

## Shadows of the Self

### Reflections on the Authority of Advance Directives

*Japa Pallikkathayil*

People are commonly taken to have the authority to issue advance directives governing their medical care should they cease to be competent. My aim in this paper is to challenge the circumstances and extent to which we should regard individuals as having authority over treatment decisions made for a future state of incompetence. I focus on the loss of competence due to dementia as an important case study. Around fifty million people worldwide are currently living with dementia, and that number is expected to rise to 152 million by 2050.<sup>1</sup> In the course of this discussion, I consider the implications of my account for other circumstances of incompetence.

An adequate account of decision-making for those with dementia should be able to do justice to two related aspects of that condition. First, notice the way in which, for practical purposes, sufferers of dementia both are and are not the same people they were before. Consider reflections of both caregivers and sufferers:

I find myself writing in the past tense, although Mom is still living. I suppose I do so because of Mom's condition now and how different she is. But although very different, mentally and physically, from what I describe above, she is still Mom. This is simply another phase of her life. However, I do miss her and it does tend to feel a little like bereavement.<sup>2</sup>

When shunned for her behaviour, Minnie would say that it was not her but the other woman! (It is true, from her perspective, that it was acted out by the disease not by her *per se*.)<sup>3</sup>

Mother has for some time referred to herself in the third person.<sup>4</sup>

My wife gets frustrated with me...and she is right to be frustrated. She asks me to put a can in the recycling...and I don't do it. She says, "I know this is because of your illness, that this is not you."<sup>5</sup>

Second, notice the way in which the "otherness" of the demented self reflects an unraveling rather than a becoming. The demented self is not a new self but rather the same old self in the process of disintegrating. I am going to begin by exploring this phenomenon. I will suggest that Christine Korsgaard's treatment of personal identity provides a helpful framework for understanding this phenomenon. I will then give an account of decision-making for dementia sufferers that is appropriately responsive to this phenomenon.

## 1. Identity and the Broken Self

It will be helpful to begin by considering a related but distinct challenge to the authority of advance directives in cases of dementia. Some worry that in cases of dementia the psychological continuity necessary for the persistence of a person is disrupted, leaving another person in her place. In these cases, an advance directive issued by a person who has ceased to exist has no authority over the treatment of the new individual. Let us call this the Other Person Problem.<sup>6</sup> As I suggested above, I do not think this is the right way to conceptualize the otherness of the demented self. But examining responses to this worry will provide some important guidance about how to think about personal identity.

I am going to consider two ways of responding to the Other Person Problem. First, one might claim that the kind of psychological continuity necessary for personal identity is not actually disrupted by dementia. Second, one might claim that psychological continuity is not actually necessary for personal identity after all. Let us consider each of these strategies in turn.

Allen Buchanan takes up the first strategy, arguing that we ought to take the degree of psychological continuity necessary for persistence to be so low that we should take cases of progressive dementia to be ones in which the same person persists until there is no person at all.<sup>7</sup> What remains after the capacities for personhood are lost may still be a living being with the capacity for pleasure and pain. In this case, the interests of this non-person may need to be weighed against the surviving interests of the person who has ceased to exist, interests, for example, in what happens to her living remains. But this will not be a case in which respecting an advance directive would involve

subjecting *a person* to another person's treatment decisions. Furthermore, in many cases, respecting an advance directive would not involve violating any obligations to the remaining non-person given the radically truncated nature of that being's interests. The painless termination of life support, for example, would not be in tension with such a being's interest in the nature of her experiential states.<sup>8</sup>

This response to the Other Person Problem invites two questions. First, given that psychological continuity comes in degrees, why take it that there should be a threshold above which identity claims are taken as settled? Second, supposing there should be some threshold, why take it to be as low as in Buchanan's argument? Buchanan's responses to both questions invoke the practical implications of a conception of personal identity. In response to the first question, Buchanan notes epistemological problems about how to distinguish degrees of psychological continuity, and coordination problems that would be created by expecting people to follow rules nuanced enough to be sensitive to degrees of continuity.<sup>9</sup>

In response to the second question about why the threshold should be low, Buchanan argues:

Some of our most important social practices and institutions—those dealing with contracts, promises, civil and criminal liability, and the assignment of moral praise and blame—apparently presuppose a view of personal identity according to which a person can survive quite radical psychological changes and hence a high degree of psychological discontinuity. If this is so, then given the value of these practices and institutions, we would have to have extraordinarily weighty reasons for giving up the view of personal identity upon which they are founded.<sup>10</sup>

Buchanan argues that there are no such weighty reasons. Moreover, although we might be able to reconstruct some of our practices in light of a high threshold, we would face the daunting challenge of responding to a vast new problem of intergenerational justice in light of “the ‘births’ of large numbers of ‘new persons’ who would as it were spring full-blown into the world and who would not, strictly speaking, be the sons, daughters, husbands, wives, or friends of anyone.”<sup>11</sup> And Buchanan suggests that nothing about the view that psychological continuity is necessary for identity forces on us such a radical revision in our thinking.

There are, then, two steps in Buchanan's defense of a low threshold. First, many of our practices presuppose that a person can survive radical

psychological changes. And second, if we reconstructed these practices in light of requiring a high degree of psychological continuity for personal identity, we would face a serious practical problem about what the “new people” these new practices acknowledge are owed.

I suggest, however, that both steps in this reply reveal a methodological tension in Buchanan’s view. In the first step, Buchanan relies on the presuppositions of our practices to support a low threshold. But there are at least some ways in which our practices also seem to presuppose that a person survives even through the late stages of dementia, the stages in which Buchanan holds that the person is gone. For example, we typically take familial obligations to the demented to last even through the late stages of that condition. Thus the presuppositions of at least some of our practices seem to indict requiring even a low degree of psychological continuity.

