Dementia and Personal Identity: Implications for Decision-Making

Winston Chiong, M.D., Ph.D.

Abstract

Changes in personality and judgment in dementia raise both practical and philosophic problems concerning personal identity. For example, if a patient's values and preference have been altered by disease, should caregivers' decisions for the patient be guided by their understanding of the patient's premorbid values and preferences, or of the patient's current values and preferences? Some authors have argued, based on philosophic accounts of personal identity, that the previously healthy person and the demented patient should be regarded as two different people; therefore, previous values and preferences lose their authority. Meanwhile, others have argued that when patients lose their ability to act autonomously, their "precedent autonomy" must be respected by following their earlier wishes. These debates can inform our thinking about the proper interpretation of advance directives, as well as other practical questions such as how carriers of dementia gene mutations should think about their futures.

Keywords:

advance directives

Alzheimer's disease

frontotemporal dementia

autonomy

presymptomatic diagnosis

34.1. Introduction

Among Americans aged 55 and older, Alzheimer's disease is more feared than any other disease, including cancer, stroke, heart disease and diabetes (Metlife Foundation, 2006). One source of this fear may be that Alzheimer's disease and other dementias threaten not only our lives, but also our identities. In the case of Alzheimer's disease, this may be because we understand and interpret our lives in the context of memory; for instance, I might feel that if I were no longer able to recall first meeting my wife, or to recognize my children, then there is an important sense in which I would no longer be me. For some of us, our sense of self may also be closely tied to personality traits (such as independence) or intellectual abilities that are compromised by dementia (Chiong, 2011). Meanwhile, frontotemporal dementia is often accompanied by more radical alterations in comportment, personal relationships, and even religious or political values (Miller et al, 2001). In advanced dementia, clinicians and caregivers may find it natural to distinguish between the demented patient and "the person he/she used to be," as though they were two different people (Kaldjian et al, 2010).

Disease-related alterations of personality and judgment raise serious problems when clinicians and caregivers must make decisions on behalf of patients. The accepted ethical standard for such judgments is that surrogates should decide according to the patient's values and preferences; or, if these are unknown, according to the patient's best interests. However, if the patient's values and preferences have been changed by the disease—such that the patient may even seem to be a different person than he or she was premorbidly—which version of the patient's values and preferences should guide the decision? Similarly, should our judgment of the patient's best interests take account of premorbid interests (such as dignity, or not being a burden on one's family) that the presently demented patient may no longer affirm, or even understand?

Questions of personal identity in dementia have troubled not only clinicians and caregivers, but also philosophers. Can we make sense of the idea that the patient with dementia is a different person than the person who existed prior to the onset of illness? In the case of gene carriers or others at risk for developing dementia, how should their prudential concern for their futures be affected by the knowledge that their values and preferences may be altered greatly by disease? And, given the importance afforded to autonomy in contemporary bioethics, how can we respect the autonomy of patients who are no longer able to exercise it? Questions like these have inspired a rich philosophical literature, which I selectively review in this chapter.

34.2 The "someone else" problem

Consider a patient with advanced dementia (for illustrative purposes I have modified an actual case that has been extensively discussed—see Firlik, 1991; Dworkin, 1993; Dresser, 1995). Prior to her illness, Margo had been a fiercely independent woman who was devoted to social causes and strongly identified with intellectual pursuits. At various times she expressed to friends that she would not want to live in a demented state; for instance, that if she were to be diagnosed with Alzheimer's disease and later were to develop a serious but treatable infection, she would want to be allowed to die, and that she would want her property to be given to a designated charity rather than devoted to her own care. Later, unfortunately, Margo develops dementia, and no longer recalls or espouses these convictions. Instead, Margo comes to derive obvious satisfaction from simple pleasures such as eating peanut butter and jelly sandwiches, reading and rereading mystery novels, listening to songs over and over, and attending an art therapy group where she paints the same picture every week. While she previously would have regarded such an existence with repugnance, she now appears quite content in it.

If Margo were to develop pneumonia, prevailing ethical standards would require us to withhold antibiotics out of respect for her previously-expressed values, despite her present interest in surviving a treatable infection and continuing a life that she now finds agreeable. However, a radical criticism of this consensus has been advanced by Rebecca Dresser, who suggests that according to some philosophic views about personal identity, Margo has been so changed by the disease process that the earlier, healthy person and the presently demented patient really are two different people. If so, then the values and preferences of the earlier person (even if formally executed in an advance directive) would have no authority in guiding Margo's present care, because they were the values and preferences of someone else (Dresser, 1986; Dresser, 1995).

