UCLA
UCLA Public Law & Legal Theory Series

Title
Autonomy, Beneficence, and the Permanently Demented

Permalink
https://escholarship.org/uc/item/2bd0063d

Author
Shiffrin, Seana

Publication Date
2004

DOI
10.1002/9780470996386.ch11

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California



UCLA | SCHOOL OF LAW

PUBLIC LAW & LEGAL RESEARCH PAPER SERIES

RESEARCH PAPER NO. 15-40

AUTONOMY, BENEFICENCE, AND THE PERMANENTLY DEMENTED

by

SEANA SHIFFRIN

Published in *Ronald Dworkin and His Critics* 195-217 (edited by Justine Burley, Blackwell Publishers, 2004).

Electronic copy available at: https://ssrn.com/abstract=2681781
Electronic copy available at: http://ssrn.com/abstract=2681781

From: DWORKIN AND HIS CRITICS with replies by Dworkin (Blackwell Publishing, 2004)
Edited by Justine Burley

11

Autonomy, Beneficence, and the Permanently Demented

Seana Valentine Shiffrin

Ronald Dworkin's *Life's Dominion*[1] addresses the debates about abortion and euthanasia that have exercised comptemporary American politics. His book and earlier articles[2] urge us to discard portraits of these issues that paint irresoluble conflicts of rights. Instead, he suggests that we should reconceptualize these debates to reflect a recognition that what propels them is an essentially religious disagreement about the intrinsic value of human life.

A human life, Dworkin thinks, has a subjective value that is a function of the interests its bearer takes in that life. Dworkin also contends that we all share the conviction that, independent of its subjective value, human life is intrinsically valuable, that is, it is valuable in itself. It may be valuable even when its bearer does not take an interest in his or her life. Independent of the bearer's attitudes, the life is miraculous and we regard it as in some way "sacred" or "inviolable." Dworkin regards the consensus that life has intrinsic value as a shared, but essentially religious, idea. Thus he sees the social disagreement over our *characterization* of this intrinsic value as a religious dispute.[3] For some the intrinsic value of each life resides in its being the product of divine creation or, alternatively, the product of natural evolutionary processes. For others, it has intrinsic value because it is a complex, vibrant product of a cultural, social, or individual human achievement (pp. 68–101).[4]

Those who regard the natural or divine contribution to life as paramount may regard abortion and euthanasia as manifesting deep disrespect for life's sanctity. These acts end lives that are the outcomes of these processes. By contrast, those who regard the human investment as the critical locus of value believe that abortion and euthanasia may not threaten or disregard the intrinsic value of human life. There is a relatively minor investment represented in the fetus, whereas the comatose, those in constant pain, or those no longer in command of their mental faculties, can make only a negligible contemporary contribution to this process.

In Dworkin's view, reconceptualizing these debates as religious disputes justifies liberal approaches that vest individuals with the ultimate decision-making authority over these matters. Principles of freedom of conscience and religious neutrality make it inappropriate for the state to take positions about why human life has intrinsic value, how much value it has, and how best to respect that value. Likewise, it would be inappropriate for the state to restrict personal liberty to protect that value. Its role must simply be to protect the rights and interests people have in their lives and in living out their convictions. Consequently, Dworkin believes that while it is legitimate for the government to encourage serious thought about the sanctity of life, the government may not force any single view about spirituality on citizens (pp. 150, 157, 168). For instance, Dworkin

contends that the government may encourage reflection about the value of human life by educating its citizens (read women) about abortion (pp. 169–70).[5] But it may not enforce any particular view about how the sanctity of life may best be respected, for example, by banning the option of abortion outright.[6]

Dworkin's philosophical and legal views about the connection between the abortion issue and the sanctity of life are rich and thought-provoking. It is no surprise that they have sparked a significant amount of critical attention and commentary.[7] His views about the connection between respect for autonomy and control over the end of life, while ground-breaking and arguably more provocative, have received less attention.[8] In this chapter I examine these views. I focus on his argument for implementing advance directives that specify what treatment should be given to people who have become permanently demented. In section I, I provide an overview of Dworkin's analysis of the problem of the permanently demented. In section II, I argue that the reasons to respect the autonomy of the fully competent fail to justify extending the range of autonomous control over self-regarding life-and-death decisions enjoyed by competent people to include prospective control over what should happen to themselves if they become permanently demented. I also argue that Dworkin's conception of autonomy is flawed. It leads him to disregard the autonomy that demented people may be capable of exerting. In the third section, I argue that Dworkin's complementary analysis of our duties of beneficence to demented people does not and cannot emerge from a neutral stance toward the sanctity of life. If I am right about the relevant considerations of either autonomy or beneficence, then there may be reason to question seriously whether we should abide by advance directives concerning conditions of dementia. At the very least, I aim to show that Dworkin's argument as it stands is incomplete. More needs to be shown if we are to implement these directives with any degree of moral confidence.

I

Dworkin argues that decisions concerning the prolongation of life and euthanasia properly rest with individuals. The government may not adopt or enforce any particular view as to how to pay tribute to the sanctity of life. It may not dictate that respect for life entails prolonging it as long as possible or, alternatively, that respect for life is best shown by preventing its degeneration into a dependent state of horrific pain, dementia, or permanent vegetation. Instead, people should be encouraged to write their own advance directives that direct what should be done should they become terminally ill, fall into a persistent vegetative state, or become demented. In Dworkin's view, it is appropriate for the state to act on reasons of beneficence and respect for individual autonomy. These considerations dictate respect for individuals' judgments about the value of their lives. Thus these directives should be implemented, whether they direct us to continue or to curtail the lives at stake.

With respect to permanently demented people in particular, Dworkin argues that their rights of autonomy and to beneficence entail that we should respect the decisions of their *past* competent selves about what should be done should they become demented. He claims that we would be mistaken to conceive of demented people as persons with rights of autonomy and interests of their own, distinct and separable from those of their previous competent selves. Rather, in Dworkin's assessment, *qua* demented people (that is, abstracting entirely from their connection to their past competent stages of life), such people have no rights of autonomy at all. They lack the requisite capacities for the attribution of such a right to make any sense.

As to what is in the best interests of permanently demented people, the argument becomes more complicated. Dworkin distinguishes between what he calls one's *experiential* and one's *critical* interests. Something is in one's experiential interest if one enjoys the experience of doing it, like listening to jazz, and runs counter to one's experiential interest if it provides an unpleasant experience such as discomfort, boredom, pain, or nausea. Dworkin believes that although we have a great yen to satisfy our experiential interests, generally, the fulfillment and frustration of these interests, considered by themselves, do not make a life better or worse. We do not, he observes, think a person's life has been a worse life if that person has undergone a great deal of painful dental work (p. 201). Critical interests are interests the satisfaction of which make a life genuinely better and the frustration of which makes a life worse. What lies in one's critical interests are things that will contribute to a good life and that one *should* want, such as interests in accomplishment, in having close relationships with one's family, and in living a morally decent life. Importantly, these interests are the kinds of interests about which people may be mistaken. Such mistakes themselves can make a life go worse.

Dworkin maintains that a demented person *qua* demented person lacks a critical interest in continuing to live. Demented people may have experiential interests in their lives. They may have good or bad experiences within them; they can, for instance, enjoy comfort and reassurance or feel pain and fear (p. 227). They cannot, however, take a view of their lives as a whole, and hence they cannot form opinions about their critical interests. That is, they cannot assess what renders their lives a success or a failure and what, within a life, pays proper heed to the sanctity of life (pp. 201, 230). Moreover, Dworkin thinks that not only are demented people incapable of understanding what is in their critical interests, they cannot do anything to further their critical interests. He claims they are incapable of the acts and attachments that give life value. Value, he notes, "cannot be poured into a life from the outside; it must be generated by the person whose life it is, and this is no longer possible for [the demented person]" (p. 230).

But, for Dworkin, we take an improper view of demented people if we regard them solely *qua* demented people. They must be understood as people whose history is composed mostly of stages of clarity and competence, but who have become demented. The autonomy rights and interests of these whole persons should be understood in the light of their life histories that encompass far more than their period of dementia. These people can be affected by what happens to them in their demented state. A competent, but terminally ill, person may regard the final painful stage of a terminal disease as not worth living and perhaps as a stage that tarnishes the image of a vital life. Likewise, Dworkin imagines that people prior to their decline into dementia may have considered carefully the effect of a demented stage at the end of their lives upon the value of their lives as a whole. They may have decided that "a life ending like that is seriously marred," because it would degrade the previous accomplishments and deliberative direction that previously characterized their lives (p. 231). Consequently, they may exercise their autonomous judgment and sign a living will directing that they be denied medical care and nutrition should they become demented. Others differ and have alternative conceptions of the proper value of life and its sanctity. They may think their appreciation of life is best paid tribute by persevering through life to the end, even with diminished capacities. They may direct that they be permitted to live on during dementia (assuming there are available funds for maintenance and care).[9] Dworkin holds that these sorts of previously issued judgments by those who have become demented reflect their assessments of their critical interests. These judgments must be honored if we aim to respect these people and to protect their interests.[10] In his view, both considerations of autonomy and beneficence counsel us to respect and implement advance directives – whether to terminate or to prolong

the lives of permanently demented people – even when we profoundly disagree with the content of these directives and even when the demented people's expressed contemporary wishes contradict these directives.

These are striking views with heady implications for the lives and deaths of demented people. In what follows, I will not examine or challenge Dworkin's views about the rights of conscious and competent individuals to continue treatment, to refuse treatment, or to end their lives through more active means. I agree that we should take seriously the idea that conscious and competent individuals have a right to die (as well as a right to live, no matter how wretched the existence). I am troubled, however, by approaches that treat the situation of permanently demented people as of a piece with these cases – that is, analyses that extend the right of competent individuals to choose their course while competent to include a right to dictate what should happen to them if they become demented. I do not intend to offer a knock-down argument against such approaches, but to present some concerns about the legitimacy of implementing advance directives about dementia, concerns that emerge from doubts about Dworkin's analysis of autonomy and beneficence.

