with respect to investment interests that is equivalent to the severely demented, does this not create unlimited and absurd moral obligations to the dead? It does not. Respect for persons is largely (though not entirely) a negative obligation, an obligation to avoid interfering, insofar as otherwise morally permitted, in the important choices of other people's lives. But other people's autonomously formed choices do not exert unlimited authority on others even when they are alive and well. We are not obligated to do absurd things for each other even while alive, so it would not follow that we are obligated to do absurd things for the dead either.

Surely there is a great deal more to be said about posthumous harm, but here I simply want to show that a possible moral equivalence between the brain-dead and the dead does not threaten my thesis with absurd implications.

The Extension View and the Equivalence Thesis

An objector might simply reject the claim that a severely demented patient such as Daniel has any surviving investment interests. Since Daniel lacks the cognitive or behavioral abilities to advance, or even to remember, his previously endorsed investment interests, it follows that he is actually not invested in them anymore at all. Thus it makes little sense to ascribe a right of "precedent autonomy" over interests that no longer exist. Similarly, the objector might argue, neither Veronica nor Christine has any investment interests either, and for the same reason: They are no longer *invested in them*.

This objection rejects the standard Extension View in the bioethical literature, which I have assumed but not defended here (cf. Dresser and Robertson 1989). For the objector who rejects the standard view that we have deontic obligations to severely demented patients with respect to advance directives or precedent autonomy, I simply concede this point and retreat to the conditional thesis: *If* one assigns any significant moral status to the severely demented with respect to investment interests (not experiential interests), then one must also do the same to the brain-dead with respect to investment interests. Recall that the Equivalence Thesis holds that, other than experiential interests, the brain-dead can be harmed and wronged if and only if, and to the extent that, the severely demented can. That statement of material equivalence is not falsified by rejecting the assertion that the severely demented have rights of precedent autonomy. It is worth noting however that law and policy in the US clearly recognize incompetence-surviving investment interests and a right of precedent autonomy over them. This is embodied in advance directives, surrogate decision-making and the substituted judgment standard, and the Patient Self-Determination Act. For those who find these practices ethically justifiable, it is arbitrary to allow a right of precedent autonomy to the severely demented but not to the brain-dead.

One might also wonder why the Equivalence Thesis should be taken to show that organs may *not* be removed from brain-dead patients; why not, instead, take it to show that organs *may* be removed from patients with dementia? But I have nowhere argued that organs may not be procured from brain-dead patients. It's not the lethality of organ procurement that creates harms and wrongs; it's the lack of valid consent. With valid consent and appropriate safeguards in place—I make no claim that I know what those would be—lethal organ procurement from patients with severe dementia would have to be ethically justified in the same way that it would be from patients who satisfy brain death criteria, at least if the Equivalence Thesis is true.

My aim in this essay, however, is not to argue for organ procurement from patients with severe dementia. It is to show that there is an important equivalence between patients with severe dementia and patients with brain death. If one cannot even in principle endorse validly consented, appropriately safeguarded lethal organ procurement from patients with severe dementia, then on pain of arbitrariness one cannot endorse validly consented, appropriately safeguarded lethal organ procurement from brain-dead patients.

Back to the Social Construction of Death and the Need for a Bright Line

Finally, one might object that law and policy need brain death as a "bright line," because without it, we will create unnecessary legal and clinical problems, and potentially undermine support for organ transplantation. For example, Magnus and colleagues write, "If brain death were not defined as death, it would be more difficult to justify routine decisions to discontinue mechanical support [for patients meeting brain death criteria]. . . . Rejecting brain death . . . might undermine support for cadaveric organ donation" (2014, 3). This objection relies on a version of the idea that death is a matter for social choice rather than biological reality, and indeed Magnus and colleagues refer to "the important social construction of death when the brain has ceased all meaningful activity" (2014, 3). In a similar fashion, Brody (1999) has argued that life and death are fuzzy sets with no precise boundary between them, so it is appropriate to choose points in the dying process as answers to normative questions, such as when organs may be removed. And Khushf (2010) has argued that biological reality underdetermines the truth of our statements about death, but considerations based on respect for persons and the need for transplant organs may appropriately be used in choosing a precise point of death from an otherwise arbitrary continuum.