In the second step of his argument, Buchanan points to a practical problem that revised practices employing a high threshold would face, namely, what to do about all the “new people.” But Buchanan’s own view faces a version of the same problem since it leaves the severely demented untethered in networks of social relationships. Although the severely demented are not persons on Buchanan’s view, we still face questions about who is responsible for their care, questions that have no straightforward answers if we take them to be “new beings.”

The methodological tension in Buchanan’s view thus lies in arguing against a high threshold by pointing to practical presuppositions and problems while allowing some practical revisions and accepting some practical problems for the sake of a low threshold. I suspect this tension arises because Buchanan begins his paper by assuming for the sake of argument that some degree of psychological continuity is necessary for personal identity. This assumption is thus not subjected to the same practical constraints as the competing proposals about how to understand this requirement.

To be clear, I do not intend this criticism of Buchanan’s view to be a criticism of a psychological continuity requirement on personal identity. We will consider that requirement more closely towards the end of this section. Rather, what I take this discussion to show is that reflection on the relationship between a conception of personal identity and our practices must begin at an earlier stage than Buchanan’s. Considering the second of the possible responses to the Other Person Problem will take us a step closer to doing that.

Ronald Dworkin takes up this strategy, denying that psychological continuity is necessary for personal identity. Dworkin takes what he calls the “assimilationist” approach to personal identity. This approach begins by working

out the conception of personal identity that is implicit in our moral and social practices. When we encounter novel cases in which claims about identity are difficult to adjudicate, we should try to extend or adapt our conception of identity “with an eye to the efficiency and fairness of the public system of prudential concern and temporal co-responsibility.”<sup>12</sup> For example, if “travel” by replicating one’s body and destroying the original became commonplace, Dworkin argues that we would take this procedure to preserve personal identity since doing so would be least disruptive to our practices.

Dworkin doubts that any such extension of our ordinary, implicit understanding of personal identity is needed in cases of dementia. He takes the judgment that the life of a single person may have a demented stage to be already firmly rooted in our public understanding of personal identity, an understanding that reflects the following rule of thumb: “if stages of human experience are connected by *either* of the two main continuities—physical or psychological—then we should assume continued personal identity, for the purposes of the twin assumptions of prudential concern and temporal co-responsibility, unless this would be irrational because prudence could take no hold or unjust because co-responsibility is unfair.”<sup>13</sup> Although psychological continuity may be disrupted in cases of dementia, physical continuity is maintained. So, Dworkin sets out to show why neither of the excepting conditions obtains in this case.

Let us begin by considering whether prudential concern is appropriate in such cases. Dworkin seems to take prudential concern to be concern about “the character and value of [one’s] life as a whole.”<sup>14</sup> And he offers as an example of when prudential concern would be “unnatural and unstructured” the following: “someone told he must be prudent for the combination of himself and someone else, whose fate he cannot connect with the value of his own life, has no *way* to be prudent as distinct from altruistic.”<sup>15</sup> In contrast, he argues that the concern that people have for what happens during the demented stages of what he proposes to consider their later lives is easily construed as concern for the character and value of their lives as a whole. This concern has the same character as many other familiar prudential concerns, for example, concerns about the manner of one’s death and concerns about what happens after one’s death, including the success or failure of one’s projects and how one is remembered. Dworkin argues that we could not regard prudential concern for a demented stage of one’s life as irrational without calling these other prudential concerns into question.

Notice two problems with this argument, both stemming from the conception of prudential concern on which Dworkin relies. First, since Dworkin

takes prudential concern to be appropriately directed at events after one's death, noticing that some event is an appropriate object of prudential concern is consistent with denying that this event is an episode in the life of a single, persisting person. One might, for example, have prudential concern in Dworkin's sense for what happens to one's children after one's death. But given where Dworkin places the burden of proof in his argument this might not be very problematic for his purposes—prudential concern gives us no reason for doubting the identity of the demented self and the competent self even if it also does not give us a positive reason for affirming that identity.

The deeper and more pressing problem for Dworkin's argument lies in the passive understanding of prudential concern. Dworkin is, at least here, primarily focused on your concern about what happens to you. But that is in many ways not the primary focus of prudential thought. Of course it matters to you how factors beyond your control shape the character and value of your life as a whole. But much of your prudential thinking is not focused on those factors but instead on how *you* shape the character and value of your life as a whole. You do this by making choices that unfold in time and thereby draw together your earlier and later selves. But this is precisely the thing you can no longer do once dementia looms in the future. You might now make plans for what will happen to your demented self, much like you might make plans for the disposition of your estate after your death. But you cannot make plans for your demented self to carry out. Your thinking will not be integrated with hers in the right way. And that constitutes a profound break with the kind of prudential thinking most of us are engaged in most of the time.

I will return to this point shortly. But first let us consider the second of the two excepting conditions that Dworkin takes it would tell against identity in cases of dementia: the unfairness of "co-responsibility." By this, I take it he means the potential unfairness of holding someone accountable for earlier choices. In particular, we normally take people's current opportunities to depend in certain ways on their past choices. Dworkin suggests the following as a case in which that would be unfair. Suppose that my body divided into two identical bodies with the same memories and mental life. Dworkin suggests that if I knew this would happen I might coherently have prudential concern for both of the lives that would follow my division, and I might even take measures to provide for them both. But requiring this "might well be thought unjust, because I would then have either to cheat my life before I was divided, in order to make available enough for two lives later, or these lives would have to share what would normally do for one."<sup>16</sup> Dworkin asserts that there is no comparable injustice in making one's provisions as a demented

person depend on the prior choices of one's competent self, though he does little to argue for this claim.<sup>17</sup>

Since this aspect of responsibility over time is rather removed from my main focus, let us grant Dworkin for the sake of argument that at least this aspect of co-responsibility is not undermined by dementia. Nevertheless, I suggest that this is not the only important aspect of responsibility through time that dementia might threaten. Consider whether it would be fair to hold your demented self accountable for keeping promises you have made, promises that you can no longer remember or no longer have the capacities needed to fulfill. This seems like a paradigmatic case in which you would be excused from keeping promises. The judgment of excuse rather than irrelevance suggests a single, persisting person. But notice that as soon as you know this excusing condition looms in the future, your ability to make promises is compromised. This is the interpersonal analogue of the problem for prudential reasoning noted above. Dementia threatens your ability to make both self- and other-directed commitments. And this marks a profound break with your future self—that is, a self who you cannot obligate.