II

This section focuses on assessing what reasons are provided by respect for autonomy to follow advance directives concerning the cessation or the continuation of the life of a demented person, especially when the demented person now expresses some will to the contrary.[11] My primary aim is to scrutinize Dworkin's analysis of autonomy and to suggest that it is overly narrow. A fuller understanding of autonomy jeopardizes Dworkin's conclusions about the binding force of advance directives in the case of dementia. Contrary to Dworkin's contention, the demented's (*qua* demented) exercise of the will may well engage some autonomy values. At the same time, Dworkin's own analysis of autonomy's value fails to provide the requisite justification for demonstrating the legitimacy of advance directives in cases involving permanent, involuntary, and significantly degenerative personality change.

There are three important questions to press concerning Dworkin's argument for honoring advance directives about dementia:

1 Does respect for autonomy provide us with a reason to implement these advance directives?
2 Does respect for autonomy provide us with *no* reason to respect the contemporary wishes of the demented?
3 Is Dworkin's analysis of our duties of beneficence to permanently demented people persuasive?

I address the first two questions in this section and the third in the following section. I will argue that with respect to the moderately demented and the severely demented, the answer to the first question is no. With respect to the second question, I will argue that the wishes of some permanently demented people concerning the continuation of their lives have some autonomy value worthy of our respect, even if these wishes do not express a reflective unified character. The severely demented who are almost vegetable-like or who have radically self-contradictory wishes may not be able to convey any clear wishes at all, in which case it may be true that they may not exercise autonomy. Still, even in the case of the severely demented, if my first argument is correct,

the advance directive issued by their prior selves may not exert much moral pull *qua* its autonomy value. In such a case, we should not be guided by the advance directive. Rather, we are left with the difficult task of divining what is in such patients' best contemporary interests. I will argue in the following section that Dworkin's rendering of beneficence does not offer plausible guidance for this task. In fact, his account is in tension with the major theme of the book.

Before proceeding to discuss autonomy in greater depth, it is worth becoming clearer about the subjects of our concern. Dworkin limits his discussion to considering those who have become *permanently* demented, that is those people who suffer from an acquired,[12] persistent, and irreversible impairment of intellectual function, characterized by compromised language skills, memory skills, visual and spatial skills, emotional reactions, personality expression, or cognition.[13] Different stages of dementia correspond to different levels of functioning. The mildly demented suffer a moderate level of memory loss, some spatiotemporal disorientation, and have enhanced difficulties handling problems. Still, they exert generally good judgment, may still perform many tasks inside and outside of the home, and generally retain the capacity for independent living with some assistance and prompting.[14] For the moderately demented, independent living may be hazardous and assistance may be required for personal care. Though they have severe memory loss, they retain some highly learned material, and often clearly recognize close friends and family members. They can be taken outside the home, but their social judgment is impaired and their interests are restricted. Many moderately demented people are in denial about their condition, although some indicate awareness that they are behaving abnormally and are undergoing a progressive intellectual decline.[15] Moderately demented patients, even in the moderately severe stage, can engage sporadically in effective communication and may still enjoy living.[16] The severely demented possess only memory fragments, require continuous supervision, cannot maintain personal hygiene, are largely incoherent or mute, and are incapable of making judgments or solving problems.

Unfortunately, it is unclear to which of these stages Dworkin intends his conclusions to apply. Dworkin declares early in his discussion that his attention is restricted to those in the late stage of dementia who have lost all memory and sense of self-continuity; cannot attend to their own needs and functions; are incapable of sustaining projects, plans, or desires of even a very simple structure; but who may harbor a desire to live (pp. 218–19). Later, however, he elaborates, arguing that:

> When a mildly demented person's choices are reasonably stable, reasonably continuous with the general character of his prior life and inconsistent only to the rough degree that the choices of fully competent people are, he can be seen as still in charge of his life and he has a right of autonomy for that reason . . . But if his choices and demands, no matter how firmly expressed, systematically or randomly contradict one another, reflecting no coherent sense of self and no discernable even short-term aims, then he has presumably lost the capacity that it is the point of autonomy to protect. (p. 225)

I will discuss Dworkin's views about autonomy later. At present, I am interested in this passage for what it reveals and fails to reveal about the subjects of Dworkin's conclusions. The passage suggests at one point that the test to determine whether advance directives should trump the patient's contemporary wishes turns upon whether the demented's judgment is discontinuous with the judgment of the patient's past selves. Only a few sentences later, the test becomes whether the contemporary judgments are themselves radically self-contradictory, a rather different criterion. Dworkin's articulated standards leave it unclear whether he intends to extend his conclusions to those moderately demented people who may express occasional firm preferences that are

not themselves radically self-contradictory, but are not continuous with their past character, are less consistent than the choices of the fully competent, and do not flow from a clearly defined, reflective character. Further, despite Dworkin's pronouncement that he will limit himself to the severely demented, the cases he discusses involve people who appear moderately demented. One patient, Margo, maintains persistent, though simple, desires and projects. She pretends to read, attends art classes, and loves sandwiches. Her medical student attendant declared that she was "undeniably one of the happiest people I have ever known" (p. 221).[17] Such examples appear to indicate a willingness on Dworkin's part to extend his conclusions to a broader category than those in late-stage dementia. In any case, whether this interpretation correctly gauges Dworkin's intention, as I will argue, his accounts of autonomy and beneficence seem subject to such extension.

To simplify syntax, I will refer to the subjects of my concern in the following way. I will refer to a person when in a stage of competence as P_1. I will refer to the person occupying the same body as P_1, but when in a stage of dementia, as P_2. I will assume, with Dworkin, that P_1 and P_2 are not different people but rather pick out different selves or different stages within the same person's life.[18]

In Dworkin's view, if prior to the onset of dementia I sign a living will that declares that my life should be terminated if I mentally deteriorate to a particular stage, then no matter how much I enjoy my life, and insist upon its continuation at the time, others will show fundamental respect for me by "allowing" my death (or by killing me, depending on the contents of the directive). In this example, P_1 issues the directive that P_2 is to be killed and P_2 pleads against this directive's implementation. Alternatively, if, when competent, I direct that I must continue living during my demented stage no matter what – even if I am in utter pain and begging for relief – then others will show me respect by refusing to allow my death; they would disrespect me by providing lethal relief.

Dworkin's justification for these conclusions runs as follows. When deliberating about what action to take concerning the continuation of the lives of demented people, we must consider their rights of autonomy as well as their rights to our beneficence. Considerations about respect for autonomy, he thinks, come down in favor of adhering to advance directives, as an analysis of two main justifications for respecting autonomy reveals. The first, the *evidentiary account*, stresses that we should respect the decisions of others and refrain from interfering with others' lives without invitation, because individuals generally make the best judgments about what is in their own best self-regarding interests (pp. 222–3). Respecting other people's autonomy is, as a rule, the best way to promote their welfare, and not simply because among their various desires is the desire for control. But these grounds give us little reason to respect the expressed wishes of demented people, because it is less plausible to think that they generally make the best judgments about what is in their best interests.

Dworkin finds the evidentiary account inadequate, however, or at least incomplete, for two reasons (pp. 223–4). First, the evidentiary view cannot explain why respect for autonomy would direct us, as he thinks it does, to permit people to behave in ways that they themselves acknowledge as contrary to their best interests. Second, it neglects those cases in which one justifiably and understandably uses one's autonomy to further the interests of others, even at one's own overall expense. These considerations lead Dworkin to posit that some other justificatory ground must support respect for autonomy. He contends that an appeal to "integrity" may supply this missing ground. On this account, autonomy commands respect because it protects the capacity to express one's character through one's actions, to create and fashion a distinctive personality that propels

one's life, and thereby to structure a life around one's own values. Dworkin argues that this view does not assume that every action a person performs will reflect consistency or structure, but that autonomy "encourages and protects people's general capacity to lead their lives out of a distinctive sense of their own character, a sense of what is important to and for them" (p. 224).[19]

Appeal to the *integrity account* clearly addresses the second shortcoming of the evidentiary account. I may wish to lead a life structured not only around furthering my own self-regarding interests, but also devoted to fulfilling interests of others. I may also reasonably commit myself to other forms of value, even if such pursuits are ultimately to my detriment or only indirectly serve my self-regarding interests. Less clear, though, is how an appeal to the integrity-based view could justify protection for acknowledged akratic behavior. One might, of course, be in some peculiar and precious way dedicated to living a rakish, imprudent life driven by a misguided but beguiling romanticism. But for those for whom akratic action is entirely out of character or unintentional, it is unclear how the integrity-based account lends support for respecting those actions. Dworkin might argue that the freedom to perform actions that are out of character is necessary for developing the ability to perform actions within character, but then, if this instrumental move works to support the integrity account, why couldn't similar arguments be made to support the evidentiary account? Perhaps the appropriate argument here 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 character and life, although the sum of these actions may produce a different product than what one hopes or intends to create.

In any case, Dworkin argues that neither account would compel us to respect a permanently demented patient's requests, expressed preferences, or decisions. Demented people lack the capacities of character structuring presupposed by the integrity account and the capacities for savvy deliberation presupposed by the evidentiary account. Both accounts point us toward following the directives issued by those people while they were in full command of their capacities. Suppose, however, that although P_1 directed that she be killed were she to become demented, that P_2 is happy and expresses desires to continue living while exhibiting terror at the prospect of her death. Shouldn't her contemporary wishes supersede her past decision? Doesn't respect for autonomy dictate that we should refrain from acting contrary to P_2's expressed wishes?

Quite the contrary, says Dworkin. Keeping the demented person alive where her precedent self has asked that she be killed in such a state "violates rather than respects her autonomy" (p. 229). He justifies this fairly chilling conclusion by asking us to reflect upon the case of a temporarily deranged Jehovah's Witness. He regards the case as analogous to the situation of the demented. Dworkin persuasively observes that it would be wrong to provide a transfusion to a demanding, yet temporarily deranged, Jehovah's Witness. It is not that agents cannot reverse or retract their previous decisions. Agents may change their minds, but only under the proper conditions. The deranged Jehovah's Witness, Dworkin observes, lacks the necessary capacity for a fresh exercise of autonomy. Her demands, therefore, cannot overturn her previous resolve. Her former decision remains in force because no new decision by a person capable of autonomous deliberation has annulled it. On the integrity-based value of autonomy, we respect the autonomy of the Jehovah's Witness by respecting her decisions about the character of her life made when she was in full control of her deliberative capacities. Likewise, Dworkin believes that the integrity-based view supports the idea that we should not indulge the deranged since they lack the necessary capacities to make decisions that contribute to the aim of building and expressing a coherent and distinctive character. Given his analysis of the Jehovah's Witness case, Dworkin thinks that it follows

that we respect the autonomy of the demented person best by allowing her to die as per her previous directive, despite any contemporary protestations or signs of happiness.

