

## INTRODUCTION

We all agree that it is generally wrong to kill persons. Although we may differ over exceptions, and over which beings qualify as persons, the judgment that it is wrong, other things equal, intentionally to kill paradigm persons represents common moral ground. At the same time, there is profound disagreement regarding the ethics of killing—or allowing to die—those beings whose moral status is less certain than that of paradigm persons: nonhuman animals, fetuses, infants, the severely retarded and the severely demented, individuals in permanent vegetative states (PVS), and others. Moreover, for those who assume that we human persons have moral status for as long as we exist, related controversies concern the boundaries of our existence: When did we come into existence, and when do we go out of existence or die? Let us refer to all these issues as *the marginal cases*.

One commonly hears that the marginal cases represent areas of moral and/or ontological indeterminacy. According to this position, our clashing beliefs about, say, abortion or meat-eating reflect diverging assumptions about moral status that lie beyond the reach of rational adjudication. Many philosophers, however, are more optimistic. A common strategy for addressing the marginal cases is to appeal to a *theory of moral status*. A relatively novel approach within this broader strategy is to support a theory of moral status with a *theory of personal identity*.

A review of Jeff McMahan, *The Ethics of Killing: Problems at the Margins of Life* (Oxford: Oxford University Press, 2002), hereafter EK, and David Boonin, *A Defense of Abortion* (Cambridge: Cambridge University Press, 2002), hereafter DA. Thanks to Jeff McMahan and the Editors of *Philosophy & Public Affairs* for helpful comments on earlier drafts of this article.

Theories of personal identity generally offer a definition of the term *person* and address the issue of *personal identity*: “In what does a person’s continuing existence over time consist?” Since the publication of Parfit’s *Reasons and Persons*,<sup>1</sup> personal identity theorists have also generally addressed the issue of *what matters in survival*: “What in our continuing existence (e.g., identity itself, psychological continuity) primarily matters from a prudential or self-interested standpoint?” Since the 1990s, the issue of *our essence* has also become prominent: “What are we human persons, most fundamentally: persons, human animals, or something else?”<sup>2</sup>

Can philosophical argumentation help us to settle some or all of the marginal cases? Two recent, outstanding works—Jeff McMahan’s *The Ethics of Killing* and David Boonin’s *A Defense of Abortion*—answer confidently in the affirmative.

McMahan defends and appeals to a theory of personal identity in developing a theory of moral status and the ethics of killing in addressing the full range of marginal cases. His book is an enormously rich contribution to personal identity theory, ethical theory, and applied ethics. Let me briefly describe the five hefty chapters, each of which could be a short book of scholarly significance.

Chapter 1, which will receive extensive consideration in the next section, addresses personal identity theory. McMahan argues that we are essentially neither souls, human animals, nor persons (defined as self-conscious beings), but rather *embodied “minds”* (embodied beings with the capacity for consciousness); our identity is a function of the continuation of this capacity. As for what matters in survival, McMahan contends that the degree to which one should be egoistically concerned about some event in one’s future varies with the degree of psychological unity between oneself now and oneself at the later time. Chapter 2 presents the most probing investigation of the harm of death of which I am aware. Its thesis is that we should understand this harm not in terms of the (prudential) value of one’s life as a whole, but in terms of one’s *time-relative interest* in continuing to live; as explained in the next section, this thesis draws from his view of what matters. Chapter 3 defends a two-tier account of the ethics of killing (as discussed more fully in this article’s

1. Derek Parfit, *Reasons and Persons* (Oxford: Clarendon, 1984).

2. See, e.g., Eric Olson, *The Human Animal* (New York: Oxford University Press, 1997); and Lynne Rudder Baker, *Persons and Bodies* (Cambridge: Cambridge University Press, 2000).

section on abortion). The ethics of killing “minded” nonpersons—in particular sentient animals and severely cognitively impaired humans—is to be governed by the time-relative interest account, which suggests that death harms nonpersons less than persons. Meanwhile, the ethics of killing persons is to be governed by a requirement of equal respect, which overrides consideration of persons’ time-relative interests in continuing to live. The highly original Chapter 4 tackles abortion, employing both the time-relative interest account and the embodied mind account of identity; secondarily, it appeals to the former account in addressing infanticide (see section on abortion). Finally, Chapter 5 draws from various parts of the theoretical apparatus in exploring marginal cases at the other end of our existence: (1) the definition of death (the embodied mind account); (2) euthanasia and assisted suicide (the two-tier account of the ethics of killing); and (3) the authority of advance directives in cases of severe dementia (the embodied mind account and the time-relative interest account). This article’s final two sections address topics (1) and (3).

Focusing primarily on abortion, Boonin develops a theory of moral status that concentrates on the possession, content, and limits of moral rights. His may be the most philosophically thorough book-length treatment of abortion to date and, perhaps, also the best. The introductory Chapter 1 presents the plan for the book, identifies reflective equilibrium as the method to be used in developing a rights theory, and states the project’s central thesis: “the moral case against abortion can be shown to be unsuccessful on terms that critics of abortion can, and already do, accept.”<sup>3</sup> Now the case against abortion can take either rights-based or non-rights-based forms. Chapters 2 through 4, on which I will focus in the section on abortion, respond to rights-based arguments against abortion on the assumption that human infants have a right to life. Chapter 2 energetically rebuts arguments that attempt to establish a fetal right to life from the time of conception. It also contends that we acquire the right to life whenever we are first able to form desires and, on the basis of actual desires, an “ideal dispositional desire” to continue living. Chapter 3 addresses such post-conception criteria as implantation, fetal movement, and organized cortical brain activity, defending the latter—the time at which fetuses become capable of having desires—as the

3. DA, p. 2.

correct criterion for possessing a right to life. In the long and ingenious Chapter 4, Boonin argues that, even if the previous two chapters are mistaken and fetuses have a right to life from conception, the impermissibility of abortion does not follow and cannot be persuasively established. Here he stresses the limits of what rights can morally impose on others. A bit anticlimactically, Chapter 5 handily rebuts a miscellany of non-rights-based arguments against abortion (none of which struck this reader as very powerful).

The present article addresses some of the central themes treated in these books. The most general theme unifying the books is that, contrary to the claim of intractability, philosophical argument can settle at least some of the marginal cases. A more specific point of convergence is that a liberal view on abortion is the only reasonable option. Meanwhile, one of McMahan's most provocative theses is that personal identity theory can illuminate the marginal cases and the connections between them. This article does not comprehensively review either the marginal cases or personal identity theory. Rather, it focuses on the three themes just mentioned with special attention to the unique contributions of the two books. First, in order to provide a context in which the importance of McMahan's work on identity can be appreciated, I describe some challenges facing the dominant psychological approach as well as an alternative view. Ultimately, I reject McMahan's embodied mind account of identity while largely endorsing his account of what matters in survival. Next, I contend that, suitably combined (and in one case modified), McMahan's and Boonin's arguments nearly clinch the case for a liberal view of abortion. Finally, I take up two other marginal cases with special attention to the role of personal identity theory and McMahan's contribution in particular: the definition of death and the authority of advance directives in cases of severe dementia.<sup>4</sup>

4. While these two issues may seem unrelated to identity and abortion, McMahan contends otherwise, and I generally concur (noting some qualifications in later sections). In McMahan's view, death is the mirror image of the beginning of our existence, our beginnings are important to the abortion issue, and both boundaries of human life are determined by our identity and essence. Meanwhile, progressive dementia is the mirror image of early psychological development, both boundaries of our psychological life prove important (whether or not *exclusively* important) to what matters in survival, which in turn illuminates abortion and the authority of advance directives in dementia cases.