When dementia looms in the future, your choices do not have the significance for your future self that they otherwise would. The same is true in the other direction. The choices of your future demented self also do not have the same significance for you that they would otherwise. Contrast two cases. Suppose that, for most of his life, Edward has worked to support equality. But after a series of misfortunes later in life, Edward becomes a vocal racist. Suppose next that Melanie has also worked to support equality for most of her life. But as Melanie slips into dementia, she begins to refer to individuals of certain racial groups pejoratively. She even begins referring to them with slurs. In these cases, Edward responds poorly to the ups and downs of life whereas Melanie's mind is degenerating. For this reason, we might wholeheartedly say, "What a shame Edward became such a racist in later life," while we may hesitate to say the same of Melanie. We might, as in some of the quotations in the introduction, suggest that Melanie's racism is "the disease talking." We must be careful with this point. There is a straightforward sense in which Melanie is doing the talking and foreseeing that one might behave in these ways is one of the horrors of anticipating dementia. But Melanie may not be blameworthy for her racism in the way that Edward is. Perhaps her new views are the product of increasing paranoia and confusion. Or perhaps the views are not so new after all but instead reflect a revealing of her true self brought about by the loss of inhibitions. Especially in this latter case, Melanie's actions may not be wholly excused by her dementia. But even if she has really

always been a racist deep down, the way in which dementia loosens her tongue mitigates her responsibility for her actions. And in this way, Melanie's actions may not reflect poorly on the person she was before in the way that Edward's do.

Let us take stock. Dworkin highlights two important dimensions of our lives in which judgments about identity figure: our prudential reasoning and our responsibility for our choices. But in neither case are the judgments of identity as firm in cases of dementia as Dworkin seems to think. There is a noteworthy way in which your prudential concern for your future demented self has something of the character of estate planning, a kind of planning that is predicated on the assumption of your absence. And judgments about responsibility are likewise a mixed bag. You cannot obligate your future demented self and her actions may not impugn your character.

For all that, I do not think this is good reason for taking the demented self to be a new person. As I noted at the outset, I think that the otherness of the demented self is best conceptualized as a different kind of disruption of identity. Rather than the Other Person Problem, we should consider what I will call the Broken Person Problem: the demented self is a self in the process of falling apart. And I suggest that Christine Korsgaard's work on personal identity can help us appreciate the nature of this problem. Korsgaard's work will enable us to overcome three shortcomings in Buchanan's and Dworkin's accounts. First, I take Korsgaard's view to provide a response to a worry about the practically focused methodology both Buchanan and Dworkin employ, namely that it is overly conservative with respect to our practices. Second, Korsgaard's view will enable us to appreciate the kind of psychological continuity that is relevant to personal identity, an important nuance that Buchanan's view misses. Finally, Korsgaard's view will enable us to begin to understand how and why bodily continuity is important for identity, a question which Dworkin's disjunctive criteria for personal identity leaves unanswered.

Korsgaard is focused on the way in which agency requires us to think of ourselves as unified. Action requires resolving conflicts between one's motivational states.<sup>18</sup> The deliberative standpoint from which you do this requires that you identify yourself as something other than any of these desires. Your actions then take you forward through time:

[T]he choice of any action, no matter how trivial, takes you some way into the future. And to the extent that you regulate your choices by identifying yourself as the one who is implementing something like a particular plan of life, you need to identify with your future in order to be *what you are even*

*now*. When the person is viewed as an agent, no clear content can be given to the idea of a merely present self.<sup>19</sup>

Action thus requires one to think of oneself as unified both synchronically and diachronically.

Many of the practices Buchanan identifies as involving a conception of personal identity are ones that we might coherently consider revising. Dworkin's focus on prudential concern and co-responsibility brings us to practices that are more deeply embedded in our thinking. But these ways of thinking might nonetheless be called into question, as Derek Parfit, for example, does.<sup>20</sup> The kind of identity judgments Korsgaard's practical reflections suggest, however, have something more like an inescapable character. Acting requires thinking of oneself as persisting through time. And it is not clear how we could coherently cease doing that. As Korsgaard puts it: "There is a necessary connection between agency and unity which requires no metaphysical support."<sup>21</sup> In this way, the practically focused methodology Korsgaard employs is not subject to a worry that the practical concerns she relies on reflect an undermotivated conservatism about our practical thinking.

Consider, then, the kind of psychological continuity that is important for identity on this view: "You are not a different person *just* because you are very different. Authorial psychological connectedness is consistent with drastic changes, provided these changes are the result of actions by the person herself or reactions for which she is responsible."<sup>22</sup> With this in mind, return to Buchanan's observation that many of our practices presuppose the possibility of persisting through radical psychological changes. But our practices are sensitive to the etiology of those changes. Consider again our responsibility for our actions. Edward's racism is attributable to him because it reflects the development of his thinking. Melanie's racism may not be attributable to her because it may reflect a malfunction in her thinking. In this way, the agential account of identity Korsgaard proposes is able to account for nuances in our downstream practices, suggesting that these practices are in these respects well-motivated.

But as Dworkin noted, there are other aspects of our practices that treat Melanie as persisting even in the late stages of dementia. Is this an aspect of our practices that can be vindicated on Korsgaard's account? It may initially appear as though not. Melanie is already at a stage in which diachronic unity is a problem for her. Her thinking is not properly integrated with the thinking of her past and future selves. And her progression through dementia may eventually problematize even synchronic unity for her. She may be reduced to

competing impulses with nothing left to adjudicate between them. It may seem, then, that this account pushes us to conclude that Melanie becomes a series of new persons until she ceases to be a person at all.