I am skeptical about the analogy between the Jehovah's Witness case and the demented patient. My unease about this case can be traced to a suspicion that Dworkin neglects an important aspect of the grounds for valuing autonomy. This aspect may be uncovered by examining the force of the Jehovah's Witness example and identifying some disanalogies with the case of the demented.

Although Dworkin does not emphasize this feature, it seems important to his example that the derangement of the Jehovah's Witness is temporary and further, that if the life is saved, the Jehovah's Witness with full capacities will reemerge. She will regard our decision to deliver a transfusion with horror and revulsion. Dworkin rightly insists that the prospect of the Witness's regret cannot be what motivates us to respect the Jehovah's Witness's prior decisions against treatment. We should accede to a lucid Jehovah's Witness's request for treatment even if we know that person will later regret the decision. Anticipated regret, then, cannot be the salient feature restraining us from providing treatment (p. 228). There is an important asymmetry, however, between the Jehovah's Witness case and the cases of dementia. In the Jehovah's Witness case, the Witness's capacity for full-blown autonomy (a capacity to render decisions which is characterized by a stable, developed character and a fully rational intellect capable of assessing options and of engaging in self-reflection) has not been permanently lost but rather lies in abeyance. We may regret that we cannot accede to the deranged Witness's request for treatment, but we refrain from doing so because we have evidence of the Jehovah's Witness's true autonomous decision. We know that she will reemerge with this full capacity and our action will have frustrated its aim. In the case of permanent dementia, though, the capacity for full-blown autonomy is lost and will not be recovered.

This places the decisions and requests of the permanently demented person in a different light. It is one thing to refuse to listen to those who have temporarily lost their senses and act extremely out of character, but who will, or can, in time, return to 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. This makes a difference. In the Jehovah's Witness case, P_1 will return, is the more capable decision maker, and will have to live with the decisions for a long period of time. Where more permanent dementia sets in, the two factors come apart. P_1 may be the more capable decision maker, but P_2 will have to live with the consequences of the decisions, not P_1. Given that the fully capacitated self has been permanently effaced, our reason for overriding the expressed will of the deranged self diminishes. Since she is now operating with all of her possible decision-making capacities characteristically in force, limited though this set may be, there may be a reason to treat her expression of her will as a retraction of her previous decision.

Dworkin contends that there is nothing about P_2's judgment that commands any respect. Her decisions do not promise to teach us much about what is in her best interests. Further, she cannot reflect upon her critical interests at all, and hence she is not capable of making the kinds of decisions that the integrity-based view recognizes as exercises of autonomy. Keeping in mind the Jehovah's Witness case, however, it may be apparent that Dworkin glosses over an important dimension of autonomy's importance: the basic value of being in control of one's experience and in not having experiences forced or imposed upon one when one's will is to the contrary. The

value of basic control over, and self-direction of, one's experiences, even momentary and slight ones, is an important ground for respect for autonomy distinct from the two foundations Dworkin discusses. Certainly, the evidentiary and integrity accounts identify strong central reasons for respecting autonomy, but the value of self-determination is not fully captured by the values pegged by these accounts. Understanding that one's choices generally best promote one's welfare or express one's character may isolate what is noble and most prized about the achievement of leading an autonomous life. Yet these values do not exhaust why we think people have some strong rights of control over their lives. There are many cases in which people do not choose what is in their best interests, and in which their choices will not contribute substantially to the conscious creation or expression of a distinctive character structured around a set of values, yet in which there is still *some* (much!) reason to respect those people's choices and to refrain from forcing a contrary decision upon them. We value and respect autonomy in part because we recognize that one's life is solely one's own and that one must bear and endure singly one's own conscious experience. Given this, the right and ability to select for oneself the contents of this experience is fitting and appropriate. To a certain extent, one is wronged if one's self-regarding experience is dictated or imposed by another. Respecting other people's decisions is not simply a way to promote their welfare or to facilitate the valuable process of their creating and expressing a distinctive character. In a more basic way, I believe it serves as an acknowledgement of the moral importance of the uniqueness and separateness of persons and the deep, irreducible fact that one's life is the only life one has. These facts do not just amount to the fact that one's *character* is distinct from that of others. There is more to one's mental life than is properly regarded as part of one's character (take, for instance, pain). Moreover, even when misdirected or confused, a life determined from within (driven by that person's will) is more valuable and distinctively human than one manipulated and steered by others.

Notably, the dispute that I am pursuing here with Dworkin is not the familiar dispute about whether autonomy has intrinsic or extrinsic value.[20] My challenge is not to the type of evaluation Dworkin assigns to autonomy (its intrinsic value) but to his identification of what it is in virtue of which autonomy has value. Dworkin has identified its intrinsically valuable features too narrowly. It is not just in virtue of autonomy's expressive nature that it is intrinsically valuable but also in virtue of its being a manifestation of individual control.

I do not dispute the integrity-based account or its relative importance. I simply wish to stress that this broader dimension of control merits recognition and exerts some moral force that makes a difference to our understanding of the conditions under which one's autonomous control extends into the future. It is difficult to elaborate fluently upon what exactly this other dimension of control consists in or why it is important. I can only gesture at what I take to be basic facts carrying moral significance: one's life and one's conscious experience are one's own, distinct from others; the simple exercise of control over one's experience reinforces the special relation one has to one's experience and has some value independent of what it accomplishes. Rather than repeating these invocations of the ineffable significance of separateness and self-determination, I will attempt to render the thought more appealing by way of two routes: first, by investigating the motivations behind the integrity account and second, by considering some overlooked aspects of the case of children, to whom demented patients are often compared.

Dworkin's integrity-based account of autonomy locates the justificatory ground for respect for one's chosen action in the capacity to have a distinctive character and to act from it. Only where this capacity exists over an extended and continuous period of time does one have the requisite abilities to participate under the banner of integrity, and only then can one claim rights to auton-

omy. But suppose we probe deeper and ask why it is important to be able to *act* out of a distinctive personality once it is formed. Why exactly is the integrity-based account compelling? The suggestion that seems correct to me, at least in part, is that autonomous action attempts to synchronize one's inner and outer life. Autonomous action permits one to try to fashion a life or even just momentary happenings and feelings that accord with one's subjective aims, desires, and visions of oneself and one's experience. Autonomy, then, facilitates a sort of unity between one's will and self-consciousness and one's endured experience, and life, more broadly conceived. When autonomy's exercise proceeds at its best, the unity one achieves will result in a distinctive, self-fashioned personality that coordinates with the life lived. This is a valuable and important accomplishment in itself, but it is hard to believe that it is a prerequisite for autonomy's having any intrinsic value, or that its value is substantially different in kind from the value of exercises of self-control that are not closely connected to the expression or creation of a distinctive character. If I am right about its elaboration, then the deeper layer of rationalization supporting the integrity account would point us beyond this account and toward the broader, "control" account of which the integrity account is a prominent, but only partial, component.

This further elaboration of the justification of the integrity account raises some difficulties, though, for the defense of advance directives about dementia. On the one hand, if autonomy's value derives from the importance of unifying one's experience with one's projected vision of oneself and one's personality, then why would we think that autonomy rights should extend control to include stages of one's life in which one's personality has abruptly ended? For in such a case, the unity cannot be achieved and the personality expression is mostly illusory and insufficiently genuine.

On the other hand, if my suggestion is right about the underlying motivations behind the appeal to integrity, then the value of autonomy may be reduced in cases in which one's capacity to form and express a personality has diminished, but it will not be extinguished. Even if people lack a coherent character structured around an ordered scheme of values, they may well have desires, aims, and preferences that could be synchronized with the content of those of their experiences that they and others can control.

There are two components of the value of autonomy on this account. First, there is the value of forming one's subjective desires and one's will, and their being unified with one's lived experience. Second, there is the value of effecting this unity through one's own action. Those with normal capacities may be capable of both and may actively contribute to the achievement of a rich and complex personal unity. This activity is usually part of the process of forging a distinctive character. Demented people may lack the abilities to effect this unity (and hence some forms of intervention may be justified to help them bring about what it is they will or want). Further, that which can be unified may be more paltry and dilute. Still, to the extent that they can express their will, some of what is of value about autonomy is at stake here. They surely can attempt to control the nature of their lived experience and to experience (although perhaps not effectuate it on their own) a unity between their subjective aims regarding their life and experience, and its reality.

Even where people are incapable of carrying out or expressing their will but we can discern its content, acknowledging the importance of this kind of unity distinguishes a type of substituted judgment from a "best interests" approach. The typical form of substituted judgment discussed in the medical ethics literature, what I will call SJ_1, refers to an attempt to assess what people would want in some circumstance were they able to assess the situation competently. This corresponds quite closely to Dworkin's understanding of what lies in an agent's best interests.[21] A distinct sort of substituted judgment, SJ_2, attempts to glean what people, as they are currently

capacitated, actually do want, though they may be incapable of expressing and implementing their desires. The former type of substitute judgment, though a popular conception, is given no support by the analysis of autonomy's value where competence will not be restored. To the contrary, where there has been significant and permanent personality change, the moral significance of the results of SJ_1 should be called into question. There may be a large gap between what people do and will continue to want, given their extant personality and abilities, and what they would want were they competent. What the competent self may want of and for the incompetent self may differ dramatically from what the incompetent self actually wants. Given that the competent self will not have to live with his or her directions and the incompetent self may not welcome them or recognize them as his or her will, it is unclear what autonomy values are furthered by SJ_1 in the case of permanent personality change.

There are, however, autonomy-based reasons to attend to the results of SJ_2. The fact that one cannot *express* or *implement* one's will does not eviscerate the value of there being some unity between one's will and one's experience. SJ_2 is simply a means of discovering what it would take to achieve this unity.