This criticism depends on complex philosophic claims about the nature of personal identity. Philosophers customarily distinguish between qualitative identity and numerical identity. To illustrate: suppose I commission an expert counterfeiter to produce a perfect replica of Matisse's *Dance (I)*. The new painting might be qualitatively identical to the original, yet the two paintings would not be numerically identical—I would not be the owner of *Dance (I)*, but merely the owner of a copy. Meanwhile, the original painting that now hangs in the Museum of Modern Art is numerically identical with the painting that Matisse produced in 1909, but it is not qualitatively identical to that painting. Not only has it undergone physical changes over time (fading of paint, expansion and contraction of materials), but also various restorers have altered it by reinforcing the canvas, applying a layer of varnish to its surface, later removing the varnish, and so on.

Thus, there are some qualitative changes that the painting can undergo, while still existing as the same (numerically identical) painting. But there are other changes that would leave us with a completely different painting, or with no painting at all; for instance, if a vandal were to scratch off the paint and paint a completely different picture on the canvas, or if it were reduced to ash in a museum fire. In these cases we would not say that Matisse's *Dance (I)* had been altered, but instead that it had been destroyed. Following Aristotle, one way of thinking about numeric identity is to consider what changes something could undergo and still exist, and what changes something could not undergo without changing into something else.

People also undergo radical changes over time—we mature from childhood to adulthood and continue to change with age, we are made up of cells that are continually dying and regenerating, and those cells are composed of molecules that are in constant exchange with the environment. These changes do not affect our numerical identities; for instance, we would ordinarily say that you were once a child, and not that there was once a child who later became you. However, Dresser suggests that in Margo's case, the changes that she has undergone in the course of disease have transformed the previously healthy, independent Margo into someone else.

In advancing this claim, Dresser appeals to one of two major philosophic views about personal identity. According to this view, termed the psychologic approach, our numerical identity over time depends upon psychologic continuity. This continuity might take the form of direct connections between psychologic states at different times; for instance, you might still recall a vivid memory from when you were 5 years old, or you might have learned something at that age that you still know. However, this continuity can also take the form of overlapping chains of such connections; for instance, even if you now cannot remember any events from when you were 5 years old, you likely can remember events from when you were 10 years old, and when you were 10 years old you could remember events from when you were 5 years old. Thus, this view can account for the claim that you are (numerically) the same person as when you were 5, even if there are no direct psychologic connections between your present self and your 5-year-old self, as long as there is a chain of intermediate stages that do bear such direct connections. (In the remainder of this article I will gloss over the distinction between continuity and connections; for further discussion see Parfit, 1984.)

The second view, termed the biologic approach, is that we are essentially organisms, and that our numerical identity over time depends upon the continuity of biologic, rather than psychologic, properties constitutive of organisms—among them being dynamic metabolic stability, goal-directedness, and organized complexity (Olson, 1997). While the psychological and biological approaches to personal identity agree about most everyday cases of personal identity, they diverge at the extremes of life. For instance, the psychologic approach implies that you were never a 3-month-old infant, and that you could not exist in a persistent vegetative state. (In a persistent vegetative state, according to the psychologic approach, you would cease to exist and would leave behind an unconscious organism that is not you.) Meanwhile, the biologic approach implies, contrary to many religions, that you could not survive after the death of your organism as a disembodied soul, or be reincarnated as another organism. (At best, there would be a soul or new organism that was psychologically continuous with you, but this would not be you.)

Dresser's claim that the healthy Margo and the patient Margo are actually two different people depends upon the psychologic, rather than the biologic, approach to personal identity. In advanced Alzheimer's disease, the ability to form new memories is extremely impaired, such that Margo might effectively retain no memories from one day to the next; the ability to access remote memories is also impaired, although not to the same extent (Sartori et al, 2004). Given these weak links of memory across points in time, there may not be a chain of psychologically continuous stages sufficient to link the patient Margo back to the healthy Margo.