I take the idea that death is socially constructed to be mistaken. Like life, death is a phenomenon that is part of the natural, physical world, and thus is amenable to naturalistic explanation using concepts and categories consistent with other well-accepted physical, chemical, and biological theories. Furthermore, even granting the possibility of vague cases, it does not follow that the brain-dead fall into some vague ontological category without further argument (Nair-Collins 2010). But the evidence shows quite clearly that these patients' status with respect to *biological* life is not at all vague—they are quite clearly biologically alive. However, let us assume for the sake of the argument that death is partially a function of human choice, and thus that we must decide on a line that will separate the living from the dead.

In evaluating some proposed line separating life from death, we must ascertain whether that line is morally arbitrary, or if it creates policies and practices that treat patients differently when they are similar in morally relevant ways; otherwise the proposed distinction will be unjust. As I've argued above, with respect to precedent autonomy and surviving investment interests, brain-dead patients are morally equivalent to patients with severe dementia. Furthermore, brain-dead patients can be harmed, wronged, and exploited by end-of-life practices, just as patients with severe dementia can. It follows that drawing the distinction between life and death at brain death is morally arbitrary, and hence unjust. Therefore, even granting the assumption that death is partly a function of human choice, the objection does not work, because *whether* brain death is a justified place to draw the line depends on normative questions about harms, wrongs, and moral status which are, again, precisely what is at issue.

In the development of law and policy *surrounding* death (which surely is a function of human choice), there is a need for clarity if not bright lines. But we must not conflate the need for *some* bright line with the need for *this particular* bright line. The need for clarity in public policy does not entail that simply any policy whatsoever would be morally justified. Many believe that brain death is an appropriate place to draw the relevant legal and policy distinctions; however, it is usually overlooked just how similar brain-dead patients are to many other patients, whom no one would call dead, and whom few would claim lack sufficient moral status as to have, at minimum, a simple right to bodily integrity. The mere existence of a (current) bright-line legal distinction is no guarantee that that distinction tracks any morally relevant differences. American law once drew a variety of bright legal lines separating people based on race or gender, but those legal distinctions, in and of themselves, were never sufficient for guaranteeing legitimate moral differences; indeed, they were unjust laws *precisely because* they drew legal distinctions that treated people differently based on morally irrelevant differences. In a similar fashion, the bright line separating patients with brain death from all other patients is morally arbitrary, and it creates unjust policies and practices.

CONCLUSION

The idea that brain-dead patients are different than all other patients—in morally significant ways that justifies treating them very differently in the clinic and in public policy—runs deep in the bioethics literature and even deeper in medical and legal practice. In this essay I've argued that the assumption is mistaken. Patients meeting brain death criteria retain significant moral standing, including a right of bodily integrity and a right of precedent autonomy over incompetence-surviving investment interests, and indeed in these respects, brain-dead patients are morally equivalent to patients with severe dementia. Furthermore, brain-dead patients can be, and many are, harmed and wronged by heart-beating organ removal in the absence of valid consent.

As a result of these arguments, I would urge that systems for valid consent for organ procurement be put in place as soon as possible. And in the meantime, we should not fool ourselves into thinking that current practices in organ procurement from brain-dead donors are morally unproblematic. The significant moral good of saving and improving the quality of many lives through organ transplantation comes with the significant moral cost of harming and wronging many organ donors.

ACKNOWLEDGMENTS

I've presented earlier versions of this essay at the inaugural meeting of the International Association for the Philosophy of Death and Dying in 2014, at the World Congress of the International Association for the Philosophy of Law and Social Philosophy in 2015, and at the annual meeting of the American Society for Bioethics and Humanities in 2015. I sincerely appreciate the opportunity for thoughtful discussion with the audience at each of these conferences. I'd especially like to thank Pablo de Lora for inviting me and arranging a special workshop on Death, Philosophy, and the Law at the International Association for the Philosophy of Law and

Social Philosophy. I also thank Marshall Kapp, Ana Smith Iltis, Frank Miller, and the anonymous peer reviewers for this Journal, each of whose incisive comments and criticisms helped to strengthen the manuscript.