PERSONAL IDENTITY THEORY: SOME LEADING  
ISSUES AND OPTIONS

Since the 1980s what we may call *the psychological approach* has dominated the literature of personal identity theory. This approach comprises various theories that assert that our identity—or continuing existence over time—is (at least partly) a function of psychological continuity. Most of these theories understand psychological continuity in terms of particular *experiential connections*, such as having an experience and later remembering it, or forming an intention and later fulfilling it.<sup>5</sup> But some of these theories focus on the continuation of *basic psychological capacities*.<sup>6</sup>

Typically, both psychological theorists (e.g., Parfit, Baker, McMahan) and their critics (e.g., Olson) use the term *person* to refer, roughly, to beings with the capacity for complex forms of consciousness such as self-awareness over time and the ability to plan.<sup>7</sup> In this article, it will be convenient to follow this conventional understanding of the term. (Of course, we can bear in mind the possibility of defining “person” differently—say, as “a being with the *potential* to develop complex forms of consciousness.”)

An emerging theme in recent literature is that psychological theories typically embrace *person essentialism*: the thesis that we, who are now human persons, are *essentially* persons—beings with the capacity for complex forms of consciousness—and therefore cannot exist at any time without being persons at that time. Although some psychological theorists embrace this thesis explicitly,<sup>8</sup> most others strongly imply it, for

5. See, e.g., H. P. Grice, “Personal Identity,” *Mind* 50 (1941): 330–50; John Perry, “Can the Self Divide?” *Journal of Philosophy* 69 (1972): 463–88; Robert Nozick, *Philosophical Explanations* (Cambridge, Mass.: Harvard University Press, 1981), ch. 1; Parfit, *Reasons and Persons*; Harold Noonan, *Personal Identity* (London: Routledge, 1989); and Raymond Martin, *Self-Concern* (Cambridge: Cambridge University Press, 1998).

6. See, e.g., Peter Unger, *Identity, Consciousness, and Value* (New York: Oxford University Press, 1990); and Baker, *Persons and Bodies*.

7. See Parfit, p. 202; Baker, p. 4; McMahan, p. 6; and Olson, p. 24. Why the *capacity* for complex forms of consciousness, rather than its actual attainment at a given time? Because everyone agrees that one can be a person while unconscious so long as one retains the relevant capacity, that is, current ability (not to be confused with the *potential* to develop some ability). The same reasoning applies to the definition of “bare subject” below.

8. See, e.g., Nozick, pp. 78–79; and Baker, pp. 5–6.

example, in asserting that you would go with your mind if it somehow parted ways with your living body.<sup>9</sup>

Let us therefore understand the “standard version” of the psychological approach as asserting the following: Our identity is a function of psychological continuity and we are essentially persons (in the sense defined above).<sup>10</sup> While the standard version has various theoretical motivations, the most powerful may be certain thought experiments, influentially developed by Locke and now the staple of argumentation for this view.<sup>11</sup> Some such thought experiments are realistic, as with “What would happen to you if you entered PVS?” But many are highly fanciful, as with the fission of one person into two, teleportation to Mars, and two people’s mental lives exchanging bodies.

Despite its domination in recent decades, this standard version faces tough challenges.<sup>12</sup> Here I will briefly outline four particularly serious problems, citing McMahan where he advances the same criticism. First, the Newborn Problem: If we are essentially persons—beings with the capacity for complex forms of consciousness—then, since newborn babies lack this capacity, none of us was ever a newborn. This is hard to believe.<sup>13</sup> Second, the standard view has yet to produce a plausible account of the relationship between a person and the human animal associated with her. You are a person. It is natural to think that you are also a human animal, a *Homo sapiens*. But person essentialism entails that you are not, because persons and human animals are not always found together; for example, during gestation and in PVS, the animal exists in the absence of any person. If you are essentially a person, you cannot *be* (identical with) the animal that precedes and may succeed you, as nothing can precede or outlast itself.<sup>14</sup> Thus, on the standard view, you, the person, *overlap* the animal for part of the animal’s existence; or perhaps the animal *constitutes* you, the

9. See, e.g., Sydney Shoemaker, “Personal Identity: A Materialist’s Account,” in *Personal Identity*, ed. Shoemaker and Richard Swinburne (Oxford: Blackwell, 1984), pp. 67–132, at pp. 108–09.

10. Naturally, the essentialist claim can be stated without using the term *person*.

11. John Locke, *Essay Concerning Human Understanding*, 2nd ed. (London: 1694), Bk. II, ch. 27.

12. For an extended critique, see Olson, ch. 3 and 4; see also my “Are We Essentially Persons? Olson, Baker, and a Reply,” *Philosophical Forum* 33 (2002): 101–20. I criticize both the standard view and McMahan’s theory in *Human Identity and Bioethics* (forthcoming: Cambridge University Press), ch. 2.

13. McMahan, p. 44.

14. Olson, p. 94.

person, just as a hunk of bronze constitutes a statue when shaped in the right way. None of these possibilities seems very promising.

The other two problems concern the intuitive case method on which the standard view relies so heavily. Use of this method, especially when cases are described in the first or second person, tends to reify the mind, treating it as a substance that can travel independently of a functioning brain: “What if you woke up in someone else’s body?” Unless substance dualism, which most of these theorists reject, is true, how *could* you wake up in someone else’s body (assuming your brain was not transplanted)? A better approach would be firmly grounded in the natural world. Finally, if we consider a sufficiently broad range of cases, the case method suggests not that we are essentially persons—as the standard view asserts—but that we are essentially *bare subjects* (in McMahan’s terms, *minds*), beings with the capacity for consciousness. After all, one can imagine losing one’s capacity for complex forms of consciousness while still existing as a bare subject; without recourse to science fiction one can imagine, and fear, becoming profoundly demented.<sup>15</sup>

We can begin to appreciate the importance of McMahan’s contribution to this literature by noting how his theory avoids some of these difficulties facing the standard version. His embodied mind account of identity asserts that we are essentially minded beings or minds. But minds, he argues, are—or are caused by—brains (more precisely, those brain parts in which consciousness is realized) functioning in certain ways. And we can plausibly individuate minds not in terms of their mental contents, but by individuating brains. Thus his “mind essentialism” suggests that we are essentially *embodied* minds—a thesis that apparently avoids reifying minds. Accordingly, the “criterion of personal identity is the continued existence and functioning, in nonbranching form, of enough of the same brain to be capable of generating consciousness or mental activity” (EK, p. 68). Further, McMahan holds that the intuitive case method supports the thesis that we are essentially *minds*—not the thesis that we are essentially persons, or beings with experiential connections over time—plausibly implying that we could survive transformation into nonpersonal minds. Moreover, his theory avoids the Newborn Problem: We came into existence when the capacity for consciousness or sentience emerged, perhaps around five months after conception, so we existed at birth.

15. McMahan, p. 66.

This theory also enjoys an advantage common to all psychological views: accommodating the thesis that a conscious being goes wherever its mind goes. Suppose your cerebrum were successfully transplanted into a decerebrate body, maintaining normal mental life, or even just the capacity for consciousness, while your original body, including brainstem, continued to breathe spontaneously and maintain other vital functions except mental life. The powerful “transplant intuition” is that you would go with your cerebrum while the now-decerebrate human animal continued to live. This suggests that you are essentially minded and distinct from the human animal.

But there remains the challenge of plausibly explaining the relationship between you and “your organism.” McMahan contends that you, the embodied mind, are *part of* the human organism with which you are associated (EK, pp. 92–93). Does this not imply that there are now two conscious beings reading this paragraph: the organism (which, having a normal human brain, is surely conscious) and the mind? That seems one too many. But McMahan argues that this implication is palatable in view of certain analogies. When you honk your car’s horn, the horn makes a noise and the car makes a noise (the same one). When a tree’s limb grows, so does the tree. Similarly, we may say that you, the mind, are conscious and so is your organism: The latter is conscious, derivatively, in virtue of having a mind that is conscious.