Korsgaard's take on the role the body plays in identity over time makes this issue even more pressing. She takes a physical criterion of identity to be contingently correct for us because, as things stand, the body is the basic kind of agent. But things might be different if technology were different. Suppose, for example, that technological advances allowed us to replicate our bodies, destroying the original, once every year after age thirty in order to preserve our youth.<sup>23</sup> Korsgaard suggests that this series of bodies would be a single, persisting person because she "would be able to carry out unified plans and projects, and have ongoing relations with other persons."<sup>24</sup> I imagine she would say the same about transportation by replication. But if we are identified with our bodies only so long as they are the basic site of agency, it is especially unclear why we would take identity to hold through the breakdown of agency in cases like Melanie's.

There is something right about this line of thought, but I believe it overlooks an important aspect of the relationship between the body and agency. Agency is an element of a properly functioning mature human body.<sup>25</sup> Our bodies set us the task of coordinating our impulses and sensations in a way that results in action. Even when a body cannot sustain agency, these impulses and sensations—the raw materials of agency—may remain. And when they do, what we have is not a new person or a new being. It is something broken.

Something may be more or less broken. Compare a drinking glass that has crack allowing water to leak out with a glass that has shattered. Likewise, agents may be more or less broken. A person in the early and middle stages of dementia may still be able to act. But diachronic unity is a problem for such a person, and so her ability to complete drawn-out or complex actions may be dwindling. And as I noted above, in the later stages of dementia, even synchronic unity may become a problem.

I do not think it is impossible to imagine psychological breaks in a single body so sharp, complete, and permanent that they are best understood as bringing about the end of one agent and the beginning of another. So, just as we may in principle outlast our bodies, our bodies may in principle outlast us, and not simply in vegetative states but as new agents. But this is not a circumstance we presently encounter in real life. Elements of agency persist even through radical neurological impairment. Those experiencing retrograde amnesia, for example, may retain procedural memory.<sup>26</sup> And even if in the late stages of dementia all that remains is a wash of experiences, these

experiences are the product of the same agential capacities for responding to one's environment that have been present all along. So, in real life we do not encounter new agents inhabiting old bodies. We encounter broken agents. With this description of the phenomenon in hand, we can turn to considering how we ought to respond to it.

## 2. The Authority of Advance Directives

As I noted at the outset, people are commonly taken to have the authority to issue advance directives governing their medical care should they cease to be competent. It will be helpful to begin by considering Dworkin's very influential defense of this position. He begins by asking why we should ever respect the decisions people make when we believe that these decisions are not in their best interests. According to what Dworkin calls the evidentiary view, "we should respect the decisions people make for themselves, even when we regard these decisions as imprudent, because each person generally knows what is in his own best interests better than anyone else."<sup>27</sup> But as Dworkin points out, this view will have trouble explaining why we would defer to people's akratic or altruistic decisions. So, Dworkin instead proposes what he calls the integrity view, according to which the value of autonomy "derives from the capacity it protects: the capacity to express one's own character—values, commitments, convictions, and critical as well as experiential interests—in the life one leads. Recognizing an individual right of autonomy makes self-creation possible."<sup>28</sup>

Dworkin holds that the seriously demented may lack this capacity, and so have no right to autonomy. But the precedent autonomy of such individuals should still be respected: "A competent person making a living will providing for his treatment if he becomes demented is making exactly the kind of judgment that autonomy, on the integrity view, most respects: a judgment about the overall shape of the kind of life he wants to have led."<sup>29</sup> This deference to one's past, competent decisions extends even to situations in which those decisions contradict one's present demented desires.

Dworkin uses a case of temporary derangement to help motivate this view. Suppose a Jehovah's Witness has an advance directive forbidding blood transfusions. If he ends up needing such a transfusion he may countermand his previous decision so long as he remains competent, even if such a reversal would reflect weakness of will. But in a state of derangement, he will not have the authority to change his mind: "Suppose we were confident that the

deranged Witness, were he to receive the transfusion and live, would become competent again and be appalled at having had a treatment he believed worse for him than dying. In those circumstances, I believe, we would violate his autonomy by giving him the transfusion.”<sup>30</sup>

Seana Shiffrin argues, however, that here Dworkin is overlooking an important difference between cases of temporary and permanent derangement:

It is one thing to refuse to listen to those who have temporarily lost their senses...and will have to live with our action. We might think metaphorically about their deranged behavior in terms of an alien self who has temporarily commandeered their bodies. We decline to obey them because a very different personality with a different form of judgment will have to live with the effects of their temporary yet powerful reign. It is both that they will regret the decision *and* that the decision was made under conditions that were not true to themselves. In the case of the permanently demented, however, the “real” self will not return and be forced to live with the consequences of a temporary period of insanity.<sup>31</sup>

As I will go on to argue, I think there is an important difference between many cases of temporary and permanent impairment. But Shiffrin focuses on the wrong difference. This can be brought out by thinking about another case. Suppose you decide to go through a serious surgery because you very much want to live. You have been made aware of what the surgery will involve, including that you will be intubated and will need to remain intubated for some time after you wake up. You also know the medications you will be on when you wake up will confuse your thinking. After the surgery, you do indeed wake up confused and you begin trying to rip out your breathing tube. If we accede to your deranged wishes, your competent self will not return and have to live with the consequences in Shiffrin’s sense since you will be dead. Yet this seems like a case in which it would be entirely reasonable to defer to the decisions of your competent self.

Dworkin’s own objection to Shiffrin’s argument emphasizes a somewhat different problem. He maintains that Shiffrin fails to take seriously the identity of the competent and demented individual.<sup>32</sup> Even in cases of permanent dementia, the earlier competent self has to live with the consequences of what happens to her when she is demented. Those choices affect the character and value of her life as a whole, the very same life as the demented person’s.