NOTES

1. While there are many interesting and important open questions regarding death and brain death, it is no longer a reasonably open empirical question as to whether mechanically ventilated brain-dead bodies preserve the capacity to maintain homeostasis of the organism as a whole. Assuming the standard homeostasis-based theory of death, it is therefore not a reasonably open empirical question as to whether brain-dead patients remain biologically living. This point has been convincingly and repeatedly made by many authors (e.g., Shewmon 2001; Miller and Truog 2012). For those readers who remain unconvinced, at minimum it must be acknowledged that this is a live option in the literature on death. It is therefore an important exercise to map out the implications if brain death is not sufficient for biological death. Finally, it may be noted that while all patients who meet diagnostic standards for brain death are in an irreversible apneic coma, not all patients in irreversible apneic coma meet criteria for brain death, since a brainstem reflex could be preserved.
2. In writing this manuscript I have benefited greatly from Davis's (2007) lucid overview of the advance directives literature. In this section I draw heavily from Davis's articulation of the literature.
3. To be clear, I am stipulating that this particular patient, Daniel, has the characteristics described in the text, of lacking self-awareness and so on. I do not assume that the condition of severe dementia necessarily entails these characteristics, because it does not ("severe dementia" is not a clearly defined nosological category in the first place). To clarify my thesis, I argue that patients with brain death are morally equivalent, in certain ways to be delineated below, to those patients with severe dementia who are like Daniel, in lacking self-awareness or a concept of self but retaining what I'm calling "bare sentience."
4. I emphasize that this is a philosophical fantasy, assumed for the purpose of conceptual analysis. The vegetative state, that is, the state of wakefulness without awareness, is not reliably clinically diagnosable, with a consistent clinical false positive misdiagnosis rate of about 30–40% (cf. Giacino et al. 2014 for a recent review).
5. Like Consuela, Chanda may have also had a surviving interest in having a say about how his life ends, in which case he would have been simultane-

ously harmed by setting back his interest in having some control at the end of life, while benefited by advancing his interest in organ donation. For this example, we may stipulate that, on balance, Chanda would have considered organ donation a benefit since it would have been more important to him that he donate his organs. And let us also stipulate that in Consuela's case the opposite holds, where having some control at the end of her life was more important to her.

6. I here leave aside an important issue, namely, that food is a right not a privilege; so if I have excess food and the stranger is hungry, it is not quite so clear that the stranger has done anything wrong in the scenario envisaged. But this is peripheral to our concerns in this essay, so let us assume that the stranger has as much access to food as do I.
7. I have presented only the briefest gloss of each of these theories, and there are a variety of subtle variations within each camp, which I won't review here. It is worth noting however that a different variety of narrative identity could plausibly take self-construction of one's own narrative to be a necessary condition for continued narrative identity. This would require a complex mental life and in particular the concept of a self, hence, this variety of narrative identity theory would be coextensive (though not intensionally equivalent) to the Lockean theory. Because of the coextensiveness, this version of narrative theory would categorize the three classes of patients in the same way as the Lockean theory would.

REFERENCES

Bernat, James L. 2006. "The Whole-Brain Concept of Death Remains Optimum Public Policy." *The Journal of Law, Medicine, & Ethics* 34 (1): 35–43.

Bernat, James, Charles Culver, and Bernard Gert. 1981. "On the Definition and Criterion of Death." *Annals of Internal Medicine* 94: 389–94.

Brody, Baruch A. 1999. "How Much of the Brain Must Be Dead?" In *The Definition of Death: Contemporary Controversies,* edited by Stuart J. Younger, Robert M. Arnold, and Renie Schapiro, 71–82. Baltimore: The Johns Hopkins University Press.

Buchanan, Allen E. 1988. "Advance Directives and the Personal Identity Problem." *Philosophy and Public Affairs* 17 (4): 277–302.