Although I believe McMahan’s view is the most attractive version of the psychological approach on offer, it has several major problems. First, the theory implies, contrary to what we take to be educated common sense, that we are not animals. Rather, we are parts of human animals and can, in principle, part ways with them (as the transplant intuition suggests). Second, if we embodied minds are not animals, what exactly are we? I do not think McMahan, who claims neutrality on the mind/brain relation (EK, p. 88), can answer cogently.

McMahan cannot say the mind is an irreducibly mental substance, entailing the thesis of substance dualism, because he attempts to refute that thesis (EK, pp. 14–24). Nor can he plausibly assert *property* dualism, which states that while all substances are physical, some properties, the conscious ones, are not. Conceptually, the mind is *that which is* conscious, the subject of consciousness. So it is an entity (substance) that has properties; it itself cannot be reduced to a set of properties. If the mind could be so reduced, then on McMahan’s view, you are merely a set of proper-

ties. This would contradict his explicit assumption that you, the mind, are a substance—and a stranger thesis is hard to imagine! Therefore, the mind on his view must be something physical.

Accordingly, McMahan might claim that the mind is reducible to the brain. But a brain can continue to exist (say, in a corpse) after losing the capacity to support consciousness, whereas, on McMahan's view, we cannot exist without this capacity. Perhaps, then, we are brains that have the capacity to generate consciousness: *functional* brains. But assuming we are substances and the brain is a substance, this thesis would imply that my brain and my functional brain are *two distinct substances*—and that is hard to believe.<sup>16</sup> It seems, therefore, that McMahan has no plausible conception of the kind of thing we essentially are.

Rather than pursue other concerns about McMahan's identity theory here, let me turn to what I take to be the most defensible approach: the biological view. This alternative approach will compete with McMahan's theory and serve as a contrasting point of reference in examining the marginal cases discussed in later sections, helping to clarify what is at stake in embracing one option over another. According to the biological view, which has been developed most comprehensively by Eric Olson,<sup>17</sup> we human persons are essentially living human animals and the criteria for our identity are biological: Human person X at one time and any Y at another time are one and the same being if and only if X's (biological) life is Y's (biological) life. We were all mindless fetuses before we became persons, and we might again exist as nonpersons in severe dementia or even as nonsentient beings in PVS. Thus, like childhood and adulthood, personhood represents a *phase* of our existence rather than an essential property (a property without which we could not exist). A major advantage of the biological view is that it describes the person/human animal relationship in a straightforward, plausible way: You, who are now a person, *are* a human animal.

Interestingly, few if any *actual* cases seriously threaten the biological view. Most of us can accept that we existed mindlessly as fetuses and may again exist mindlessly in PVS. Hypothetical cases provide the strongest challenges to the biological view while supporting psychological theories. Although many of these cases are of dubious intelligibility, the

16. Eric Olson has made essentially the same point (personal correspondence).

17. Olson, *The Human Animal*.

cerebrum-transplant scenario is surely intelligible, so it cannot be dismissed as irrelevant.

I submit that, in interpreting the transplant case, psychological theories conflate *numerical identity* and *patterns of identification*. Because we identify so strongly with our mental lives, considering them so important, we identify with—and would even like to be—whoever has (what was originally) our cerebrum. In another sense of “identity” that presupposes numerical identity—namely, *narrative identity* (as discussed below)—our values and continued capacity for consciousness prove crucial in ordinary cases. But neither our values nor this capacity are criterial for our numerical identity or our essence. We need not assume, with McMahan, that our numerical identity and what matters in survival must be very tightly linked.

In addressing the issue of what matters, McMahan refers to the *degree of psychological unity over time*, which involves three factors: “the proportion of the mental life that is sustained over that period, the richness or density of that mental life, and the degree of internal reference among the various earlier and later mental states [an example of strong internal reference: a detailed memory of an earlier experience]” (EK, p. 75). Dogs, for example, manifest considerably less psychological unity over time than normal humans do. McMahan then advances this thesis about what matters in survival:

The relation that is constitutive of identity—sufficient physical and functional continuity of the areas of the brain in which consciousness is realized in order for those areas to retain the capacity to support consciousness—is both a necessary and a sufficient condition of a *minimal* degree of rational egoistic concern. Beyond that, the *degree* of egoistic concern that it is rational to have about the future may vary with the degree of physical, functional, and organizational continuity in [the relevant brain parts, corresponding to] the degree of psychological unity within the life.<sup>18</sup>

Thus, the extent to which you should be prudentially concerned about some possible event in your future, say, retiring with little savings, varies not only with the positive or negative value of that event, as we normally think, but also with the extent to which the *prudential unity relations* (the relations that ground rational egoistic concern about the future) will hold between yourself now and yourself then. Hence a discount rate

18. EK, p. 79, emphases mine.

for decreased prudential unity. This discounting determines the strength of one's *time-relative interests* regarding possible future events, such as your present interest in retiring with sufficient funds. Now, if identity were the sole prudential unity relation, then one's time-relative interests would simply be one's interests, which are not relativized to some particular time but concern one's life as a whole (EK, p. 80). Where prudential unity relations are very weak, as with a late fetus in relation to itself in the future, one's time-relative interests can diverge sharply from one's interests, and prudential evaluation should rest on the former. (This thesis will carry heavy normative freight in the contexts of abortion and decision making for demented patients.)

Why believe that psychological unity affects the degree of rational egoistic concern about a future event, setting up the contrast between one's interests and one's time-relative interests? This thesis best explains certain prudential judgments: for example, that a barely sentient creature with little psychological unity over time has little stake in continuing to exist (as opposed to having a good quality of life while she does exist); or that the death of a fetus who had attained the capacity for consciousness and had good prospects in life is not prudentially tragic to the degree that the death of a ten-year-old child normally is; or that it *may* be rational for you to choose five more years of normal life over many more years of higher quality life in which you would be psychologically discontinuous from yourself now.<sup>19</sup>

McMahan is on to something. After suggesting a different way of thinking about egoistic concern, I will show how this other approach can be combined with McMahan's.

Earlier I argued that the transplant intuition informed us about our patterns of identification, not about our numerical identity. I think the biological view provides the correct account of *numerical* identity. But, as Marya Schechtman has argued, a distinct sense of "identity" is more closely tied to our patterns of identification, values, and sense of self: narrative identity.<sup>20</sup> Narrative identity involves our self-conceptions, our sense of what is most important to who we are; it becomes threatened in an identity crisis. An appropriate answer to the question "Who am I?" will normally address what is most salient in one's sense of self: perhaps parenthood, a highly selective personal history, one's major occupation,

19. Ibid, pp. 75–78. The latter possibility is illustrated in the case of The Cure, p. 77.

20. Schechtman, *The Constitution of Selves* (Ithaca, N.Y.: Cornell University Press, 1996).

core values, and so forth. My narrative identity is largely determined by my self-told story about my life (though other people figure in the story and shape the way I tell it to myself).

For present purposes, narrative identity is important for two reasons. First, it is a familiar sense of “identity” that people find germane in everyday life. Second, narrative identity is closely connected with egoistic concern. Prospectively considering our own futures (later we will consider our past interest in continuing fetal life), most of us value not mere survival—or numerical identity—but survival with the capacities for consciousness and action. (Those who do value mere survival, say religious thinkers who believe they would be better off alive in PVS than dead, surely value survival with these capacities *considerably more*.) Further, we want our present self-narratives to continue to unfold and include the future experiences and actions, maintaining psychological continuity. And for those of us whose basic needs are met, giving us a modicum of control over our futures, *we want to make something of our lives and become certain sorts of people*. For us, then, much of what matters in survival is its making possible *projects of self-creation*.<sup>21</sup> Self-creation projects flow from narrative identity and, as they do so, continue to write and often edit the narratives from which they flow. In sum, then, much of what matters (to most of us, anyway) is our continuing existence *as persons*—beings with the capacity for complex forms of consciousness—with unfolding self-narratives and, if possible, success in self-creation. But we cannot continue to exist as persons of any kind unless we continue to exist. Therefore narrative identity, as I understand it, presupposes numerical identity.<sup>22</sup>

This account of what matters can be combined with McMahan’s. Addressing the question “How concerned should we be with an event in our future?” McMahan answers with a discount rate based on the degree

21. On the topic of self-creation I have learned much from Jonathan Glover, *I: The Philosophy and Psychology of Personal Identity* (London: Penguin, 1988), part 2.