Both forms of substituted judgment are not fully "substitutable" for the first person exercise of autonomy. Absent from both forms is the control over one's life that is achieved through self-implementation.[22] On both the integrity and the control accounts of autonomy, there is a special significance to one's engaging in processes of *self*-creation and *self*-direction. These accounts may recognize autonomy value in some types of substituted judgment, but they may still exhibit a principled preference for actual or first personal judgments and implementation.

The point I am trying to make may be approached from a different direction. As I remarked earlier, demented people are often analogized with children to make obvious the permissibility of paternalism toward the demented. But this analogy should set off some alarms. If we were to accept the evidentiary and integrity accounts as the sole underpinnings of the value of autonomy, then shouldn't we begin to wonder why we respect children's decisions as often and in the ways we do? Much attention is paid to the permissibility of paternalistic interference for children. It is routinely emphasized that children lack proper deliberative capacities for autonomous decision making. Strangely, there is little attention to the fact that paternalistic interference is not continuously exercised over many aspects of children's lives. Further, many paternalistic actions are often exercised with a measure of reluctance and regret (albeit outweighed by the good we are accomplishing), as though we are both impeding and overcoming something of value.

Why do we respect so many of children's decisions and preferences about their own lives? Often, the evidentiary or the integrity accounts will not supply a persuasive answer. Children do not always or consistently know what is in their best interests. Moreover, few of them conceive of or have the capacity to conceive of their actions as expressing a sense of self or a character constructed to reflect a chosen structure of values. As in Dworkin's description of the demented, many of their decisions may not flow from a coherent character or form part of a reflective project to construct such a character around a system of values. Nonetheless, absent substantial risk or miscalculation of possible benefit, children's decisions and preferences exert substantial moral weight and command our respect (though how much weight is recognized varies from parent to parent), even though their deliberative and reflective capacities are diminished and insufficiently developed to be convincingly assimilated into the evidentiary or integrity models of valuable autonomous behavior. Even where we override their preferences, we feel the residual need to justify our control and interferences. Our justifications are not always simply educatory but often apologetic and justificatory in tone and content.

Surely, some of why we respect children's decisions can be explained by the thought that decision making helps them to develop the proper skills and capacities required for "true" autonomous action later. This, however, cannot be the full explanation, and it certainly is not the simplest and most straightforward explanation in all cases. Not all of the everyday, trivial decisions children make are necessary for this learning process nor pertinent to the development of their characters. Further, it would not explain the residual justificatory obligation in cases in which justified paternalism takes place. Most pertinent to my claim, though, is the reaction we have to terminally ill children who we know will not grow into adulthood and will never have the fully developed intellectual and intentional capacities and the moral sensibilities associated with "full autonomy" and its richest realization. This case provides a more germane analogy to the case of dementia than that supplied by an analysis of our general treatment of children with full futures. In this case, I do not think that our conviction that we should accede to many of the terminally ill children's plans and demands would wane, even in cases where we thought the plans were seriously (though not dangerously) flawed. Respect for the fact that the children's lives is theirs would exert force over us even where their capacities were not fully developed and could not become fully developed through practice. Our reactions to this case show, I think, that we recognize that there is some value simply to expressing one's will and thereby determining the nature of one's experience, even where such control does not reliably maximize one's interests or contribute to the conditions necessary for a life with full Dworkinian integrity and unity.

Dworkin argues that showing respect for autonomy of people who have become demented requires respecting their *precedent autonomy* by respecting the decisions they forged before the decline into dementia. He also argues that precedent autonomy surmounts any interests to the contrary these people might now have. I contest both of these claims. As I have been arguing, the decision that is the product of precedent autonomy may conflict with the preferences that the demented person currently exhibits or expresses. The precedent decision may direct that people be killed or allowed to die should they become demented, although while demented they resist this decision and affirm a willingness to live. Alternatively, prior to dementia, they may direct that no assistance be given to them to enable them to die, although during their dementia, while in pain and distress, they repeatedly articulate a will to die. Dworkin insists that respect for autonomy entails following these precedent decisions because the person cannot retract this decision through a fresh exercise of the will. I am less sure of this conclusion. There is some value to the will that demented people express. These expressions may command our respect as exercises of autonomy, albeit etiolated ones, and they may suffice to nullify previously made decisions. Exercise of this type of autonomy might not command respect where fuller capacities for autonomy were present but suppressed (as in the case of the temporarily deranged). But if this control autonomy is the only type available, given the person's capacities, there may be better reasons to take heed of it. Unlike the Jehovah's Witness case, a self with an altogether different will is not going to have to live with the results of the action undertaken. That a person's full capacities have been permanently stripped from him or her provides reason to challenge the continuing force of a past decision. It may also lend greater weight to the exercise of whatever capacities remain.

One might speculate that what really motivates the suggestion that there is reason to honor the demented person's preferences is a rather different objection: that the demented person is a different person from the issuer of the advance directive. If the person has changed this considerably, perhaps it is questionable whether the demented person is indeed the same person as the person who wrote the advance directive. Not only have there been significant physical and psychological changes in the person, but these changes did not occur under conditions in which it

was possible for the person undergoing change to subject it to conscious direction or control.[23] If the demented person is not the same person who wrote the advance directive, then there is no autonomy-centered reason to respect the advance directive. The power of autonomous decision does not extend this far.

Dworkin acknowledges that his conclusions depend upon the assumption that personal identity survives dementia. He frequently reiterates that demented people are not new, distinct persons from the persons with heightened intellectual powers who once occupied their bodies. Rather, their dementia represents a stage of an entire life of those who have become demented. Instead of examining Dworkin's defense of this position, I am more interested in disputing a potential objection that the autonomy considerations I have forwarded disguise and depend upon psychological reductionist views about personal identity. To dispel this suspicion, let us take it as a given that, for Dworkin's reasons or for others, personal identity is preserved where a person declines into dementia, even if there is a radical, abrupt change in character.

I resist the claim that if personal identity is maintained, then it follows that respect for autonomy entails that P_1's decisions about P_2 rightly govern P_2. There is something puzzling about Dworkin's conjunction of claims that P_2 is a person, that P_2's decisions do not manifest autonomy values, and that P_1's decisions should legitimately control P_2. Dworkin's justification for ignoring P_2's decisions derives from the fact that P_2's capacities are diminished and her personality is shifting and unstable. But if shifts in personality are sufficient to disqualify one from exercising respectable autonomy, then it is unclear why the major shift in personality and in capacities undergone by P would not pose a barrier to justifying the extension of P_1's power over P_2. If the value of P_1's autonomy is explained in virtue of the expression of P_1's *continuing* character, then dementia cases should trouble us, for they are marked by a sharp, uncontrollable, and unmeditated disruptive change in character. Given this disruption, it is puzzling why Dworkin would think it was obvious that P_2's life was properly governed by P_1's decisions. For what happens to P_2 and what P_2 does cannot seriously be taken as a *direct* expression of P_1's character. That P_1 could indirectly express her character by manipulating P_2 does not justify such actions, any more than it would justify coercing a third party on the grounds that such coercion might express something about oneself. The directness of expression is a vital condition for underwriting protections for autonomy under the integrity account.

Strangely, it is often assumed that if nonreductionism about personal identity were true, we could confidently explain such things as binding promises over our future selves no matter what their features and differences. It may not be so simple if the justification for respecting autonomy does not coincide with the account of what explains continuity of personal identity.[24] Suppose, for instance, the justification for respecting autonomy emanates from considerations relating to the expression of personality or character, and the correct account of personal identity appeals to bodily continuity, or continuity of the brain. Then there is a further question beyond whether and how autonomy has value: namely, we must confront the issue as to what the proper temporal boundaries governing autonomous decisions are. Depending on the account of autonomy, it may not be obvious that the temporal range of one's control should extend to encompass the entire span of one's existence as the same person over time. Even if psychological reductionists like Parfit are mistaken about what personal identity consists in, they may still be right about the conditions necessary for assessing responsibility.[25]

Thus far I have argued that precedent autonomy may lack the range necessary to justify the application of advance directives to demented people. Second, I have argued that, contrary to Dworkin's claims, there is no warrant to conclude that the expression of the demented person's

will has *no* autonomy value whatsoever. We have some reason to implement their will, even when it conflicts with the directions of their past competent selves. I have not attempted to show that the will of demented people should command anywhere near the respect we owe to the competent. As with children, the will of demented people may have some autonomy value, but that may not tell the whole story. The autonomy of demented people may command some of our respect, but they need greater guidance about what sorts of things will achieve their ends and greater help to implement them. Further, given their diminished capacities for deliberation, their autonomy could, in appropriate circumstances, be trumped by our concern for their interests. If either of the critical tasks I have been pursuing is successful, then we must investigate what is in the interest of the demented person.[26] In the next section, I argue that Dworkin's account of beneficence cannot salvage his defense of advance directives and that the account is in some tension with his stance of religious neutrality.

III

Dworkin recognizes that apart from any respect we owe to the demented person's autonomy, whether previously or contemporaneously exerted, we must also explore what duties of beneficence we have toward the demented. Not surprisingly, Dworkin believes that considerations of autonomy and beneficence both support honoring the advance directive. Even if precedent autonomy as valued by the integrity account cannot extend its reign over P_2, the directive may be the best available evidence of P's interests.

Dworkin argues persuasively that P_1 may feel that P has strong "critical" interests concerning what happens to P_2. Some may feel that a life of unity and integrity may be spoiled if a prolonged stage of dementia is tagged on to the end of it. As Dworkin puts it, "as we judge a literary work, . . . [a] bad ending mars what went before" (p. 27). Even so, it remains to be shown that P_1's interests concerning P_2 are relevant to our beneficent concern for P_2.

Dworkin's effort to meet this burden may be reconstructed in the following way. He acknowledges that demented people experience pleasure and pain and therefore have contemporary experiential interests. Demented people also have critical interests, although they are incapable of sound reflection about them. These critical interests, however, attach only to P_2 *qua* P, that is, through her connection to P_1; she has none *qua* P_2 alone. These critical interests of P_2 arise in virtue of P_1's more complex character; P_1's vision of the unity, integrity, and character of her life; and her previously entertained views about how her life expresses and respects the sanctity of life. Dworkin argues that P_2 has no critical interests in any forward-looking way. That is, P_2 *qua* P_2 has no positive critical interest in the continuation of her life since P_2 is incapable of forming and pursuing the acts and attachments that give life value. The value of a life, Dworkin insists, must be generated from the inside, and this is no longer possible for P_2 to do, given her diminished capacities (p. 230). Her continued existence can only serve her critical interests insofar as her existence would be a symbol of P_1's persistence or tenacity, qualities that P_1 more resolutely and deliberately manifested.