Buchanan, Allen E., and Dan W. Brock. 1990. *Deciding for Others: The Ethics of Surrogate Decision Making.* Cambridge: Cambridge University Press.

Chiong, Winston. 2005. "Brain Death Without Definitions." *Hastings Center Report* 35 (6): 20–30.

Davis, John K. 2007. "Precedent Autonomy, Advance Directives, and End-of-Life Care." In *The Oxford Handbook of Ethics,* edited by Bonnie Steinbock, 349–74. New York: Oxford University Press.

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

Dresser, Rebecca S., and John A. Robertson. 1989. "Quality of Life and Non-Treatment Decisions for Incompetent Patients: A Critique of the Orthodox Approach." *The Journal of Law, Medicine, & Ethics* 17 (3): 234–44.

Dworkin, Ronald. 1994. *Life's Dominion: An Argument About Abortion, Euthanasia, and Individual Freedom.* New York: Vintage Books.

Feinberg, Joel. 1986. *The Moral Limits of the Criminal Law, v.1: Harm to Others.* New York: Oxford University Press.

Giacino, Joseph T., Joseph J. Fins, Steven Laureys, and Nicholas D. Schiff. 2014. "Disorders of Consciousness After Acquired Brain Injury: The State of the Science." *Nature Reviews Neurology* 10: 99–114.

Gostin, Lawrence O. 2014. "Legal and Ethical Responsibilities Following Brain Death: The McMath and Munoz Cases." *Journal of the American Medical Association* 311 (9): 903–04.

Green, Michael B., and Daniel Wikler. 1980. "Brain Death and Personal Identity." *Philosophy & Public Affairs* 9 (2): 105–33.

Halevy, Amir, and Baruch Brody. 1993. "Brain Death: Reconciling Definition, Criteria, and Tests." *Annals of Internal Medicine* 19: 519–25.

Iltis, Ana S. 2015. "Organ Donation, Brain Death and the Family: Valid Informed Consent." *Journal of Law, Medicine, and Ethics*. 43 (2): 369–82.

Joffe, Ari. 2010. "Are Recent Defences of the Brain Death Concept Adequate?" *Bioethics* 24 (2): 47–53.

Khushf, George. 2010 "A Matter of Respect: A Defense of the Dead Donor Rule and of a 'Whole-Brain' Criterion for Determination of Death." *Journal of Medicine and Philosophy* 35 (3): 330–64.

Kuczewski , Mark G. 1999. "Commentary: Narrative Views of Personal Identity and Substituted Judgment in Surrogate Decision Making." *Journal of Law, Medicine, & Ethics* 27 (1): 32–36.

Locke, John. 1690. *An Essay Concerning Human Understanding*.

MacIntyre, Alasdair. 1984. *After Virtue.* Notre Dame: University of Notre Dame Press.

Mangus, David C., Benjamin S. Wilfond, and Arthur L. Caplan. 2014. "Accepting Brain Death." *New England Journal of Medicine* 370 (10): 891–94.

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

Mill, John S. 1859. "On Liberty." In *Collected Works of John Stuart Mill, Volume XVIII—Essays on Politics and Society Part I (On Liberty),* 1977 edited by John M. Robson. Toronto: University of Toronto Press.

Mill, John S. 1861. "Utilitarianism." In *Collected Works of John Stuart Mill, Volume X—Essays on Ethics, Religion, and Society,* 1985 edited by John M. Robson. Toronto: University of Toronto Press.

Miller, Franklin G., and Robert D. Truog. 2009. "The Incoherence of Determining Death by Neurological Criteria: A Commentary on 'Controversies in the Determination of Death,' A White Paper by the President's Council on Bioethics." *Kennedy Institute of Ethics Journal* 19 (2): 185–93.

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.

Nair-Collins, Michael. 2010. "Death, Brain Death, and the Limits of Science: Why the Whole-Brain Concept of Death Is a Flawed Public Policy." *Journal of Law, Medicine, and Ethics* 38(3): 667–83.