22. Thus our patterns of identification in the transplant case are misleading. Although one tends to identify with whoever receives one’s cerebrum in the thought-experiment, one would not really be that person and would not have strictly egoistic reasons to care about her welfare. Thus, in this odd hypothetical case, patterns of identification diverge from both forms of identity. Admittedly, this is counterintuitive. One might even wonder why it is worthwhile to distinguish narrative from numerical identity if the former presupposes the latter and can diverge from our intuitive sense of identification. Let me mention four reasons. First, the distinction is edifying because numerical identity, which analytic philosophers emphasize, differs from the more everyday narrative sense of “identity,”

of psychological unity between oneself now and oneself later. Our reflections on narrative identity and self-creation speak, at a very general level, to the *content* of what typically matters to us in survival. The two views are broadly compatible—later I note one major discrepancy—and mutually reinforcing. A high degree of psychological unity over time guarantees the persistence of a person with an ongoing self-narrative (assuming numerical identity is maintained); and to the extent that the agent fulfills earlier formed aspirations, contributing to self-creation, this deepens psychological unity. Near the end of the paper, however, I will argue that accepting the importance of narrative identity enables us to improve upon McMahan's account of what matters.

Thus, while rejecting McMahan's assumption that an account of what matters must closely track *numerical* identity, I largely embrace his account of what matters and integrate it with a view about narrative identity and self-creation. On the view I recommend, which will serve as a point of contrast with McMahan's, we are essentially human animals. Regarding what matters in survival, *even if mere survival counts positively (as some people believe)*—an issue I leave open—*survival as persons with continuing self-narratives counts much more.*<sup>23</sup>

Having sketched several leading issues and options in personal identity theory, and McMahan's important place in this literature, let us now turn to abortion. Can personal identity theory illuminate the morality of abortion? If not, can philosophical argument resolve this issue in another way?

---

which is closely connected with what matters in survival. Second, the two senses of "identity" are relevant to different issues in bioethics, a thesis I develop in *Human Identity and Bioethics*. Third, in the world as we know it (where cerebrum transplants do not occur), patterns of identification track narrative identity very reliably, more reliably than they track numerical identity since some people do not identify with future stages of themselves in which they are permanently comatose, for example. Indeed, the transplant intuition may simply be an overgeneralization from what is ordinarily true: that patterns of identification track narrative identity, which entails—in presupposing—numerical identity. Fourth, as we will see, an understanding of narrative identity and the closely related phenomenon of self-creation will enable us to improve upon McMahan's account of what matters; considerations of numerical identity alone would be insufficient for this purpose.

23. People's intuitions clash over whether numerical identity or mere survival is a sufficient basis for any degree of egoistic concern. For a strong defense of the thesis that we should respect people's differing intuitions, and ask what *does* matter to people rather than what *should* matter to them, see Martin, ch. 1.

## ABORTION

Although the moral debate about abortion may appear to be at an impasse, McMahan and Boonin believe this issue is resolvable. In order to convey the flavor and importance of their contributions to this discussion, I will begin this section by sketching what I regard as the most powerful argument against abortion before evaluating how the authors can and do reply to it. Throughout the section, the challenge of this anti-abortion argument will serve as a test of the strength of the authors' arguments.

Why is it ordinarily wrong to kill paradigm persons like you or me? According to what we may call the *Valuable Futures Argument*, it is because killing us would deprive us of *valuable futures*, which would contain all of our personal projects, enjoyments, and other valued experiences plausibly believed to make human life valuable (in normal circumstances).<sup>24</sup> This account of the wrongness of killing explains why we regard killing as such a terrible crime while accommodating our sense that death ordinarily harms the person who dies. Further, this account plausibly implies that, since killing infants would deprive them of valuable futures, doing so is (normally) wrong. But a human fetus—at least after implantation, when twinning is no longer possible—is an individual that can ordinarily, if permitted to live, *grow into* a person who has the sorts of experiences we value so highly. So the fetus too has a valuable future and therefore abortion is morally comparable to killing paradigm persons, that is, wrong in ordinary circumstances. (Here we need not determine which circumstances constitute exceptions to the moral presumption against abortion.) This is a substantial case against abortion without religious assumptions, equivocation on the moral and descriptive senses of such terms as *human being* or *person*, or dubious appeal to the fetus's potential. Can either McMahan or Boonin defeat this argument?

McMahan's reply has two parts, the second of which is discussed later. Consider fetuses that have not yet attained the capacity for consciousness: call them *early fetuses* and call abortions in their case *early abortions*. Recall that, for McMahan, you and I are essentially minded beings who came into existence when the fetus acquired the capacity for consciousness. Neurological evidence suggests this capacity emerges between the twentieth and twenty-eighth week of gestation—and even if we draw the line conservatively at five months—99 percent of all abortions in the

24. The most influential version of this argument appears in Donald Marquis, "Why Abortion is Immoral," *Journal of Philosophy* 86 (1989): 183–202.

United States occur before this time (EK, p. 268). Now the thesis of mind essentialism implies that early fetuses, lacking minds, *cannot become minded beings*, since it asserts that anything that is ever minded is always minded. Thus, early abortions do not kill beings with significant moral status, making these abortions “relevantly like contraception and wholly unlike the killing of a person” (EK, p. 267). The Valuable Futures Argument, therefore, trips on the mistaken assumption that the early fetus will develop into a minded being. Because it will not, the early fetus does *not* have a valuable future.

If McMahan is right that the early fetus can never become one of us, then his identity theory undermines not just the Valuable Futures Argument but also every other anti-abortion argument that assumes we were once early fetuses, which is to say virtually every significant argument against abortion. This would be an astonishing result that flowed directly from his identity theory. If, however, we are essentially human animals rather than essentially minded beings, then even early fetuses (ordinarily) have valuable futures.

Boonin, assuming for the sake of argument that we were once early fetuses,<sup>25</sup> challenges the Valuable Futures Argument with the following reflections (DA, pp. 56–85). Why does it matter if A deprives B of her future? How, precisely, is that future of value to B? One possible response is that it contains experiences (projects, and so forth) of a sort that B *now values* or (if B lives) *will come to value*. We cannot restrict valuable futures to futures that one now values, one might think, because we need to explain the wrongness of killing *the suicidal teenager* and *the temporarily comatose adult*. For, while neither now values his future, both will (if not killed) come to value it. Yet, Boonin argues, we need not invoke future valuing to explain these cases. The temporarily comatose adult *now desires*<sup>26</sup>—not occurrently, but *dispositionally*—to survive and have those future experiences, just as you (dispositionally) believe that  $2 + 2 = 4$  even when you are not entertaining this belief. The suicidal teenager, meanwhile, requires us to consider *ideal*, not just actual, desires. Ideal desires are those we would have if our actual desires were corrected for cognitive distortions such as faulty reasoning or lack of relevant information. Sup-

25. Boonin, pp. 49–56, 283.

26. Boonin switches from the language of *valuings* to that of *desires* because Marquis, whose version of the valuable futures argument he addresses, treats the terms as interchangeable, and because Boonin finds the language of *desires* more natural for the argument he develops (p. 63).

pose our teenager is so biochemically depressed that, even if prevented from suicide, he will never value his experiences. He still has a right to life, a right not to be killed, yet a view based on actual desires cannot explain why. By contrast, the present-ideal-dispositional-desire version of the valuable futures criterion affirms his right to life: “[His] desires about his personal future would include the desire that it be preserved if his desires were formed in the absence of the chemical imbalance that prevents him from having this desire” (DA, pp. 76–77).