One problem here is that, on the psychologic approach, it's not clear exactly *how much* continuity is required for someone to remain numerically identical over time. (A similar problem might arise in some cases for the biologic continuity of organisms within the biologic approach: see Chiong, 2005.) If a low degree of psychologic continuity is sufficient, then Margo could remain the same person over time despite severe loss of memory. If a high degree of psychologic continuity is required, this condition would support Dresser's claim that the healthy Margo and the patient Margo are two different people, but might have other implausible implications (Buchanan, 1988). For instance, as Margo remembers very few events from one day to the next, such a high threshold might imply that Margo is actually a new person every day.

34.3 Is personal identity really what matters?

Many philosophers accept an even more counterintuitive view: that (numerical) personal identity is not really what matters. This claim initially sounds strange because many important ethical and legal practices presuppose personal identity over time. For instance, we punish people for crimes that they have committed in the past, and we believe that it would be wrong to punish someone for a crime that someone else had committed. We also think that each of us has special reasons to be prudentially concerned about our own welfare in the future, and that this gives us reasons in the present to do things like save money, quit smoking, and fill out advance directives. However, many philosophers hold that this commonsense view is not precisely correct.

We might start by examining why philosophers who accept the psychologic approach claim that numerical identity is not what matters. Numerical identity is transitive: if A is numerically identical with B, and B is numerically identical with C, then A must be numerically identical with C. Derek Parfit has considered ingenious science-fiction cases in which someone becomes psychologically continuous with two future people—for instance, by entering a matter replicator, or by having his cerebrum divided with each hemisphere transplanted into a different body (Parfit, 1984). The two future people cannot be numerically identical with each other, so by the transitivity of identity they cannot both be numerically identical with the original person. However, Parfit argues that the original person should be prudentially concerned for both of the two future people. Thus, he concludes that identity is not what matters; instead, what matters is psychologic continuity.

Some philosophers who accept the biologic approach also claim that personal identity is not what matters. For instance, according to the biologic approach, if you were to suffer an irreversible neurological injury that leaves you in a vegetative state, you would still survive as a permanently unconscious organism. Many people think that being permanently vegetative would be just as bad as dying, and would have no more concern for the subsequent survival of their unconscious organisms than they would about the disposition of their corpses. Those who accept the biologic approach can accommodate these convictions by claiming that, while the unconscious organism would be numerically identical with you, personal identity need not be what matters for prudential concern (Olson, 1997). Thus, someone who accepts the biologic approach as an account of our numerical identity could also agree with Parfit that what matters for the purposes of moral responsibility or prudential concern is psychologic continuity.

If Parfit is correct that psychologic continuity and not personal identity is what matters, this assertion may allow for a more satisfying interpretation of Dresser's claims about Margo. One important difference between psychological continuity and personal identity is that psychologic continuity admits of degrees, while personal identity is all-or-none. For instance, you today are probably *more* psychologically continuous with yourself as a ten-year-old than with yourself as a five-year-old child, but it would not make sense to say that you are more numerically identical with yourself at ten than with yourself at five.

Because psychologic continuity admits of degrees, Parfit argues that prudential concern and moral responsibility also admit of degrees; furthermore, when psychologic continuity is reduced between two stages in a person's life, it would be reasonable for that person to be less prudentially concerned about, and to be held less morally responsible for the actions of, their future or past self. For example, some people believe that if you must choose between a small benefit now and a larger benefit in the future, rationality requires that you wait for the larger benefit. Parfit denies this belief.

If you expect your future self to be very psychologically different from the way you are now, Parfit claims that it would be reasonable to discount benefits to this future self; not because these benefits would be temporally distant, but because they would accrue to a later self who is very different from the person you are today. Similarly, Parfit argues that someone who had previously committed a crime and later underwent a great psychologic change (through the passage of time, or a personal epiphany) would deserve less punishment, and that this circumstance may provide a rationale for statutes of limitations on criminal penalties.

In Margo's case, it may not be plausible to claim that Margo in her demented state is actually a different person than she was before her illness; however, there is much less psychologic continuity between these two stages in her life than there would ordinarily be. This psychologic discontinuity explains why it may seem natural to think of Margo before and after her illness as two different people, whether or not this claim can be philosophically defended. Furthermore, given this psychologic discontinuity, her previously expressed values and preferences would seem to merit less weight in deciding how to treat Margo now that she has changed. After all, consider the difference between these three cases in which patients express the desire to have medical care discontinued:

Contemporaneous expression of values and preferences: A cognitively intact but ventilator-dependent woman with ALS tells her physician that she no longer wants to be kept alive in her condition. This refusal of treatment would, of course, give us very strong reasons to discontinue mechanical ventilation, since keeping her alive would require imposing an intervention on her that she does not want and that she has expressly refused.