Prima facie, this position seems like a strange one for Dworkin to occupy. Given his professed aim to avoid partisanship about the sanctity of life, it seems odd for him to say that the life of the demented cannot be made valuable enough for it to be in the demented person's critical interest *qua* demented person for it to continue. I will return to this difficulty later on, but I will here articulate a possible line of reply to continue the exposition of Dworkin's view of beneficence.

Perhaps Dworkin would attempt to evade this difficulty by arguing that the reason why continued life cannot lie within the critical interests of demented people is not that their life itself is not valuable, but due to some fact about the *nature* of a critical interest. Dworkin's thought may be that for a person to have a critical interest, it must be possible at least for it to be grasped, recognized, and adopted by that person. Since demented people cannot have a perspective on what it is that makes their lives go well, they cannot, therefore, *qua* demented people, have critical interests.[27]

If demented people *qua* demented are unable to have critical interests, then conflicts may arise between P_2's contemporary experiential interests and P_1's critical interests (as well as with P_1's precedent autonomy). But no conflict can arise between the *critical* interests of P_1 and P_2. This latter claim, that P_2 cannot generate critical interests that diverge from P_1's, permits Dworkin to advance a reconciliatory strategy that eliminates the possibility of conflict between autonomy values and beneficence. Dworkin argues that since precedent autonomy will likely be guided by P_1's assessment of her interests and since such decisions will reflect P_1's weighing (or otherwise taking account of) of her critical interests against P_2's experiential interests, there is no conflict between respecting precedent autonomy and discharging our duties of beneficence toward P_2. "Once we accept that we must judge [the demented person's] interests as she did when competent to do so – then the conflict between autonomy and beneficence seems to disappear" (p. 231). Beneficence, in Dworkin's view, provides us with a separate positive ground for respecting advance directives, because their implementation will further the critical interests of P.

There are a number of difficult issues over which this argument too hastily glides. First, there is the question as to whether – if beneficence is our aim – we should place priority on the critical interests of P_1 over the experiential interests of P_2. Dworkin does not provide a strong justification for sharply discounting P_2's experiential interests. He merely notes that these interests would have been taken into account by P_1 when contemplating the advance directive. If P_1 has legitimate authority over P_2, because of P_1's autonomy, that might give us reason to act as though P_1's decision was well-considered and took all relevant factors into account. Surely, though, it does not settle that the *beneficent* thing to do, therefore, is to satisfy the critical interest of P_1. There are reasons to doubt the evidentiary strength of the advance directive. Although people may generally assess their own interests well, imagining what it would be like to be demented and to endure experiences not entirely accessible to mental and emotional processing is not an easy task. There are reasons to doubt how perceptive P_1's judgment would be about the importance and nature of these experiential interests. Dworkin's claim cannot be that P_1's assessment just makes it the case that the denial or frustration of the assessed critical interests is *perforce* counter to P's best interests. Such a move would represent backsliding on his previous acknowledgement that one can get it wrong as to what is in one's critical interests (see, e.g., p. 202), and would be inconsistent with his critique of the purely evidentiary account of autonomy.

Dworkin needs to argue that critical interests just matter more than experiential interests do, full stop. This may often be true, but the case of demented people challenges this generality. For reasons similar to those already discussed, we should worry about whether the critical interests of P_1 should exert much moral force, given P's abrupt change in personality. Even if on the best account of personal identity, P_1 and P_2 are the same person, given the dramatic personality change P has undergone, it is unclear whether we should think of P_1's interest in how P_2 is treated as highly morally salient. P_1's interest may be entirely understandable. P_1 and P_2 share a body, a history, and some psychological features and habits; they are the same person, and this alone could surely make it understandable that P_1 takes an interest in P_2's condition. Further, P's circle of

friends and family may transfer their attachment to P_1 to include concern for P_2 as well, despite P's lack of a continuous character. That does not mean, though, that we should afford much weight to the interests of P_1 in our treatment decisions of P_2. By analogy, we would not be compelled to give much weight to an understandable interest P_1 might take in her doppelganger, given the inevitable connections others might make between them, if our aim were to act beneficently toward the doppelganger. Because P_2's psychological composition differs so radically from P_1's and because P_2 cannot self-consciously contribute to the life project envisioned by P_1, it is not obvious that the beneficent concern for P we evince during P_2's life should naturally focus upon P_1's interests. That is, where there has been substantial involuntary psychological change, maintenance of metaphysical personal identity may not be what is relevant from the normative, beneficent viewpoint.

The persuasiveness of this point about interests may well depend upon the plausibility of the earlier discussion about the temporally restricted range of P_1's autonomy rights and interests. Dworkin's account of beneficence is vulnerable in independent ways, though. Suppose Dworkin is right that P_1's critical interests should exert some legitimate moral force over and above the weight that the interests of others in P_2 exert. Another issue stands in the way of affirming the propriety of implementing advance directives, namely whether P_1's critical interests displace or outweigh P_2's experiential interests.

We might be inclined, with Dworkin, to think critical interests have greater weight. They reflect a more sweeping, deliberate, and deep vision of a life. They concern what makes a life go well, not simply what makes it enjoyable. In the life of a person with a multifaceted personality and an ability to be self-reflective, whether one's whole life goes well seems obviously of greater moment than whether fleeting pleasures are enjoyed or pains endured. Further, reflective people's ability to understand their critical interests and their importance may allow them to cope with and rationalize the frustration of experiential interests. Demented people, however, do not have a personality as fitted to coping with the frustration of experiential interest in this way. They cannot appreciate the significance of their critical interests as P_1 conceives them, nor can they grasp their connection to themselves. It seems cruel in such cases to force such people to live through agony so that they will fulfill a critical interest of living a life marked by fortitude and perseverance, an interest they no longer recognize, accept, or even understand. Likewise, it seems barbaric to ignore the experiential interests of people whose lives are filled with simple joys and delights and to cut these lives off short – to kill them – so that the entire life of P is unsullied by a lingering period of intellectual deterioration and decline. If all that people are capable of is awareness of their experiential desires, there are good reasons for such interests to play a more prominent role in one's analysis of beneficence than they do against the backdrop of deeper critical interests.[28] Whether or not it is appallingly cruel to implement a directive that runs counter to the only remaining interests a person has, surely it stretches the notion of beneficence to a breaking point to disregard these experiential interests in the name of kindness.

Dworkin considers the objection that experiential interests may take on greater significance in circumstances of dementia, but his replies sidestep the point. He acknowledges that demented people lack a sense of their own critical interests, but he analogizes to the case of the permanently vegetative and argues that their obliviousness to their critical interests does not provide us with "a good reason for ignoring their fate" (p. 232). Of course, I do not advocate disregard for demented people's fate. Quite the contrary. What is at issue is what proper regard for their fate consists in and whether it requires placing very strong weight on their contemporary desires and interests. Analogizing to the case of the persistently vegetative is inapt. The persistently

vegetative person has no consciousness whatsoever and hence no experiential interests. We may be convinced that advance directives should govern our actions toward persistently vegetative patients but nothing would follow about demented patients, because the complexity of their case involves their having ongoing, conscious lives.

In any case, I suspect that the situation of demented people is even more difficult. So far, I have argued that Dworkin fails to give a persuasive interpretation of beneficence that would support implementing advanced directives as such. The experiential interests of demented people may conflict with the directive's dictates. These interests plausibly exert more moral force than Dworkin admits.

There is yet a further problem with Dworkin's analysis of beneficence. Dworkin is overcasual in his dismissal of possibility that P_2 may have critical interests of her own that are independent of P_1's assessed critical interests.

Dworkin's denial that P_2 may have critical interests of her own is unconvincing. He only directly addresses the case in which the advance directive instructs the termination of P's life pursuant to P_1's critical interests. Here, he takes pains to stress that the demented person cannot have a positive critical interest in continuing her life since the life itself cannot manifest any internally generated value.

But the plausibility of such a maneuver, I suspect, derives from the narrow range of examples Dworkin attends to at this point. While advancing this claim about critical interests, Dworkin addresses only the question of whether it could be in one's positive critical interest to continue living. This, however, leaves open a significant hole. Even if positive value cannot be "poured into a life" from outside it, because active agency, identification, and reflection are necessary to grasp and realize the positive value of the life, surely it must still be possible for a life to go badly due to externally originating forces that have an effect on the person: that is, it surely is still possible for negative value to be brought into a life from the outside. Indeed, Dworkin must concede the possibility of external forces affecting the value of a life if he is to make any sense of the practice of writing and respecting advance directives to protect one's critical interests. As Dworkin argues, how one dies and whether one's body persists past the point of one's mental lucidity can be a matter within one's critical interests, even if one's death or one's continued existence is not self-inflicted or self-generated; he makes this point even about persistent vegetative patients who are not aware of their existence, much less able to contribute to it. If so, then it seems surely plausible to say that if a demented person's life is full of pain and misery and empty of the compensation of pleasure, that sort of existence – taken in itself – is not simply a hedonic disaster but is not a worthwhile life at all. It is in the critical interest of that person not to continue it.[29] If the contents of a life can run counter to one's critical interests, independent of one's awareness or ability to conceive of them, then it is not only possible for demented people to have critical interests, but it is possible for the critical interests of the demented person, along with that person's experiential interests, to conflict with the critical interests of P_1.

We cannot easily conclude that the beneficent thing to do is to further the critical interests of the demented person as that person conceived them prior to the onset of dementia, for they may have changed dramatically along with the person's capacity to conceive of them. Before affirming that implementation of the advance directive is permissible, much less required, we would have to consider whether there are any reasons to hold the previously conceived of critical interests as exerting force, and then we would have to explain why they would trump the person's contemporary critical interests. I do not know how any general argument, abstracting from the particulars of a specific case, would go and I doubt one could be easily constructed.