Boonin turns next to fetuses and infants. Since early fetuses, lacking the capacity for consciousness, have no actual desires, they also have no ideal desires; the latter are based on correcting the former for cognitive distortions. But infants have some actual desires, such as desires to enjoy warm sensations and to satisfy their hunger. Moreover, if the infant had greater understanding, he would “surely desire that his future personal life be preserved since he would understand that this is necessary for him to enjoy the experiences that he does already desire to enjoy” (DA, p. 84). So, since the plausibly specified valuable futures criterion applies to infants—and, he later argues (DA, pp. 115–29), to late fetuses—but not to early fetuses, the latter lack a right to life while infants and late fetuses have this right. Thus, Boonin concludes, a defensible account of the wrongness of killing undermines the Valuable Futures Argument in the case of early fetuses.

Although ingenious, Boonin’s reasoning seems vulnerable on several counts. We may reasonably assume that an infant has certain desires connected with her experiential well-being, such as desires for a warm feeling to continue, for Mother to cuddle her, for hunger to end, and for pain to cease. Notice, however, that while satisfying the first two desires requires the infant’s continued existence, satisfying the last two is achievable by death. Presumably, many infantile desires are negative experiential ones, like the last two. Moreover, at least younger infants have no plans that require continued existence. Thus, it is very questionable whether an ideal desire based on the infant’s present actual desires would favor continued existence over nonexistence.

Second, it is doubtful that any desire-based account can adequately capture our understanding of well-being or prudential value.<sup>27</sup> And

27. See, e.g., T. M. Scanlon, “Preference and Urgency,” *Journal of Philosophy* 72 (1975): 655–69; L. W. Sumner, *Welfare, Happiness, and Ethics* (Oxford: Clarendon, 1996), ch. 5, 6; and Martha Nussbaum, *Women and Human Development: The Capabilities Approach* (Cambridge: Cambridge University Press, 2000), ch. 2.

desire-based accounts seem even less likely to account for our sense of the wrongness of killing. McMahan's two-tier account of the ethics of killing may be more promising. In this account, killing *persons*—one tier—is *prima facie* wrong because it violates a requirement of respect, not because it frustrates a desire to live. The requirement of respect, he argues, is motivated by the conviction that, other things (e.g., number of victims, lack of consent) being equal, acts of killing persons are equally wrong.<sup>28</sup> Meanwhile, the morality of killing *minded nonpersons*—the other tier—such as late fetuses and newborns is determined by the time-relative interest account (on which more below).<sup>29</sup>

Third, if Boonin's argument really showed that late fetuses and infants have a right to life equal to yours and mine, it would apparently imply the same for animals who have actual experiential desires: These desires imply an ideal desire to continue existing, grounding an equal right to life.<sup>30</sup> Many will find this result incredible. Boonin might reply that turtles and sparrows, unlike human fetuses and infants, have futures less valuable than ours. But this would suggest that desires play no real role in grounding a right to life, the crucial factor being possession of a future relevantly similar to ours, in which case early fetuses would have rights to life.

If, in view of such difficulties, we reject desire-based accounts of prudential value and the wrongness of killing, we could embrace McMahan's two-tier account of the ethics of killing. But, since infants, lacking self-awareness, are not persons on his view, the killing of infants is to be governed by the time-relative interest account, arguably leaving the morality of infanticide uncomfortably open. Another possibility is to retain the valuable futures criterion but interpret it in terms of *opportunities* as opposed to desires. Although Boonin does not consider this interpretation, he *should* because it represents a stronger version of the Valuable Futures Argument; and, again, we are trying to determine whether the authors can defeat the strongest anti-abortion argument. On this version of the argument, one can be harmed by lost opportunities for valu-

28. By contrast, harm-based accounts of the wrongness of killing imply, for example, that killing old people is less *prima facie* wrong than killing young people, who generally have more to lose from death.

29. McMahan, pp. 240–46.

30. Unlike some theorists (see, e.g., Mary Anne Warren, "Human and Animal Rights Compared," in *Environmental Philosophy*, ed. Robert Elliot and Arran Gare [University Park: Penn State University Press, 1983], pp. 110–25), Boonin uses the term "right" such that if A and B both have rights to X, these rights are equal in strength.

able experiences even if one is not aware of the loss and does not, in any meaningful sense, desire what was lost. Then, since early fetuses will normally lose many valuable opportunities if they die, this approach will confer on them a right to life. The Valuable Futures Argument against abortion remains standing.

Suppose this argument established that fetuses have a right to life. As Boonin emphasizes, it would not follow that abortion is morally impermissible. That conclusion requires another premise: *If fetuses have a right to life, then abortion is impermissible*. Much of Boonin's achievement consists in his resourceful challenges to this apparently innocent assumption. Expanding and improving upon Judith Jarvis Thomson's efforts to the same end, Boonin devotes his long Chapter 4 to "the Good Samaritan Argument," arguing for its validity and tirelessly defending it against objections.<sup>31</sup>

For the sake of argument Boonin grants that the fetus, from the time of conception, has a right to life. But he denies that one's having this right entails that others have a duty to assist one in remaining alive. Consider Thomson's famous case: You wake up to find yourself in a hospital, hooked up to a violinist who can survive his kidney ailment only if you, who alone have the right blood type, remain in the hospital bed for nine months with the violinist plugged into you. He surely has a right to life. And to unplug him in less than nine months would guarantee his death. It nevertheless seems clear that you are not obliged to undergo such a great burden to sustain his life. For while being such a good samaritan would be praiseworthy, going to such lengths to assist another (whom one has not consented to help) is beyond the call of duty. It is not that your rights to liberty and bodily integrity override his right to life; rather, there is no conflict of rights because the violinist's right to life *does not encompass a right to use your kidneys*. The Good Samaritan Argument (GSA) claims that, even if the fetus has a right to life, unwanted pregnancy is relevantly similar to the situation involving you and the violinist: terminating pregnancy through abortion, like unplugging the violinist, is permissible (DA, pp. 135–39).

31. Thomson, "A Defense of Abortion," *Philosophy & Public Affairs* 1 (1971): 47–66. Boonin notes that, among other defenses of this line of argument, E. M. Kamm's treatment in *Creation and Abortion* (New York: Oxford University Press, 1992), which he cites repeatedly, is especially important (DA, p. 134).

Critics of the GSA have two main options: (1) embrace the counterintuitive judgment that it would be impermissible to unplug the violinist while he needs your assistance; or (2) identify a morally relevant disanalogy between pregnancy and the violinist scenario. Boonin persuasively counters the first option. Devoting most of the chapter to the second, he argues patiently and sometimes brilliantly against alleged disanalogies. He is very convincing, for example, that the pregnant woman has not tacitly consented to assume the burdens of continued pregnancy even if she voluntarily had sex and chose not to use birth control. He is also persuasive that the killing/letting die distinction will not by itself support a relevant disanalogy. Still, I think his defense of the GSA *might* break down.

According to *the responsibility objection*, in non-rape cases, that is, where a woman has voluntarily had sex and becomes pregnant, she is responsible (along with the biological father) for the situation in which the fetus needs “life support.” Boonin helpfully distinguishes two senses of responsibility: (1) one is responsible for the fact that A exists; and (2) one is responsible for the fact that, given that A exists (anyway), A needs your assistance. In the violinist case, you are responsible in neither of these senses; similarly with pregnancy due to rape. In non-rape cases, a woman is responsible in sense (1)—she helped cause the fetus’s existence—but not in sense (2), since the fetus could not have existed without needing the woman’s life support.