Nair-Collins, Michael. 2013. "Brain Death, Paternalism, and the Language of 'Death'." *Kennedy Institute of Ethics Journal* 23(1): 53–104.

Nair-Collins, Michael. 2015. "Taking Science Seriously in the Debate on Death and Organ Transplantation." *Hastings Center Report* 45: 38–48.

Nair-Collins, Michael, Sydney R. Green, and Angelina R. Sutin. 2015. "Abandoning the Dead Donor Rule? A National Survey of Public Views on Death and Organ Donation," *Journal of Medical Ethics* 41(4): 297–302.

Nozick, Robert. 1974. *Anarchy, State, and Utopia.* Oxford: Basil Blackwell.

Parfit, Derek. 1984. *Reasons and Persons.* Oxford: Oxford University Press.

President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Biobehavioral Research. 1981. *Defining Death: A Report on the Medical, Legal, and Ethical Issues in the Determination of Death.* Washington, DC: U.S. Government Printing Office.

President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Biobehavioral Research. 1983. *Deciding to Forego Life-Sustaining Treatment.* Washington, DC: U.S. Government Printing Office.

President's Council on Bioethics. 2008. *Controversies in the Determination of Death: A White Paper by the President's Council on Bioethics.* Washington, DC: U.S. Department of Health and Human Services.

Rady, Mohamed Y., Joan L. McGregor, and Joseph L. Verheijde. 2012. "Mass Media Campaigns and Organ Donation: Managing Conflicting Messages and Interests." *Medicine, Health Care and Philosophy* 15 (2): 229–41

Regan, Tom. 1983/2004. *The Case for Animal Rights.* Berkeley: University of California Press.

Rhoden, Nancy K. 1990. “The Limits of Legal Objectivity.” *North Carolina Law Review* 68: 845–65.

Robertson, John A. 1999. “The Dead Donor Rule.” *Hastings Center Report* 29 (6): 6–14.

Shah, Seema K, Robert D. Truog, and Franklin G. Miller. 2011. “Death and Legal Fictions.” *Journal of Medical Ethics* 37 (12): 719–22.

Shemie, Sam D., Laura Hornby, Andrew Baker, et al. 2014. “International Guideline Development for the Determination of Death.” *Intensive Care Medicine* 40 (6): 788–97.

Shewmon, D. Alan. 1998. “Chronic ‘Brain-Death’: Meta-Analysis and Conceptual Consequences.” *Neurology* 51: 1538–45.

Shewmon, D. Alan. 1999. “Spinal Shock and ‘Brain Death’: Somatic Pathophysiological Equivalence and Implications for the Integrative-Unity Rationale.” *Spinal Cord* 37: 313–24.

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): 457–78.

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): 256–98.

Truog, Robert D., and J. C. Fackler. 1992. “Rethinking Brain Death.” *Critical Care Medicine* 20: 1705–13.

Uzgalis, William. 2014. “John Locke. Supplement: The Immateriality of the Soul and Personal Identity.” *The Stanford Encyclopedia of Philosophy*. Edited by Edward N. Zalta http://plato.stanford.edu/archives/win2014/entries/locke/.

Veatch, Robert M. 2015. “Killing by Organ Procurement: Brain-Based Death and Legal Fictions.” *Journal of Medicine and Philosophy* 40: 289–311.

Wijdicks, Eelco F. M. 2001. “The Diagnosis of Brain Death.” *New England Journal of Medicine* 344 (16): 1215–21.

Woien, Sandra, Mohamed Y. Rady, Joseph L. Verheijde, and Joan McGregor. 2006. “Organ Procurement Organizations Internet Enrollment for Organ Donation: Abandoning Informed Consent.” *BMC Medical Ethics* 7 (14).

Zamperetti, Nereo, Rinaldo Bellomo, Carlo A. Defanti, and Nicola Latronico. 2004. “Irreversible Apneic Coma 35 Years Later: Towards a More Rigorous Definition of Brain Death?” *Intensive Care Medicine* 30: 1715–22.