I take this objection to Shiffrin to have some force. It is striking that she takes respect for autonomy to provide no reason to implement advance directives in cases of permanent dementia. And it is hard to see how that

could be consistent with taking dementia to be the last stage of the competent person's own life.<sup>33</sup> But it also seems as though Dworkin's view might be charged with a similar mistake. On Dworkin's view, there seems to be no significant difference in the force of an advance directive in cases in which the person is in a permanent vegetative state and in cases in which the person is permanently demented. To be sure, in the latter case, the person may still have interests in the quality of her experiences. But the value of those experiences is taken to have already been accounted for by her previously competent self. Whether some outcome is or is not worth suffering for is a judgment that has already been rendered. Thus, on Dworkin's view, even when considering what is required by beneficence rather than autonomy, the point of view of the demented individual is practically inert. In this way, despite his insistence that the same person persists through dementia, Dworkin treats a person as gone as soon as competence is permanently lost.

Shiffrin and Dworkin both fail to appreciate the sense in which a demented agent is an agent in the process of falling apart. Shiffrin fails to take a competent person's decisions to have any import for how demented individuals should be treated. Dworkin fails to take the exercises of agency of which the demented self is still capable to be worthy of any respect. In the next section, I offer a view that does better on both of these fronts.

### 3. Rethinking Advance Directives

To begin, let us consider a different kind of temporary impairment than the one that Dworkin and Shiffrin have in mind. Suppose Greg is a devout Jehovah's Witness who has had a traumatic brain injury. Greg did not need a transfusion in the course of his treatment for this injury. His treatment has proceeded in accordance with the instructions on the advance directive card he carried with him. As Greg's recovery begins, the extent of his injuries become apparent. Greg struggles with some memory loss and some difficulty forming new memories. He has lost facility with many concepts and must relearn many basic skills. Doctors believe that he may eventually regain enough of his mental faculties to be competent to make his own medical decisions, but he is not presently so and it may be many months, and maybe even a year or two before he improves to that extent. Greg has a generally positive attitude toward his present circumstances. Although he experiences some understandable frustrations during his therapy, he takes pride in his accomplishments and looks forward to conquering further challenges. He also enjoys the time he spends talking with his family and friends.

On the other hand, although Greg remembers some aspects of his religious beliefs, they now strike him as confusing and somewhat frightening. And he expresses a preference not to participate in the religious practices that were once very important to him.

Now suppose that Greg has an accident resulting in massive blood loss and a transfusion is needed to save his life. Should his caregivers still defer to the decisions of his earlier, competent self? This seems like a much harder question to answer than when we imagined a Jehovah's Witness in a fit of derangement, as in the case Dworkin and Shiffrin considered. I suspect that is because Greg is living a life that includes exercises of agency that are worthy of respect. Greg understands the basic goals of therapy and embraces them. Greg participates in his relationships with others. These are not the ravings of an unhinged person. At the same time, the religious convictions that motivated his earlier competent decisions have no place in his present point of view. His previous treatment instructions are thus at odds with the life he is now living.

One might be tempted to explain hesitation about Greg's case in a different way. Although Greg's competence is expected to return, he may very well have different values by then. Greg's case suggests the possibility of two rather different competent points of view. But that possibility was present even in the case of the deranged Witness. Suppose we knew, as is not implausible, that near-death experiences unsettle people's values in some non-trivial number of cases. We might then think it is an open question whether the Witness would indeed be appalled if his competence returned after a transfusion is performed in defiance of his earlier decisions. But I do not think that the openness of that question weakens the case for respecting the Witness's advance directive in that case. The right to live one's life as one sees fit is supposed to protect even choices that may very well be regretted.

Let us return, then, to the thought that the key difference between the original case of the deranged Witness and Greg's case is that the latter but not the former is living a life that includes exercises of agency that are worthy of respect. Greg thus has a point of view that may suggest a course of treatment even if he does not understand his treatment options well enough to be able to decide between them. And this difference between the cases makes no reference to Greg's prospects for recovering competence. I therefore take that feature of Greg's case to be inessential. And I take this to reveal that the important difference is not between cases of temporary and permanent impairment, but between cases in which respect-worthy exercises of agency remain and those in which they do not.

So far, this has just been a diagnosis of why we might hesitate to abide by Greg's earlier competent decisions in his present circumstances. Now I turn to

arguing that this hesitation is well-founded. This will involve gesturing toward a somewhat different reason for deferring to people's choices than Dworkin's.<sup>34</sup> Dworkin emphasizes the value of expressions of character that reflect one's values and commitments, especially those that are oriented toward enacting a view of one's life as a whole. Although I do not doubt that there is value in such expressions, I doubt that this is the value that is operative in our deference to people's choices. Rather, that deference is primarily oriented toward realizing a certain kind of relationship—a relationship of mutual independence. I defer to your choices because I recognize that you and I have the same standing to determine how we each live our lives. And this is true even if we very often fail to live up to our own values and commitments, and rarely act with a view to our lives as a whole.<sup>35</sup> Rather, our claim to stand in relationships of mutual independence is grounded in the much more basic capacity to act in ways that are responsive to the reasons we take to bear on our choices.

I take this view of the grounds of deference to provide an attractive explanation of the expanding scope of the deference we afford to children as they grow into maturity. One ground for deferring to the choices of children even if we believe them to be misguided is educative. But I do not take this consideration to be exhaustive.<sup>36</sup> Instead, notice that it may be appropriate to begin treating children as independent in certain limited domains before it is appropriate to treat them as independent full stop.<sup>37</sup> I suggest doing so reflects acknowledgment that their capacity to reason about the relevant domain is adequate to the task. And this can be so even if they are still not capable of envisioning their lives as whole and aiming to bring about lives with a certain character.