Boonin then considers cases where one is responsible for A’s existence in sense (1) but in a different way, by extending rather than creating life, and which also vary as to whether in addition sense (2) applies (DA, pp. 172–75). For example, suppose that you, A’s doctor, saved his life seven years ago in the only way possible: by giving him a drug that cured his disease but would predictably cause kidney failure seven years later. Now, as you anticipated, you alone can save him from his new ailment by giving him the use of your kidneys for nine months. Surely you do not act wrongly if you refuse to accept this burden. In this case you are responsible for A’s present existence, since you saved him, but not for his neediness, given that he exists, because he couldn’t have survived to the present day in a non-needy state. Boonin concludes that “where you are responsible in sense (1) and not in sense (2) for the fact that another now stands in need of your assistance . . . the individual in need has not acquired the right to your assistance” (DA, p. 174). If this is correct, the responsibility objection does not work in non-rape cases, and, of course, it does not apply in rape cases. It

would follow that the GSA establishes the permissibility of abortion (assuming the responsibility objection is the GSA's strongest challenge).

But has Boonin defeated the responsibility objection? Stressing the special parent/child relationship, McMahan states a ground for doubt: "[I]t is very hard to believe that it is permissible to kill one's own child in order to avoid the burden of providing the aid that one has caused it to need."<sup>32</sup> Recall that we are assuming, both for the sake of argument and because the Valuable Futures Argument remains standing, that the fetus has a right to life, as an ordinary child does. And, except in cases of hysterectomy or hysterectomy (which present special risks to the pregnant woman), current methods of abortion involve killing the fetus. Finally, while the pregnant woman is not responsible for the fetus's neediness in sense (2)—it is false that the fetus would have existed even if she had not caused its neediness—*she and the biological father have through voluntary action caused the fetus to exist in a state of need, even if they tried everything short of abstinence to avoid this result*. Boonin's senses (1) and (2) of responsibility present a false dichotomy. Here is a sense (3): one is responsible for the fact that A exists in a state in which A needs your assistance.<sup>33</sup> In non-rape cases one is responsible in this sense, which *arguably* grounds a right to one's assistance, at least where A (who, we assume, has a right to life) is one's own child. If so, the GSA is insufficient to justify abortion in non-rape cases.

But we have assumed that the fetus has a right to life. Does it? Or can the Valuable Futures Argument be defeated after all? I believe so. Note that this argument makes two controversial assumptions. First, it assumes that each of us was once an early fetus (in accordance with the biological view of our identity). Equally controversially, the Valuable Futures Argument assumes that identity is the sole prudential unity relation, the sole basis for egoistic concern, suggesting that evaluation of a fetus's future should assume a whole-lifetime perspective. From this perspective, abortion entails (in normal circumstances) a tremendous loss, the loss of a valuable future, with no discounting of this loss. The Valuable Futures Argument takes this prudential claim to justify the moral thesis that abortion is comparable to killing paradigm people. But notice that from a whole-lifetime perspective, the younger one is, other things equal, the more of a valuable future one loses from death. Thus the aforementioned prudential claim implies,

32. McMahan, p. 398. His supporting argument, which explores the Dependent Child case (pp. 392–98), is a response to the GSA in general, not to Boonin's handling of the responsibility objection.

33. Cf. McMahan, p. 366.

strangely, that an early fetus is harmed *more* by death than is a baby, who is harmed more by death than is a five- or twenty-year-old, even if we judge that such differences do not matter to the ethics of killing and assign each individual an equal right to life (EK, pp. 270–71). But just the opposite seems true: Other things equal, the five- or twenty-year-old loses more from death than does a baby, who loses more from death than does an early fetus. And this suggests that the harm of death is a function not only of lost opportunities but also of the way in which one is psychologically “invested” in, or connected with, one’s future. McMahan’s time-relative interest account explains our intuitive judgments about the harm of death, and it can illuminate the morality of abortion.

Let us consider both early fetuses, which have yet to achieve sentience or the capacity for consciousness, and late fetuses, who have this capacity. If the biological view of identity is correct, even the early fetus has a valuable future. But the early fetus has *no* psychological states connecting it with the later psychological being it can become. If McMahan is right that psychological unity is, like identity, a prudential unity relation, then *the complete lack of psychological unity between the early fetus and later minded being requires a very heavy discounting of the value of its future in considering the fetus’s stake in continuing life*. That is, the proper basis for assessing the harm of death to the early fetus is its weak time-relative interest in remaining alive. So, on the view I am recommending, because the fetus is identical to the later personal human animal, it does have *some* (time-relative) interest in remaining alive—this is what is right about the Valuable Futures Argument—but its interest is quite weak, much weaker than yours or mine, due to the absence of psychological unity. Another possibility, from a biological approach, is to drop the claim that bare identity (with psychological life only in the future) is a prudential unity relation, in which case the early fetus would have *no* interest in remaining alive. I do not embrace this second possibility, because I tend to believe that the fetus’s numerical identity with a later being whose life is very valuable confers some interest in continuing to exist. But this is a highly debatable matter where people’s intuitions are likely to clash. Importantly, whichever specification of the biological view is correct, even if the early fetus has an interest in remaining alive, it would be too weak to ground a right to life, so the interests of the pregnant woman could easily justify abortion.

What about late fetuses? According to McMahan, the late fetus is one of us, a minded being. But, while abortion causes it to lose a large

amount of good (a valuable future), its relative psychological isolation from the person it could become—for example, its inability to have intentions about that later time—entails that its time-relative interest in continuing to live is rather weak. Because the late fetus's cognitive capacities are much lower than those of many animals, who lie outside the threshold of respect that grounds (equal) rights to life, the morality of aborting late fetuses concerns both their weak time-relative interests to live and other individuals' interests (EK, pp. 329–30). Since the late fetus has no right to life, its time-relative interest in surviving is easily outweighed by the Thomson-like consideration that we are not morally required to make major sacrifices, especially of bodily integrity, to help others. (We left open the possibility that the responsibility objection defeated the GSA only because we were assuming the fetus had a right to life.) Thus, “even if the Thomson argument does not provide an adequate defense of abortion on its own, it can be combined with the argument I have advanced; and the two arguments together seem wholly decisive” (EK, p. 398).

Considering the common belief that the morality of abortion is intractable, McMahan's optimistic thesis merits careful scrutiny. Interestingly, Boonin cannot accept this thesis. Since he holds that the late fetus has a right to life in virtue of sentience, which gives rise to actual desires and the ideal desire to live, Boonin needs the GSA to justify late abortions. I suggested, however, that his desire-based account of the wrongness of killing is inadequate and that McMahan's two-tier framework is more plausible. On the latter framework, the tier of persons, who enjoy an equal right to life, excludes late fetuses, but it also excludes infants. Some might prefer a modified, one-tier framework that places late fetuses and infants within the morality of respect, giving them a right to life; in that case, the time-relative interest account would determine the harm of death in the case of such minded nonpersons, but not the morality of killing them. Those attracted to this revised framework will need to think carefully about the apparent implication that sentient nonhuman animals have an equal right to life. Those who find this implication too radical may prefer the time-relative interest account as a basis for determining the morality of killing animals, but this approach must deal plausibly and coherently with infanticide.<sup>34</sup>

34. McMahan argues in several ways that his approach accommodates our sense that infanticide is generally worse than abortion. See pp. 338–45.

Thus I am uncertain what conclusions to draw about late abortion. But I think Boonin and McMahan have done much to illuminate the morality of early abortion, which represents nearly all the abortions performed in the United States. Again, I leave for other scholars to consider whether the Thomson–Boonin GSA really does clinch the permissibility of both early and late abortions. If so, then Boonin’s efforts to specify the limits of what a fetal right to life could morally impose on others will have proven very fruitful. I also leave open whether McMahan’s time-relative interest account successfully demonstrates the permissibility of late abortion. But I have suggested that his account justifies early abortions, overcoming the surprisingly enduring Valuable Futures Argument, even though McMahan himself, believing he had justified early abortions with his identity theory, did not apply his account in this way.