Temporally distant expression of values and preferences, without interval change: A healthy man reads *The Diving Bell and the Butterfly* and tells his wife that he would not want to be kept alive in a locked-in state. Five years later, he suffers a basilar artery stroke resulting in a total locked-in state. If there has been no major change in his personality and cognitive ability in these five years, the man's earlier statement to his wife also gives us very strong reasons to discontinue or withhold life-sustaining treatments. For instance, since his values and preferences have not changed greatly in this time, his earlier statement provides strong evidence that he does not want such treatments now, and would refuse such treatments if he could be asked.

Temporally distant expression of values and preferences, with interval change: In Margo's case, a healthy woman (perhaps after visiting an elderly relative with Alzheimer's disease) tells her friends that she would not want to be kept alive in a demented state. She then develops Alzheimer's disease with relatively rapid progression, and within five years she exhibits the personality and cognitive changes described previously. If she later develops pneumonia, her previous statements do not give us exactly the same reasons to withhold treatment that we observed in the previous cases. For instance, given the change in her overall outlook, we cannot conclude from what she said before that she now does not want treatment with antibiotics, nor that she would refuse such treatment if she could be asked. Instead, her overall contentment with her quality of life suggests that she would want treatment in order to continue living.

More broadly, we might distinguish between two different ways of understanding claims like the brainstem stroke victim's earlier statement that he would want life-sustaining treatments to be withheld if he were locked-in. One way of understanding this is as a prediction about what he in the future, as a locked-in patient, would want. A second way of understanding this is as a statement of what he now, as a healthy person, would want regarding the future possibility of being locked-in. These are, in fact, two different claims; for instance, he is much more likely to be mistaken about the first statement than about the second. We usually do not take notice of this difference in cases like this, however, because we expect his values and preferences to remain fundamentally similar across both points in time.

In Margo's case, however, there is a much greater difference between (1) a prediction, when healthy, of what she would want in a demented state, and (2) a stated preference, as a healthy woman, regarding her future care in the event that she becomes demented. We should not interpret Margo in her healthy state to have been making claim (1), the predictive claim—in fact, precisely what she finds horrible about the prospect of dementia may be the fear that her values and preferences will be altered by disease. (After all, if she could confidently predict that she would refuse life-sustaining treatment when demented, then she would have less cause to worry about being kept alive in a demented state.) We must instead interpret her as making claim (2), a statement about her values and preferences as a healthy woman; and this would be even clearer if she were to formally execute her preference in an advance directive.

The question remains: how much weight should these values and preferences be given, if they are no longer the values and preferences of the patient? Unlike the ALS and locked-in cases, if Margo develops pneumonia when demented and is given antibiotics, we would not be imposing a treatment on someone who does not want treatment and would refuse treatment if she could be asked. Instead, we would be treating someone who previously did not want treatment for her present condition, but now presumably does. Why, then, should we honor her previously expressed values and preferences at the expense of her current ones?

The standard response to this question is that, when she previously expressed the desire not to receive life-sustaining treatment, Margo was still capable of autonomy, whereas now as a demented patient she is not. Thus, some have argued, the appropriate way to respect Margo's autonomy is to fulfill the wishes she had when she was still capable, even if we do so at the expense of her present wishes. I consider the debate over this claim in Section 34.5; however, in the next section I digress briefly to consider the related question of what sort of concern the healthy Margo should have had about her future as a patient with dementia, who would not share her current values and preferences, if personal identity is not really what matters.