We may have judgments of the same kind about Greg. There are limited domains within which it is appropriate to defer to Greg's choices. We have reason, for example, to defer to his choices about when to receive visitors because there is nothing importantly lacking in his ability to think through this matter for himself. He may sometimes regret spending too much or not enough time with those who want to see him. But he has enough sense of what seeing visitors involves to defer to his judgment about this aspect of his life. I take this line of thought to confirm that there really is a sense in which Greg continues to live his life even in his impaired state. He makes decisions that are worthy of our deference. And the point of view that emerges in the course of living his life is one that speaks in favor of saving his life even if that requires transfusion.

How then are we to adjudicate between this point of view and his earlier one? Simply deferring to the earlier point of view overlooks the way in which Greg is still living a life. But Greg is not in a position to integrate his current

point of view with his previous one. He has ceased to have the religious beliefs that motivated his advance directive, but not because he rejects them—he does not understand them well enough to do that. This might incline us to think that the task of integration is best left up to Greg’s earlier competent self. But that is also not a self that can complete the task of integration. Greg’s earlier self cannot choose to remain steadfast in his religious convictions even through impairment. Remaining steadfast is something that only a later self can accomplish.

Greg never occupies a point of view that is adequate to the task of integrating his earlier and later selves. This is the way in which his condition reflects a breakdown in agency. And this is the reason that independence with respect to this task is inappropriate. Integrating Greg’s earlier and later points of view is something that we have to do for him.

The order in which our lives are lived generally privileges our later selves with respect to the task of integration. To be sure, as we saw above, one begins unifying one’s present and future selves by making plans and having commitments. But as the future unfolds, one is in a position to respond to one’s past plans and commitments in a variety of ways, including by rejecting them. One plausible way of integrating Greg’s earlier and later selves involves taking the viewpoint that emerges from the respect-worthy exercises of agency Greg continues to engage in to constitute a rejection of his previous commitments. This would suggest deferring to his later point of view when making treatment decisions for him.

Notice, however, that this is not the only way the task of integration might be completed. Greg’s loss of religious conviction is bound up with confusion about the meaning of religious claims. For this reason, one might see his reluctance to affirm those claims as merely “the brain damage talking.” On this view, although Greg engages in discrete exercises of agency that are worthy of deference, the point of view that emerges in the course of those exercises of agency is not itself worthy of deference. Hence the task of integration does not after all involve reconciling two points of view. Whether we see Greg’s loss of religious conviction as part of a new point of view that is worthy of deference or as a malfunction in his thinking thus has tremendous significance for the treatment decisions that it would be appropriate to make on his behalf.

This suggests the following conclusions about the authority of advance directives. The authority of advance directives is limited when one continues to live a life that includes respect-worthy exercises of agency. In those cases, there is an important question about whether those exercises of agency add

up to a new point of view that is worthy of deference. There may be cases in which the answer to this question is relatively clear. But I take it that answering this question will very often require an exercise of judgment on the part of decision-makers. And this suggests that in cases like Greg's the real power the past self has over the future self lies in the choice of a surrogate decision-maker who is tasked with interpreting and integrating one's past and present points of view.

I initially presented Greg's case as one of prolonged but temporary impairment. But I took what was relevant to Greg's case to be orthogonal to the temporary aspect of his condition. And for that reason, the above reflections suggest that there may be cases in which deference to the point of view of someone suffering from dementia is appropriate even when it conflicts with an earlier competent decision. This may be especially so for certain kinds of dementia. Alzheimer's disease, for example, is thought to display a pattern of cognitive decline that mirrors cognitive development in children.<sup>38</sup> So, just as increasing domains of deference are appropriate for children as they develop, decreasing domains of deference may be appropriate for Alzheimer's patients as they decline. But that leaves room for the possibility that such a patient may be living a life even when they are no longer competent, a life that may involve a point of view that is at odds with their earlier competent judgments.

Suppose Christina thinks that a life in which she is unable to care for herself in many important respects would not be worth living. For that reason, she makes an advance directive indicating that she should not be given life-saving treatments if she were permanently in such a condition. She is diagnosed with dementia and as her disease progresses she is able to do less and less on her own. For example, she can no longer bathe or cook for herself. But she expresses minimal frustration with respect to these losses. And she enjoys seeing her family, listening to her favorite music, and eating her favorite meals. She also expresses fear at the prospect of dying. The way Christina lives her life in the limited ways in which she is able to suggests a point of view that treats her life as worth living.<sup>39</sup> A decision-maker might reasonably take this as a ground for deferring to Christina's later point of view. Though as we saw above, this involves an important interpretative choice.

Consider also a case with the opposite structure. Suppose John's religious convictions lead him to regard life as sacred in a way that is in tension with withholding life-saving treatments. His advance directive reflects these convictions and requests that such treatments be administered. But during the course of his decline, his wife dies. After that, he ceases to show much interest in the activities of which he is still capable and begins to remark that he would

welcome death. When asked about his religious convictions, he responds, “I don’t know about all that. I just know that I want to die.”<sup>40</sup> John here evinces a point of view that is in tension with his advance directive. And the same considerations that suggested deference to later points of view in the previous cases are applicable here too.

But not all cases will be like this. The activities of a demented individual may not suggest a stable, coherent point of view to which to defer. And here it will be important to recall that there are not two lives under consideration—one pre- and one post-dementia. There is just one life. And where there are no competing viewpoints, there is no task of integration. All that remains for us to do in these cases is to respect the agent’s final point of view. Thus, this view continues to countenance the authority of advance directives in an important class of cases.