Thus, I agree in part with McMahan’s thesis that personal identity theory can illuminate the morality of abortion. A plausible theory of what matters in survival—a part of personal identity theory, broadly construed—proves very important. But I disagree that the study of our identity and essence does much to settle the issue of early abortions; the biological view keeps the issue open (leaving space for the time-relative interest account to weigh in). Then again, the only way to justify my judgment here is to explore the implications and overall cogency of McMahan’s embodied mind account and those of leading competitors. What we cannot afford to do is ignore personal identity theory.

#### THE DEFINITION OF DEATH

We have explored the extent to which the truth about our beginnings bears on the abortion issue. Mirroring the metaphysical issue of our beginnings is that of our “endings.” And just as a theory of our identity and essence speaks to our beginnings, it also speaks to the end of our lives, or, equivalently, the definition of human death.

Since death can be defined at different levels, let us distinguish *conceptualizations* of death, which attempt to capture the general concept (either for humans or generally), and *standards* of death, which provide more specific criteria suitable for legal purposes. According to the traditional *cardiopulmonary* standard, human death is the irreversible cessation of heartbeat and breathing. In recent decades, the development of mechanical respirators and other devices that prolong cardiopulmonary

function, a growing interest in organ transplantation, and two landmark reports<sup>35</sup> propelled the widespread acceptance of the *whole-brain* standard: human death as the irreversible cessation of all functions of the entire brain, including the brainstem (the functioning of which is necessary for *spontaneous* cardiopulmonary function). All U.S. states now legally recognize both cardiopulmonary and whole-brain standards, so death can be declared as soon as it is determined that at least one of the standards has been met.

Interestingly, both traditionalists and supporters of the whole-brain standard commonly cite a *biological* conceptualization: As a biological phenomenon, death must be understood in a way that applies to all organisms, namely, as *the irreversible cessation of integrated functioning of the organism as a whole*.<sup>36</sup> An alternative conceptualization views human death as the *irreversible demise of the person or subject*. This conceptualization supports the *higher-brain* standard: human death as *the irreversible cessation of the capacity for consciousness*. Somewhat radically, this standard deems a PVS patient dead even though PVS is compatible with spontaneous cardiopulmonary function.

We can begin to discern the salience of identity theory, and the importance of McMahan's contribution, to the present issue by examining this reconstruction of a line of reasoning that supports the higher-brain standard:

- (1) For human persons,<sup>37</sup> irreversible loss of the capacity for consciousness is loss of personhood;
- (2) For human persons, loss of personhood is death.

35. Ad Hoc Committee of the Harvard Medical School to Examine the Definition of Death, "A Definition of Irreversible Coma," *Journal of the American Medical Association* 205 (1968): 337–40; U.S. President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, *Defining Death* (Washington, D.C.: Government Printing Office, 1981).

36. See, e.g., Lawrence Becker, "Human Being: The Boundaries of the Concept," *Philosophy & Public Affairs* 4 (1975): 334–59, at p. 353; Charles Culver and Bernard Gert, *Philosophy in Medicine* (New York: Oxford University Press, 1982), p. 181; and President's Commission, pp. 31–43.

37. Human *persons* are specified because this argument concerns beings, like us, who are at some time persons or at least conscious beings. Some who reason along these lines acknowledge that after one of us dies a human organism may continue to live, and that anencephalic infants, which are never conscious, are born alive. See, e.g., Michael Green and Daniel Wikler, "Brain Death and Personal Identity," *Philosophy & Public Affairs* 9 (1980): 105–33, at p. 128.

Ergo (3) for human persons, irreversible loss of the capacity for consciousness is death.<sup>38</sup>

By assuming that we cannot exist at any time without being persons at that time, premise (2) assumes person essentialism. Remarkably, commentators who reason this way typically seem unaware of making a contestable move in personal identity theory.

McMahan's discussion of death is a welcome addition to the literature that defends the higher-brain standard. Understanding that the definition of human death concerns our numerical identity, he employs his embodied mind account to investigate this topic in an original way. For reasons of space, I confine my reply to three points of interest, all of which connect importantly to public policy.

First, McMahan contends that there are *two* correct conceptualizations of human death: one for the human organism and one for the person or, more precisely, the mind (EK, pp. 423–24). That is because we minds are not identical to our organisms; and since both minds and organisms are living substances, both die when they pass out of existence. Thus, while you will die when your capacity for consciousness is irreversibly lost, your organism may continue to live in virtue of cardiopulmonary function until it suffers a terminal cardiac arrest (EK, p. 439). Hence two deaths.

In contrast, if the biological view of our identity is correct, the implications for human death are straightforward. As a human animal (organism), you will die when the animal does. Since we conceptualize human death in organismic terms, the animal dies when it irreversibly ceases to function as an integrated unit, a system whose functioning depends on the interaction of its various subsystems, though what standard best fits this conceptualization remains debatable. Because minds are not living substances distinct from the animals that have minds, the passing of a mind is not literally a form of death. It is rather the loss of a form of functioning that we typically treasure as, *inter alia*, a source of our self-narratives.

38. See, e.g., Edward Bartlett and Stuart Youngner, "Human Death and the Destruction of the Neocortex," in *Death*, ed. Richard Zaner (Dordrecht, Netherlands: Kluwer, 1988), pp. 199–215, at pp. 210–11; Roland Puccetti, "Does Anyone Survive Neocortical Death?" in Zaner, pp. 75–90, at pp. 84–85, 87; H. Tristram Engelhardt, *The Foundations of Bioethics*, 2nd ed. (New York: Oxford University Press, 1996), pp. 241–50; and Ben Rich, "Postmodern Personhood," *Bioethics* 11 (1997): 206–16, at p. 212.

But even if McMahan is wrong that we die when we “lose our minds,” his critique of the whole-brain standard proves important to the debate over standards. McMahan argues that while brain death—the irreversible cessation of functioning of all parts of the brain—allegedly makes sense of the idea that we are organisms, it is neither necessary nor sufficient for the death of the organism (EK, pp. 426–38). Confronted with the fact that most organisms lack brains, brain-death theorists respond that the brain is salient (where a creature has one) because it *regulates the organism's integrated functioning*. But, McMahan replies, brain-dead bodies can, with mechanical support, maintain integrated functioning. For example, pregnant brain-dead women have maintained bodily functions for months, carrying fetuses to term; and a brain-dead boy has over several years grown, overcome infections, and healed wounds. Brain-death theorists may object that in these cases it is mechanical support that regulates integrated functioning. But much of the relevant functioning, such as gaining weight or redistributing blood flow to the fetus, is directed internally. Besides, brain-dead patients are no more dependent on external support than are ordinary fetuses, which depend on maternal bodily functions yet clearly count as alive. So brain death is not sufficient for the irreversible cessation of integrated bodily functioning. Nor is it necessary. If a brain were removed from a living human body and somehow kept alive, the organism could lose integrated functioning and die without the brain's dying.

Considerations such as these suggest that the whole-brain standard is incoherent. McMahan is not alone in advancing such arguments,<sup>39</sup> but his critique is especially cogent. Because it is also detachable from his embodied mind thesis, it proves important for those who accept the biological approach, which leads, I suggest, to the conclusion that the correct standard for human death is an updated version of the traditional criterion: death as *the irreversible cessation of circulatory-respiratory function* (a wording that stresses that integrated functioning occurs throughout the body and cannot be reduced to heart and lung function).<sup>40</sup> If our legal standard for death should rest solely on the nature of death—a big “if”

39. See, e.g., Bartlett and Youngner; Robert Veatch, “The Impending Collapse of the Whole-Brain-Oriented Definition of Death,” *Hastings Center Report* 23 (July–August 1993): 18–24; and D. Alan Shewmon, “The Brain and Somatic Integration: Insights into the Standard Biological Rationale for Equating ‘Brain Death’ with Death,” *Journal of Medicine and Philosophy* 26 (2001): 457–78.