Further, if Dworkin concedes the possibility that externally originating events can run counter to one's critical interests, then soon we will be led to ask why it could not be possible for externally originating events to further one's critical interests – to make one's life go better – although one made no positive, reflective contribution to their furtherance. As Dworkin explains it, whether something is in one's critical interest depends upon whether it makes one's life as a human being go well. So the question we should ask initially is not what demented people are capable of constructing out of their lives, but whether the life of a demented person *qua* demented person can go well or badly at all. I have tried to suggest some ways in which demented people can have terrible lives, such as when they are kept in constant pain out of which they can make no sense in order to serve purposes they cannot understand, much less identify with. Surely it is also possible that the life of demented people can go well in ways not orchestrated or even understood by those people. They may thrive to the best of the abilities they have. This could render their lives good ones – certainly better than lives characterized by frustration and despair.[30] Moreover, their continued livelihood may provide solace and joy to their loved ones. They may receive care and concern from other human beings. Participating in these relationships can make one's life go well and may make persistence part of one's critical interests even if this life is not directed by any self-conscious vision or plan. It seems possible, then, that continuing or ending life could impact the critical interest of a demented person *qua* demented person. Contrary to Dworkin's suggestion, there is no easy assurance available to us that the path of beneficence is drawn by people's conception of their critical interests when they were lucid. And if that path is not clearly drawn, then the claim that precedent autonomy and beneficience coincide is in trouble. The critical as well as the experiential interests of demented people can diverge from the interests of those people as they conceived them prior to dementia. These freshly generated interests must command some of our concern. If that concern is directed at the demented patient, our beneficent impulses may well conflict with precedent autonomy.

These intuitive worries may be grounded in a more explicit account of the difficulty that Dworkin's claims about the demented's critical interests place him in. Dworkin has explicitly ruled out taking a position on the sanctity of life. He has also persuasively argued that part of what makes a life go well may depend upon whether it reflects and respects the sanctity of life. If Dworkin consistently refrains from taking a position on the conditions under which life has intrinsic value, how intrinsic value is properly respected, and specifically, whether it has value only when enjoyed or affirmed by reflective beings, then he will not be able to rule out the claim that the demented person has a critical interest in living in such a way that reflects and respects the sanctity of life. Thus to remain consistent with his stance of neutrality, Dworkin cannot, for instance, eliminate the possibility that demented people have a critical interest in continuing their lives even if those lives lack much intentional reflection, direction, or internal drive. This is not to say that Dworkin's implicit substantive (in his view, religious) position about critical interests is thereby proven to be wrong. There is much to be said for his underlying view that what makes a life positively valuable depends on internal generative activity and self-awareness. This view may indeed reflect the proper understanding of what is intrinsically valuable about human life. The problem, though, is that this view cannot underwrite his arguments for advance directives if these arguments are supposed to fit into a framework that maintains its religious agnosticism.

If Dworkin's conclusions about advance directives are to be defended, a different line of justification needs to be presented. Dworkin seeks to defend advance directives without taking a controversial stand on the value or sanctity of life. He thinks this can be accomplished by embarking

on a neutral analysis of autonomy and the nature of interests, critical and experiential. But if I am right that an analysis of the nature of critical interests will not assure us of the coincidence of considerations of beneficence and autonomy, then Dworkin will have to make arguments about the substantive content of these interests. He will have to argue more directly that the life led by a demented person cannot matter and cannot have value. He must argue that the life that is led will not contain experiences of a type that are worthy, but not because the nature of critical interests and valuable lives are such that it is impossible for external events to impact them. Dworkin will have to argue, rather, that the capacities of a demented person are simply too impoverished to befit a human being, and that such a life is not worthwhile. This argument will surely engage him directly in the debate about the nature of the sanctity of life.

Perhaps one could defend this somewhat exalted standard of what it takes to make a life worthwhile. On that issue, I will not comment. I will instead make some concluding remarks about the theoretical position Dworkin is placed in. Dworkin could depart from neutrality. To defend the implementation of advance directives calling for termination of the demented life, he could take the view that the critical interests a demented person may have in continuing life are insufficiently important to make the life worthwhile. This would expose that his defense of advance directives does not sit comfortably with a position of religious neutrality (as Dworkin conceives of it), but rather flows from a substantive view about which kinds of lives have objective value. Further, such a defense of directives would be importantly partial. If Dworkin adopts the substantive defense, then he will be harder pressed to defend, as a matter of beneficence, implementing any particular advance directives that require that the demented life be prolonged, despite the critical and experiential interests of the demented patient. If Dworkin abjures taking the substantive line on the sanctity of life, then the argument for implementing advance directives falls importantly short. For, if my arguments have merit, then it is possible for the demented person to have interests that deeply conflict with the interests and decisions the person rendered when lucid. It is then a much more difficult task than Dworkin allows to reach the conclusion that the beneficent way to treat demented people is to disregard their contemporary interests and to further their prior interests and decisions. Fulfillment of past relinquished desires may play some role in one's contemporary interests, but we tend to believe that their importance diminishes significantly if they are replaced by disparate contemporary interests.

At their strongest, the arguments I have presented show that Dworkin is mistaken to hold that we treat demented people with most respect by implementing advance directives. Far from being morally required to do so, we may not even be morally permitted to do so. The contemporary interests and the will of demented people may clash with their exercise of precedent autonomy and the critical interests they had before they became demented. We show respect for the demented by allowing their contemporary interests and their voiced will to exert substantial influence on our deliberations.[31] At their weakest, the arguments show that Dworkin's neutral defense of advance directives is importantly incomplete. For it to be convincing, we either need some account about how we may resolve the clash between the contemporary interests and the past interests of the person that favors the latter, or we require a revamped account of autonomy that can more convincingly explain why prior decisions merit respect and should be implemented even when they conflict with beneficence properly understood. This, as I have suggested, is a particularly difficult claim to sustain on the integrity account of autonomy as well as upon the alternative control account I have forwarded. In any case, Dworkin's reconciliatory approach leads him to avoid tackling these difficult balancing questions. If I am correct, we will have to face them.

Acknowledgement

I am grateful to Torin Alter, Justine Burley, Ruth Chang, Laura Cherry, Sean Foran, Marc Lange, Herbert Morris, Michael Otsuka, Eric Rakowski, Steven Shiffrin, and the members of the UCLA Law and Philosophy Discussion Group for helpful conversations and written comments about this material. Enid Colson provided valuable research and bibliographic assistance.

Notes

1 Ronald Dworkin, *Life's Dominion* (New York: Knopf, 1993). Page references in the text and notes are to this book unless otherwise specified.

2 See e.g., "Autonomy and the Demented Self," *The Milbank Quarterly*, vol. 64, 1987, pp. 4–16; "Philosophical Issues Concerning the Rights of Patients Suffering Serious Permanent Dementia," in *Philosophical, Legal and Social Aspects of Surrogate Decisionmaking for Elderly Individuals: Contractor Documents for the Office of Technology Assessment* (PB87-234126) (Washington, DC: Office of Technology, May 1986); "The Right to Death," *New York Review of Books*, January 1991, pp. 14–17; "Unenumerated Rights: Whether and How *Roe v. Wade* Should Be Overruled," *University of Chicago Law Review*, vol. 59, Winter 1992, pp. 381–432; "The Center Holds!" *New York Review of Books*, August 1992, pp. 112–15. Relevant chapters also appear in *Freedom's Law* (Cambridge, MA: Harvard University Press, 1996), and in *Sovereign Virtue* (Cambridge, MA: Harvard University Press, 2000).

3 For criticisms of this conception of the religious and of his interpretation of the First Amendment see Student Note, "Inside Out, Within and Beyond, or Backwards?" *Harvard Law Review*, vol. 107, 1994, pp. 943–7, and Eric Rakowski, "The Sanctity of Human Life," *Yale Law Review*, vol. 103, 1994, p. 2049.

4 Dworkin characterizes the dispute about the sanctity of life as one regarding the "relative moral importance of natural and human contributions to the inviolability of human lives" (p. 91). Presumably, though, another aspect of the dispute concerns the relative weight that should be given to the intrinsic, impersonal value of life when it conflicts with the subjective value of a life. This is at least some of what is at issue in debates about the right to die.

5 This "encouragement" may include the demand that they come to a definite decision about their pregnancies by a particular date. Dworkin's view that late-term abortions may be prohibited, because it is reasonable to expect women who seriously reflect to come to a decision earlier, is based on his dubious presupposition that women who elect late abortions do not take the sanctity of life seriously. But information relevant to one's decision may not be available until late into the pregnancy and relevant factors may change within a pregnancy. Elsewhere, Dworkin acknowledges that acting on reasons relating to the future health of one's child or one's familial or financial situation can reflect a serious view about the sanctity of life; why should the emergence of these reasons late into pregnancy make acting upon them frivolous? Finally, consider a person who, from the onset of pregnancy, struggled hard to consider the relevant issues and to make a commitment to the future child. As the pregnancy progressed and she experienced its mounting burdens, she realized that she could not make the kind of full-fledged good-faith commitment to the life it merited, so she aborted out of respect for the importance of a life's going well. As Dworkin recognizes, philosophers who devote their lives to understanding the significance of life frequently change their minds and have not been able to reach a stable consensus within four to five months. If those who study the subject believe the problem to be extremely difficult and perhaps intractable, why is it reasonable to expect faster deliberation on the part of pregnant women in particular? Why is lengthy deliberation ending in late-term abortion a sign that the woman was "indifferent to the moral and social meaning of her act" (p. 170)?

6 Laurence Tribe originally made an argument along these lines in "Foreword: Toward A Model of Roles in the Due Process of Life and Law," *Harvard Law Review*, vol. 87, 1973, pp. 1–53, although he retreated from the position in *American Constitutional Law*, 2nd edn. (New York: Foundation Press, 1988), p. 1350.

7 See e.g., T. M. Scanlon, "Ronald Dworkin and the 'Sanctity of Life'," *New York Review of Books*, vol. 40, July 1993, pp. 45–51; Galen Strawson, "The Termination of Wrongs and Rights," *The Independent*, June 1993; Student Note, "Inside Out, Within and Beyond, or Backwards?," op. cit.; and Laurence Tribe, "On the Edges of Life and Death," *New York Times Book Review*, May 16, 1993, Section 7, pp. 1, 41. Since the present essay was first written more has been written on these subjects and Dworkin's positions.