#### 4. Conclusion

I have argued that dementia involves a quite literal falling apart. It is no surprise then that so many people fear and dread such a condition. And a natural response to this fear is to try to take control of the future now—to bind everyone else to ensure that you live the kind of life you think you ought to. But I have tried to show that there are limits to the extent that one can do so. Insofar as you go on living, you cannot freeze yourself in place. The point of view you may one day have may be strikingly different from the point of view you have now. And that point of view may emerge in circumstances in which you are not fully in charge of yourself. The most you can do to prepare for such a future is to entrust your care to people who can try to hold together the parts of a self that is breaking.<sup>41</sup>

#### Notes

1. “Dementia,” World Health Organization, September 2, 2021, <https://www.who.int/news-room/fact-sheets/detail/dementia>.
2. Jennifer Davies, “We Don’t Know What is Going Through her Mind,” in Lucy Whitman, ed., *Telling Tales About Dementia: Experiences of Caring* (London: Jessica Kingsley Publishers, 2010), 35–6.
3. U Hla Htay, “We Learn to Enter her World,” in Whitman, *Telling Tales About Dementia*, 60.
4. David Shenk, *The Forgetting: Understanding Alzheimer’s; A Biography of a Disease* (London: HarperCollins, 2002), 191.

5. Shenk, *The Forgetting*, 32.
6. See Rebecca Dresser, "Life, Death, and Incompetent Patients: Conceptual Infirmities and Hidden Values in the Law," *Arizona Law Review* 28 (1986): 373–405, for an early articulation of this problem.
7. Allen Buchanan, "Advance Directives and the Personal Identity Problem," *Philosophy & Public Affairs* 17 (1988): 277–302.
8. Buchanan curiously does not consider an advance directive subjecting the remaining non-person to painful continued existence. In light of this kind of case, it is not clear that there is as little room for conflict between an advance directive and the interests of a remaining non-person as Buchanan supposes.
9. Note, though, that Buchanan's final position allows for the consideration of degrees of continuity below the threshold needed for personal identity. That is, above the threshold identity claims are an all-or-nothing matter, but below the threshold, we can acknowledge the practical import of degrees of continuity while denying that identity is preserved. Given that this is so, I am not sure how deep the epistemological and coordination problems associated with tracking degrees of continuity are supposed to be.
10. Buchanan, "Advance Directives and the Personal Identity Problem," 292.
11. Buchanan, "Advance Directives and the Personal Identity Problem," 293.
12. Ronald Dworkin, "Philosophical Issues Concerning the Rights of Patients Suffering Serious Permanent Dementia," written for the United States Congress Office of Technology Assessment (1986), 103.
13. Dworkin, "Philosophical Issues," 104.
14. Dworkin, "Philosophical Issues," 105.
15. Dworkin, "Philosophical Issues," 104.
16. Dworkin, "Philosophical Issues," 107.
17. Some care is needed to understand this claim. Dworkin's argument is situated in the context of an argument about how the provisions for demented people should depend on the results of a hypothetical insurance scheme in which the preferences of their earlier competent selves figure. So the relationship between the choices of the earlier self and the provisions for the later self is actually rather complicated. But these complications are not important for my purposes.
18. Mere desires may not suffice for action even when they are not in conflict. If so, then there is an even richer role for the deliberative standpoint to play than I have indicated in the main text. Thanks to Tamar Schapiro for bringing this to my attention.
19. Christine M. Korsgaard, "Personal Identity and the Unity of Agency: A Kantian Response to Parfit," in *Creating the Kingdom of Ends* (Cambridge: Cambridge University Press, 1996), 372.
20. See Derek Parfit, *Reasons and Persons* (Oxford: Clarendon Press, 1984), chs. 14 and 15.
21. Korsgaard, "Personal Identity and the Unity of Agency," 374.
22. Korsgaard, "Personal Identity and the Unity of Agency," 379–80.
23. Korsgaard borrows this thought experiment from Thomas Nagel.
24. Korsgaard, "Personal Identity and the Unity of Agency," 374.
25. Although not a point Korsgaard draws on in her discussion of personal identity, I take this point to be congenial to her larger project. See, for example, Christine M. Korsgaard, *Sources of Normativity* (Cambridge: Cambridge University Press, 1996), 149–50.

26. For a helpful discussion of how the fictional depictions of amnesia and related disorders compare with their real-life counterparts, see Sebastian Dieguez and Jean-Marie Annoni, "Stranger than Fiction: Literary and Clinical Amnesia," in J. Bogousslavsky and S. Dieguez, eds., *Literary Medicine: Brain Disease and Doctors in Novels, Theater, and Film* (Basel: Karger, 2013), 137–68.
27. Ronald Dworkin, *Life's Dominion: An Argument About Abortion, Euthanasia, and Individual Freedom* (New York: Knopf, 1993), 223.
28. Dworkin, *Life's Dominion*, 224.
29. Dworkin, *Life's Dominion*, 226.
30. Dworkin, *Life's Dominion*, 227.
31. Seana Shiffrin, "Autonomy, Beneficence, and the Permanently Demented," in Justine Burley, ed., *Dworkin and His Critics: With Replies by Dworkin* (Malden, MA: Blackwell Publishing, 2004), 202.
32. Ronald Dworkin, "Reply to Critics," in Burley, *Dworkin and His Critics*, 367–8.
33. For Shiffrin's discussion of this issue, see "Autonomy, Beneficence, and the Permanently Demented," 207–8.
34. Both Shiffrin and Agnieszka Jaworska also argue for somewhat different conceptions of autonomy in response to Dworkin. Jaworska argues that demented individuals may retain the capacity to value even after they have lost the ability to implement their values. And she argues that this capacity to value is worthy of respect in a way that Dworkin's emphasis on shaping one's life as a whole overlooks. See Agnieszka Jaworska, "Respecting the Margins of Agency: Alzheimer's Patients and the Capacity to Value," *Philosophy & Public Affairs* 28 (1999): 105–38. Shiffrin, on the other hand, focuses on the importance of simply having control over one's experiences even if that does not involve imbuing one's experiences with meaning. This suggests that respect for autonomy may be appropriate even after the capacity to value has been lost. See Shiffrin, "Autonomy, Beneficence, and the Permanently Demented." I suspect the account of autonomy I give in the main text is more akin to Jaworska's than Shiffrin's, though I will not attempt a detailed comparison here. But neither Jaworska nor Shiffrin has a satisfying account of how to reconcile respect for the autonomy of a competent person with respect for the kinds of autonomy still possible for people with dementia. As we have seen, Shiffrin problematically rejects the question by denying that respect for the autonomy of the competent person is still relevant to the treatment of the demented person. Jaworska is less committal on this point but much of her discussion points in the same direction.
35. Shiffrin also notices that Dworkin's integrity account seems to do less well protecting akratic actions than it purports to. She suggests on his behalf that perhaps the reason to defer even to these choices is that "without the freedom to deviate from one's values, the significance of one's accomplishment in creating and sustaining a life guided by certain values is lessened. Perhaps also all self-propelled actions contribute to a self-fashioned life character and life, although the sum of these actions may produce a different product than what one hopes or intends": Shiffrin, "Autonomy, Beneficence, and the Permanently Demented," 201. These may well be ways of enabling the integrity view to get the right answers about deference to akratic or unprincipled actions. But I nonetheless think that it focuses on the wrong reasons for these answers, and I try to motivate an alternative in the text.