40. See Shewmon, pp. 472–73.

since, arguably, pragmatic considerations also deserve a role—then the legal changes of recent decades have been seriously mistaken.

One pragmatic consideration takes us to a third point of interest. Transplantable organs such as livers, kidneys, and hearts are more viable the “fresher” they are. This fact helped to motivate the whole-brain standard several decades ago; today it motivates the more radical higher-brain standard. But the motivation depends on the dead-donor rule: Vital organs may be harvested only from dead bodies. If vital organs could be taken from living individuals who had provided their informed consent before irreversibly losing consciousness—despite the fact that taking such organs would kill the donors—then our policies for harvesting organs would not depend on legal standards for death. Maybe the dead-donor rule should go.

McMahan thinks so. He also thinks, again, that we should accept the higher-brain standard. Before examining the implications of his embodied mind account (EK, pp. 444–50), let us distinguish (1) what he calls *PVS*, in which cerebral death has destroyed someone’s capacity for consciousness, entailing her death though the organism lives on; and (2) *deep coma*, in which a damaged reticular formation (part of the brainstem that serves as an on/off switch for consciousness without affecting its contents) causes loss of the capacity of consciousness that is presently irreversible but reversible in principle, with sufficient technology. A PVS patient is dead. So, assuming she consented to donation, taking her organs is unproblematic. Although the organism remains alive and will be killed by the procedure, it cannot have interests or rights because it lacks the capacity for consciousness. Now consider someone in deep coma. Because in this state there is physical, functional, and organizational continuity of the brain parts in which consciousness is realized, this person is alive. Nevertheless, taking her organs is permissible, McMahan contends, if she consented, because this would count as justified voluntary euthanasia.<sup>41</sup> Thus, “patients in a deep coma should be treated in much the same way that we ought to treat patients in a PVS” (EK, p. 450).

Although a proponent of the biological view will reject McMahan’s claim that PVS patients are dead, there remains the idea that both they and patients in deep coma are appropriate organ sources. Surely, the strongest arguments for preserving the dead-donor rule appeal to slippery slope concerns: the feared consequences of qualifying laws that

41. He defends the permissibility of voluntary euthanasia later (EK, pp. 455–85).

prohibit the intentional killing of innocent human beings. The continuing debate should take into account McMahan's arguments, and not just those addressed here.

#### ADVANCE DIRECTIVES IN CASES OF SEVERE DEMENTIA

If the biological view of our identity and essence is correct, then our psychological life can end before we die, just as we come into being before having psychological life. The latter, meanwhile, is central (whether or not *exclusively* important) to what matters in survival. As discussed earlier, psychological development is critical in determining the prudential unity relations binding a human organism at two different times, with implications for the morality of abortion. Because progressive dementia roughly mirrors early psychological development, perhaps focusing on prudential unity relations near the end of life will, as McMahan thinks, have implications for the authority of advance directives in cases of severe dementia. Let us see.

Suppose Ann, at age sixty, displays initial signs of progressive dementia. Still competent, she carefully considers her values and options before autonomously completing an advance directive regarding her medical care for any future periods of prolonged incompetence. Ten years later Ann is so severely demented she lacks the psychological capacities that constitute personhood, although she retains the capacity for bare consciousness.

Consider this argument:

- (1) Ann-at-60 is a person.
- (2) Ann-at-70 is not a person.

Ergo (3) Ann-at-70 is numerically distinct from (specifically, a successor to) Ann-at-60.<sup>42</sup>

Its importance lies in the fact that an individual's advance directive is supposed to guide future medical decisions regarding *that* individual, not someone else. Since the argument asserts that Ann-at-70 is someone other

42. See, e.g., Allen Buchanan and Dan Brock, *Deciding for Others* (Cambridge: Cambridge University Press, 1989), pp. 162–66. In discussing cases like Ann's, the authors refer repeatedly to the demented nonperson as a "successor" to the agent of the directive, an agent who "no longer exists."

than Ann-at-60, it suggests a problem—"the someone else problem"—for the directive's authority in guiding decisions about Ann-at-70's care.<sup>43</sup>

But the conclusion in (3) does not follow. Perhaps the two "Ann-stages" are the same individual even if only one is a person. To make sense of the argument, let us bring out the suppressed premise: (1a) Ann-at-60 is essentially a person. This premise would follow from person essentialism. I argued earlier that person essentialism is mistaken.

McMahan's nuanced discussion of advance directives<sup>44</sup> draws significantly from his personal identity theory. Note, first, that his thesis of mind essentialism—that we are essentially beings with the capacity for consciousness—would not support the above reasoning. Ann, on his view, survives the changes of dementia so long as she is sentient. She is therefore (numerically) the same individual throughout. (The biological view reaches the same conclusion, since Ann persists as a single human animal.) Here I will zero in on a single aspect of McMahan's approach.

Let us say that since her teens, Ann-at-60 has strongly valued the intellectual life. She mulls over the possibility of becoming so demented that she can no longer enjoy intellectual pursuits, nor remember much of her past, nor plan more than a minute into the future, yet she is also contented to live rather obliviously in the moment ("pleasantly demented"). Clearly, the prudential unity relations binding her now to herself in this possible future circumstance are quite weak, much weaker than those binding her now to herself when she was in grade school. How prudentially concerned should Ann be about what might happen to her in this deeply demented state?

Recall that McMahan defends a discount rate for decreased prudential unity, as exists in the present case, and that the discounting determines the strength of one's time-relative interests. Nor is this discounting merely permissible; it is required (EK, p. 496). Suppose Ann-at-60 cares greatly about what happens to her in the imagined scenario. Even though she would be very different (qualitatively) in such a state, she strongly desires to be provided life-extending care so long as she seems, at any given time, contented. Moreover, she is willing to commit in advance much of her savings, which would otherwise go to her nephew, to pay

43. See my "Advance Directives, Dementia, and 'the Someone Else Problem,'" *Bioethics* 13 (1999): 373–91.

44. McMahan, pp. 497–503. Its merits include a probing discussion of Ronald Dworkin's influential approach to the authority of advance directives in dementia cases.

for care that her health insurance will not cover. On McMahan's view, apparently, the intensity of her prudential concern is irrational, because her current time-relative interest in what happens in that later possible scenario is relatively weak.

Let me briefly indicate a line of rebuttal, which will qualify my acceptance of McMahan's account of what matters. Recall the concept of narrative identity, which concerns our self-conceptions. Now imagine a situation in which Ann *identifies strongly* with that later self (with whom, again, she is numerically identical): "Sure, I wouldn't be an intellectual anymore, but people change and I've always wanted to be happy." There is a crucial asymmetry between Ann as a fetus or infant and Ann at the cusp of dementia: The latter is an autonomous decision maker, someone with her own values, the author of her self-narrative. It is up to her, I suggest, to determine how much she identifies with a later deeply demented stage of herself and, therefore, up to her to decide how much she cares about what happens to her then. Thus, I think the time-relative interest account should become *optional* in the case of autonomous persons considering their own interests.

This asymmetry between the nonautonomous and autonomous regarding the basis for determining their interests parallels a familiar asymmetry in medical ethics. Medical decisions for children and adults who have never been autonomous are to be made on the basis of their best interests. For those who are autonomous—roughly, competent adults—decisions should be based on their informed consent or refusal. To override their autonomous decisions would be objectionably paternalistic. So is McMahan's requiring the time-relative interest account as the basis for egoistic concern in the case of autonomous choosers.

#### CONCLUSION

In conclusion, McMahan's book demonstrates that we cannot ignore personal identity theory in examining the marginal cases, while his and Boonin's forceful arguments about abortion substantially advance the case for a liberal position.