8 Since I wrote this, two prominent and excellent exceptions have come to my attention: see Sanford H. Kadish, "Letting Patients Die: Legal and Moral Reflections," *California Law Review*, vol. 80, July 1992, pp. 857–88; Agniezka Jaworska, "Respecting the Margins of Agency: Alzheimer's Patients and the Capacity to Value," *Philosophy and Public Affairs*, vol. 28, 1999, pp. 105–38. I regret there is inadequate space here to discuss my substantial agreement with them as well as my few reservations.

9 Dworkin discusses financing care in "Philosophical Issues Concerning the Rights of Patients Suffering Serious Permanent Dementia," op. cit.

10 This argument assumes that there is a higher order critical interest (or a condition on critical interests) that people have in living their lives according to their own assessments of their critical interests, even when these assessments are mistaken. Dworkin does not argue directly for this premise in the book, although he hints at it at p. 106. A more extended argument for "the constitutive view of value," the view that nothing may contribute to the value of a person's life without that person's endorsement, appears in his "Foundations of Liberal Equality," in Grethe Peterson (ed.), *The Tanner Lectures on Human Values*, vol. XI (Salt Lake City: University of Utah Press, 1990), pp. 75–86.

11 There may be some question about whether the seriously demented have a will or can exhibit preferences about the continuation of their lives. Some in the advanced stages of dementia may not be capable of the resistance I imagine. Only some of my conclusions will apply to their cases. Dworkin, however, does not limit his conclusions to those people and insists that out of respect for their autonomy, we may best respect people by implementing these directives even when they do express "choices and demands" (p. 225) and "wishes" (p. 228) to the contrary in their demented state. Later, I discuss his characterization of the demented. I will not take on the more challenging task of ascertaining how other decisions regarding the content of the lives of the demented should be made, should they continue to live, and who should make them.

12 This rules out those people who are born mentally impaired. My focus is on those who once functioned normally and who undergo mental decline.

13 See Jeffrey Cummings et al., *Dementia: A Clinical Approach*, 2nd edn. (Boston: Butterworth-Heineman, 1992), pp. 1–2.

14 See A. F. Jorm, *The Epidemiology of Alzheimer's Disease and Related Disorders* (London: Chapman and Hall, 1990), pp. 26–7.

15 See Joseph Foley, "The Experience of Being Demented," in Robert Binstock et al. (eds.), *Dementia and Aging* (Baltimore: Johns Hopkins Press, 1992), chapter 2, pp. 30–43.

16 See, e.g., Barry Reisberg et al., "The Final Stages of Alzheimer's Disease: Issues for the Patient, Family and Professional Community," in Richard Mayeux et al. (eds.), *Alzheimer's Disease and Related Disorders* (Springfield, IL: Charles Thomas, 1988), pp. 12–16.

17 Other patients Dworkin describes as subject to his conclusions could not run a tub but recognized family members and were enheartened to see them. They could engage in simple activities with pleasure.

18 Dworkin, "Philosophical Issues Concerning the Rights of Patients Suffering Serious Permanent Dementia," op. cit. Other commentators have challenged advance directives on the grounds that the demented person is not identical to the author of the directive and on the grounds that the demented being is not a person at all. See e.g., Dan Brock, *Life and Death* (Cambridge, UK: Cambridge University

Press, 1993), pp. 356–87; Dan Brock and Allen Buchanan, *Deciding for Others* (Cambridge, UK: Cambridge University Press, 1990), and Eric Rakowski in this volume. I will argue that the central moral issue does not turn upon whether personal identity is maintained or disrupted, but hinges upon the abrupt involuntary change in personality.

19 For a similar account, see Gerald Dworkin, *The Theory and Practice of Autonomy* (Cambridge, UK: Cambridge University Press, 1988), pp. 17, 20, 26.

20 The evidentiary account treasures autonomy in the latter way, for its usefulness as a causal lever shunting agents toward their greatest welfare. One might read the "integrity" account as indirectly instrumental too, as prizing autonomy because it allows expression of one's character, where such robust expression over a lifetime ultimately promotes a person's welfare. Such a reading is possible, but I think mistaken. If integrity's value were instrumental, one would expect to encounter worries about whether bad decisions should be respected, even if their detrimental impact on the person's welfare were far greater than the self-expression they facilitated. The integrity account, as I understand it, does not aim to identify the goal that autonomy serves but instead identifies what it is about autonomy that makes it intrinsically valuable. It is in virtue of autonomy's exercise being a way in which one's character is created and expressed that it has intrinsic value and should be respected.

21 Not all take a "constructivist" approach to delineating what is in a person's best interests. In more objective accounts of interest, the contrast between substituted judgment and beneficence is sharper.

22 Thus I disagree with the claim that actual consent enjoys no intrinsically privileged moral status over hypothetical consent. See Eric Rakowski, "Taking and Saving Lives," *Columbia Law Review*, vol. 93, 1993, pp. 1114–15.

23 The lack of control over tremendous personality change is relevant in distinguishing between the case of dementia and more commonplace forms of character transformation and evolution.

24 Robert Adams and Jennifer Whiting both allude to related doubts about whether the truth of nonreductionism would provide justificatory reason for harboring concern for our future selves and for others' future selves. Robert Adams, "Should Ethics be More Impersonal?" *Philosophical Review*, vol. 98, 1989, pp. 439–84 at 455; Jennifer Whiting, "Friends and Future Selves," *Philosophical Review*, vol. 95, 1986, pp. 547–80 at 547.

25 Parfit and his critics frequently write as though his ethical theses about selves hold only if his theory of personal identity is true. See chapters 14 and 15 of *Reasons and Persons* (Oxford: Oxford University Press, 1984). But Parfit's arguments about personal identity often proceed by asking what is important about ourselves. On these issues, Parfit may well be right. Even if the metaphysical inferences do not follow, his normative insights about the significance of psychological change, especially severe disruptive change, might still hold.

26 This task would also present itself if there were no advance directive available. Dworkin suggests that a substituted judgment inquiry asking what the person would have chosen would both respect the person's autonomy and that the identical inquiry would ascertain the person's best interests (pp. 191–2). He does not take a stand on what to do if, contrary to his presuppositions, what is in the person's best interests differs from what that person would have chosen. Elsewhere, Dworkin criticizes substituted judgment as not promoting autonomy values at all. See "Autonomy and the Demented Self," pp. 14–15. As I argue in the text, substituted judgment as it is typically conceived, SJ_1, carries little autonomy value. SJ_2 has some autonomy value, but its results should not necessarily wholly determine our actions with respect to the permanently incompetent if there is a sufficiently serious conflict with the person's best interests.

27 This interpretation finds support in his arguments in favor of constitutive over additive models of value in his "Foundations of Liberal Equality," op. cit., pp. 77–8.

28 There is another argument for placing greater weight on the experiential interests of demented people over their critical interests. When a life has the potential to go superbly, experiential frustrations seem of paltry importance in comparison with the caliber of critical interests at stake. If due to constrained

capacities or limited circumstances, a life can go better or worse but cannot be a very good life, then the significance of experiential interests might be thought to be more closely comparable to the significance of critical interests. How excellent the life can be, it might be suggested, has a bearing upon whether critical interests automatically overwhelm experiential interests.

29 See also "Foundations of Liberal Equality," op. cit.: "Our lives may go badly . . . not just because we are unwilling or unable properly to respond to the circumstances we have, but because we have the wrong circumstances" (p. 69).

30 Such a story seems consistent with, and even reminiscent of, some of Dworkin's approving remarks about the model of challenge, the idea that living well is a matter of successfully meeting the challenges that a life presents. See "Foundations of Liberal Equality," op. cit., pp. 57, 60, 67. Dworkin's related arguments for the constitutive model of value, the view that nothing contributes to the value of a person's life unless it is endorsed by that person, do not actually rule out the possibility that various activities may lie within P_2's critical interests. Dworkin insists that it could not further one's critical interests to lead a life that one despised and actively reviled as unworthy (ibid., p. 76). Dworkin's arguments appeal to our intuition that a life cannot go well if it is actively *rejected* by a person. This intuition, however, is entirely consistent with the belief that demented people's lives may go well and may go better than they otherwise would if they contain certain features and activities, even if the people lack the capacity to recognize that their way of life furthers their critical interests.

31 I have not, however, discussed how one should balance the autonomy values against any conflicting contemporary interests. Such a task would have to be taken up to give a full account of the proper treatment of the demented.

Part III

Seana Valentine Shiffrin

How should we respect the autonomy of a person who has become seriously and permanently demented? What is in the best interests of such a person? We must confront a preliminary but crucial metaphysical and ethical issue. Is a person suffering from serious and permanent dementia the same person as the competent person who became demented? If someone remains the same person after dementia, then we must ask about the autonomy or interests of that single person, and we must consider every part of that person's life as bearing on any such question. If, on the other hand, we accept that the person after serious dementia is a different person from the competent person that existed before, then we have two sets of questions about autonomy and interests to answer. In either case, we have to resolve conflicts, but the character of the conflicts is different.

If only one person is involved, we have to decide what is in his overall best interests when he has different and competing interests at different stages of his life. That is a familiar sort of conflict that everyone has to resolve: is it in my interest overall to sacrifice exciting travel now in order to prepare for a career that will flourish only later? Dementia may make the conflict more searing or even tragic, because the interests in question may be more antagonistic. But it does not make the conflict different in character. If serious dementia ends the life of one person and begins the life of another, however, then any conflict is between the autonomies and interests of two people, and raises questions of fairness. We might have to consider whether granting a certain form of autonomy to the competent person is cheating the demented person of his or her own autonomy, for example, or whether it is fair or unfair to act in the interests of one person at the cost of damaging the interests of another. These, too, are familiar questions in more pedestrian conflicts. They are questions of justice however, and though dementia may also make these questions more agonizing, they remain questions of justice rather than, as in the case of a single person, questions of intrapersonal balance.