36. As Shiffrin helpfully points out, the educative reason for deferring to children's choices will not explain our deference to the choices of children who are terminally ill. But I take the conception of autonomy Shiffrin proposes to make sense of these cases to be too minimal to provide plausible grounds for our deference to people's choices. See Shiffrin, "Autonomy, Beneficence, and the Permanently Demented," 205–6.
37. For a helpful discussion of these "domains of discretion," see Tamar Schapiro, "What Is a Child?" *Thinking: The Journal of Philosophy of Children* 15 (2001): 4–15.
38. Shenk, *The Forgetting*, ch. 8.
39. This case is drawn loosely from Dworkin's presentation of the real-life case of Margo. See Dworkin, *Life's Dominion*, 220–1.
40. This case is drawn loosely from Jaworska's presentation of the real-life case of Mr. O'Conner. See Jaworksa, "Respecting the Margins of Agency," 107.
41. I am indebted to Ryan Davis, Kyla Ebels-Duggan, Tamar Schapiro, and the participants at the Workshop in Ethics and Political Philosophy at Brigham Young University for helpful comments on an earlier draft of this paper. I have also benefitted tremendously from many illuminating discussions with the participants in my graduate seminar on personal identity: Anwar ul Haq, Stephen Mackereth, Caleb Reidy, Rebecca Thomas, and Daniel Webber.

## Bibliography

- Buchanan, Allen. "Advance Directives and the Personal Identity Problem." *Philosophy & Public Affairs* 17 (1988): 277–302.
- Burley, Justine, ed. *Dworkin and His Critics: With Replies by Dworkin*. Malden, MA: Blackwell Publishing, 2004.
- Davies, Jennifer. "We Don't Know What Is Going Through Her Mind." In Lucy Whitman, ed., *Telling Tales About Dementia: Experiences of Caring*, 35–9. London: Jessica Kingsley Publishers, 2010.
- Dieguez, Sebastian and Jean-Marie Annoni. "Stranger than Fiction: Literary and Clinical Amnesia." In J. Bogousslavsky and S. Dieguez, eds., *Literary Medicine: Brain Disease and Doctors in Novels, Theater, and Film*, 137–68. Basel: Karger, 2013.
- Dresser, Rebecca. "Life, Death, and Incompetent Patients: Conceptual Infirmities and Hidden Values in the Law." *Arizona Law Review* 28 (1986): 373–405.
- Dworkin, Ronald. "Philosophical Issues Concerning the Rights of Patients Suffering Serious Permanent Dementia." Written for the United States Congress Office of Technology Assessment (1986). <https://digital.library.unt.edu/ark:/67531/metadc97370/>.
- Dworkin, Ronald. *Life's Dominion: An Argument About Abortion, Euthanasia, and Individual Freedom*. New York: Knopf, 1993.
- Dworkin, Ronald. "Replies to Critics." In Justine Burley, ed., *Dworkin and His Critics: With Replies by Dworkin*, 339–95. Oxford: Blackwell Publishing, 2004.

- Htay, U Hla. "We Learn to Enter Her World." In Lucy Whitman, ed., *Telling Tales About Dementia: Experiences of Caring*, 58–64. London: Jessica Kingsley Publishers, 2010.
- Jaworska, Agnieszka. "Respecting the Margins of Agency: Alzheimer's Patients and the Capacity to Value." *Philosophy & Public Affairs* 28 (1999): 105–38.
- Korsgaard, Christine M. "Personal Identity and the Unity of Agency: A Kantian Response to Parfit." In *Creating the Kingdom of Ends*, 363–98. Cambridge: Cambridge University Press, 1996.
- Korsgaard, Christine M. *The Sources of Normativity*. Cambridge: Cambridge University Press, 1996.
- Parfit, Derek. *Reasons and Persons*. Oxford: Clarendon Press, 1984.
- Schapiro, Tamar. "What Is a Child?" *Thinking: The Journal of Philosophy of Children* 1 (2001): 4–15.
- Shenk, David. *The Forgetting: Understanding Alzheimer's; A Biography of a Disease*. London: HarperCollins, 2002.
- Shiffrin, Seana. "Autonomy, Beneficence, and the Permanently Demented." In Justine Burley, ed., *Dworkin and His Critics: With Replies by Dworkin*, 195–217. Oxford: Blackwell Publishing, 2004.
- Whitman, Lucy, ed. *Telling Tales About Dementia: Experiences of Caring*. London: Jessica Kingsley Publishers, 2010.
- World Health Organization. "Dementia." September 2, 2021. <https://www.who.int/news-room/fact-sheets/detail/dementia>.