I believe that the competent person and the demented person are the same person: the same, single object of ethical inquiry. The life of each is part of the life of the other, and only one person's autonomy and overall best interests are at stake. My arguments in *Life's Dominion* about the treatment of seriously demented people begin with that assumption. In her tightly argued and powerful chapter in this volume, Seana Shiffrin says that she is willing, for purposes of her discussion, to assume that I am right in this. In fact, however, much of her argument seems to me to reject the single-person assumption, and to depend for its force on rejecting it.[81] She speaks throughout of P_1 and P_2: the first is a person P before, and the second that person after, serious dementia. The distinction between the two Ps might, of course, be only a device to distinguish different stages of the same person's life. But the distinction actually plays a much more important role in her argument than that.

Shiffrin discusses at length my example of a Jehovah's Witness who calls for a blood transfusion when she becomes temporarily deranged. Shiffrin agrees with me that it would be an insult to the Witness's autonomy to perform the transfusion, but she insists that this is because such temporarily deranged people "will, or can, in time, return to their senses and will have to live with our action" (p. 202). But in the case of permanent dementia, she says, the competent person will not have to live with the consequences of what happens to the demented person, because the competent person will have disappeared, never to return. "P_1 may be the more capable decision maker, but P_2 will have to live with the consequences of the decision, not P_1" (p. 202). She repeats the point later: "Unlike the Jehovah's Witness case, a self with an altogether different will is not going to have to live with the results of the action undertaken" (p. 206). Soon after she states that she accepts, for the sake of argument, my assumption that personal identity is preserved after dementia, she directs readers' attention to "the interests of demented people" (p. 209) as if that could be considered separately from the interests of those people when competent. Indeed, she wonders whether "P_1's interests concerning P_2 are *relevant* to our beneficent concern for P_2," (p. 208, italics added) and calls it "odd" for me "to say that the life of the demented cannot be made valuable enough for it to be in the demented person's critical interest *qua* demented person for it to continue" (p. 208). She says that P_1's interest in P_2 "may be entirely understandable" since they "share a body, a history, and some psychological features and habits; they are the same person" and since "P's circle of friends and family may transfer their attachment to P_1 to include concern for P_2 as well" (pp. 209–10). But, however understandable, P_1's concern for P_2 is "by analogy," it is like the concern someone might have for his or her "doppelganger," and "we would not be compelled to give much weight to an understandable interest P_1 might take in her doppelganger, given the inevitable connections others might make between them, if our aim were to act beneficently toward the doppelganger" (p. 210). None of this is consistent with taking seriously the ethical identity of the people she calls P_1 and P_2. Questions about the autonomy or interests of a single person must be resolved within the life of that person. As I said, there are not two autonomies or the autonomies of two different people to balance or adjudicate, and though it does make sense to speak of the different interests of that single person at different times, these are all necessarily relevant to the further, inescapable, question of what is in the person's best interests overall.

Shiffrin's repeated remark that P_1 need not "live with" events in P_2's life is particularly striking and, I believe, misconceived. I distinguish, as she acknowledges, between two kinds of interests people have – critical interests and experiential interests. People's experiential interests are their interests in enjoying or avoiding certain kinds of experience: their interests in having pleasure, for example, and avoiding pain. Their critical interests are their interests in having a life that is a good one judged as a whole. Of course experiences that occur in a person's life after he has

become seriously demented are not experienced by that person in a competent state. But it does not follow, and it is opposed to most people's own judgment, that what happens to someone after he become demented cannot affect his critical interests.

That distinction is evident when we consider how a person's interests are affected by events after his death. Experiential interests end with death: burial or cremation is not either in or against the experiential interests of the deceased. But, as I assume throughout *Life's Dominion*, a person's critical interests can be and often are affected by what happens after his death. It was against Hector's interests that his corpse was dragged around the walls of Troy, for instance, and it would make an author's life less successful if his entire oeuvre was destroyed after his death. People often take steps to guard against ignominy or to secure or protect their reputation after their death, and they think they are acting in their own, not other people's, interests.

A person's experiential interests also end if he falls into the state I discussed in chapter 7 of *Life's Dominion*: a permanent vegetative state. But most people think their critical interests survive that event and become, if anything, more dramatic. Some think it important that they be kept alive as long as possible in a vegetative state, and others think it crucial, in their own interests, that they die as soon as possible. That is why some people make "living wills" providing that life support be terminated in that event, and why the Supreme Court insisted that such wills, if properly executed, be respected.[82]

It seems *a fortiori* plain, therefore, that a person's critical interests survive even serious and permanent dementia. A competent person looking ahead to impending dementia may be concerned about the quality of experience he will have: these experiences will, after all, be happening to him. That concern may affect the wishes he expresses about how he is treated then. Shiffrin is of course right to say that he is no longer competent when he is having demented experiences; in that sense, she is right to say that P_1 does not "live with" what happens to P_2. But since the critical interests of P_2 and P_1 are the same, P_1 (which is simply one of P's three names) certainly does "live with" those events in a more important sense: in one way or another they affect what kind of a life he or she will have had. People's dread of and preparation for dementia would be inexplicable without that assumption. They do not dread and prepare for events in the life of a doppelganger, but in their own lives; they act not in someone else's interests but in their own.

Now consider the Jehovah's Witness case again. Shiffrin says she agrees with me that it would be wrong to give a fervent Witness a blood transfusion when she is deranged even if she then asks for it. I said, in constructing the example, that the Witness's derangement is temporary; Shiffrin says that fact is crucial, but not because she may later regret that we complied with her deranged wishes. After all, she might later regret that, and writhe in agonized shame, even if she asked for it in a fit of weakness while still competent: our prediction of regret would not be enough to justify refusing her the transfusion in that moment of weakness. The fact that the derangement is temporary is crucial, Shiffrin says, because it means that the competent Witness will have to "live with" the fact of the transfusion. But since she does not mean that the Witness will have to live with it in the experiential sense – the experience of transfusion will not have occurred in a competent moment – she must mean that the Witness has to live *through* it: that it is a fact in her life that has, at least as she judges it when she is competent, soiled that life. That is exactly true of the person who dreads living on in dementia: he dreads that a stage of his life that he lives through will soil that life. He knows, or strongly believes, that that stage will occur; that is why he tries to guard against it. He cannot know that it has occurred, and in that way he is different from the Witness in my example. But the kind of pain he might feel at contemplating his dementia does not depend on whether that experience occurs before or after the event. In any case, Shiffrin says

that the difference has nothing to do with any such sensations or experiences, and it is therefore hard to see why it matters at all. In both cases, someone's life goes worse as he judges it because his competent wishes have been disobeyed. Shiffrin's argument does not sustain her claim of disanalogy.

She has two further arguments against my position, but these, too, depend on not taking seriously enough the identity of a person before and after dementia. First, she says that my understanding of the point of autonomy is too narrow. Autonomy does serve the "integrity" value I describe, she says, and that value, she agrees, is not available to someone while he is seriously demented. But autonomy also provides a value that she thinks is available then, which is "the basic value of being in control of one's experience and in not having experiences forced or imposed on one when one's will is to the contrary" (p. 202). Elsewhere she says that "the simple exercise of control over one's experience reinforces the special relation one has to one's experience" and confirms "the uniqueness and separateness of persons" (p. 203).

There are difficulties in this idea: autonomy alone is hardly enough to guarantee that experience will not contradict will, and control over one's experience is certainly not necessary to the personal distinctness of experience, which is, after all, never more evident than in torture. It is unclear, moreover, how far recognizing this part of autonomy's value is simply recognizing the experiential satisfaction of having one's own way and the badness of frustration. These are the values that best explain the deference to the wishes of young children that Shiffrin cites, and they are values that I mention in explaining why we should often allow demented people to do as they wish even when their acts seem pointless or inconsistent.[83]

I see no harm in treating these experiential interests as among autonomy's benefits on at least some occasions, however, even though I did not recognize them in that way. That would still leave open the question of how we best respect a single person's autonomy overall and we might still think, as I do, that we best respect this by giving powerful weight to an exercise of what I called "precedent" autonomy, that is, the autonomy exercised by a person when most competent to exercise it. After all, someone who is fully competent must feel frustration and a sense of lack of control when he is told that he cannot guard against what he regards as the coming savage indignity of dementia, that he cannot bring his life under the control of his will in that way. If that frustration counts distinctly at all, under the title of autonomy, then it counts as least as much in favor of precedent autonomy as in favor of the paler and dubious autonomy, grounded only in that way, that he might be able to exercise after deep dementia sets in. This is not a matter, to repeat, of justice: not a matter of cheating one person of autonomy to aggrandize the autonomy of another. It is just a matter of deciding what set of policies do best to protect a single person's autonomy overall.

Shiffrin believes, finally, that I confront a dilemma when I speak of that single person's best interests as distinct from his autonomy. How can I judge that it would be in the best critical interests of someone in serious and permanent dementia to die as soon as possible without taking sides, as I say I do not, on the controversy about whether suicide does in fact make the life of such a person go critically better? But I do not offer that judgment. I say that respecting the precedent autonomy of a patient who has expressed a clear wish to die in those circumstances does not conflict "with her critical interests as she herself conceived them when she was competent to do so."[84] I said that those called upon to act as she requested might have a different opinion, but that insisting on their own opinion would be "an unacceptable form of moral paternalism."[85] So I am not taking sides on the underlying ethical issue. We often defer to the judgments of competent people, about where their interests lie when we aim to act in their interests, as a fiduciary does, for example.

That deference does not collapse the distinction between beneficence and autonomy, because respect for autonomy requires us to follow someone's directions that are not even intended to be in his own interests.

I close by trying to guard against a possible misunderstanding. I did not say, as some of some of Shiffrin's comments might be read to suggest, that it would be all things considered right to deny a demented patient life-saving treatment or to kill him if he had requested this while competent. Since the point is important I shall repeat what I said:

> We might consider it morally unforgivable not to try to save the life of someone who plainly enjoys her life, no matter how demented she is, and we might think it beyond imagining that we should actually kill her. We might hate living in a community whose officials might make or license either of those decisions. We might have other good reasons for treating [a demented person] as she now wishes, rather than as, in my imaginary case, she once asked. But still, that violates rather than respects her autonomy.[86]