34.4 Dementia and concern for one's own future

Margo, when she was healthy, did not know that she would later develop Alzheimer's disease; however, with recent research on genetics and biomarkers in dementia, many healthy people do know or will soon learn that they are at high risk of developing dementia. For instance, people with a family history of early-onset dementia with an autosomal dominant inheritance pattern (such as with amyloid precursor protein, presenilin-1 and -2, tau, and progranulin mutations) and an affected parent may conclude that they have a roughly 50% chance of developing dementia provided that they live long enough. Some people without family histories of dementia have learned that they are at increased genetic risk after undergoing testing for the apolipoprotein E4 genotype, which is available both from clinical laboratories and from direct-to-consumer services such as 23andMe. In recent years, our conception of Alzheimer's disease and other dementias has expanded to include a preclinical stage of histopathologic changes without clinical manifestations (Sperling et al, 2011); earlier this year, the amyloid-binding agent florbetapir was approved by the FDA, allowing for *in vivo* diagnosis of Alzheimer's-related neuropathology (hence, "preclinical Alzheimer's disease") even in cognitively normal people.

These new possibilities for presymptomatic diagnosis raise special ethical problems for patients and for clinicians (Karlawish, 2011). In dementia, discussion of these questions has been enriched by an extensive literature on presymptomatic genetic testing for Huntington's disease; much of this literature concerns reproductive decision-making or the risk of psychiatric harms such as depression, anxiety and suicidality (Hawkins et al, 2011). In this section, however, I consider a somewhat more abstract question: when people learn that they will develop dementia, how should this affect their attitudes about their future selves? In particular, would the fact that one's future self will be psychologically very different give one reason to care less about one's future well-being?

These issues are starker when we consider illnesses like frontotemporal dementia, which can cause radical alterations in patients' personal relationships, emotions, and overall outlook. Consider, then, a different case: Jonathan is a conscientious 30-year-old man with a family history of behavioral decline typically occurring in the fifth and sixth decades of life, often marked by alcoholism, inability to maintain stable employment, and abandonment of family and other personal obligations. His mother had exhibited such symptoms beginning in her late 40s, losing her job and developing unusual preoccupations with certain foods and with collecting trinkets. In genetic testing as part of a research protocol she was found to have a tau mutation; she was later institutionalized and died at the age of 60. Jonathan decides to undergo genetic testing because he and his wife are deliberating about whether to have children; after extensive discussion with a genetic counselor he is tested and is found to carry the same mutation as his mother.

Clearly, this knowledge will affect Jonathan and the way he thinks about his future, his wife's future, and the future of his family. How should this knowledge affect how he thinks about his future self; and in particular, his prudential concern for his own future well-being? In the future, unless there is a scientific breakthrough, Jonathan will be psychologically very different from the way he is now, even if he does remain numerically the same person. Parfit argues for the cautious conclusion that it would be reasonable for Jonathan to care less about his future well-being than someone without this mutation; however, his arguments appear to support an even stronger claim, that Jonathan *should* care less about his future well-being than someone without this mutation (McMahan, 2002). If psychologic continuity is what matters for prudential concern, and this continuity will be profoundly disrupted by disease, then it seems that Jonathan should be less prudentially concerned about his future self.

Before evaluating this claim, we should be clear about what it does not mean. It does not mean that Jonathan should care less about the fact that he will develop dementia. On the contrary, Jonathan may reasonably regard this prospect as a great tragedy. Not only will he develop a progressive and (given the limits of current therapy) incurably fatal disease, but before dying he will lose touch with relationships and commitments that are very important to him now. Second, it does not mean that he has less reason to care about the things that he currently cares about, although it will affect how his care is expressed in action. For instance, he may be less loving towards his wife in the future than he is now, but this does not give him reason to love her less now. It might give him reason to be more expressive of his love for her in the present, and to take financial or legal measures in the present to protect her in the event of his own decline.

Instead, the claim is that Jonathan has less reason to care about the well-being of his future self than he would if he did not have this prognosis, because he will be so profoundly changed from the person he is now. At first, this claim may seem implausible. For instance, suppose that Jonathan were to learn that he will also develop a rheumatologic condition that will cause him to suffer from excruciating joint pain near the end of his life. It seems strange to suggest that he should care less about the prospect of suffering this future pain simply because his values and preferences will have changed.

The claim may be more plausible when we consider other elements of well-being, besides freedom from pain. While there is considerable philosophic controversy about what makes someone better- or worse-off (Nussbaum & Sen, 1993), there is general agreement that a person's well-being depends at least in part on that person's values and preferences. For instance, if someone strongly values involvement in a religious or ethnic community, then being estranged from this community would be a great harm; whereas if someone else does not share this value, then being excluded would not matter as much (or at all) to his or her well-being. This fact can also be true of less defining interests. For instance, if someone enjoys surfing but for career reasons must move far from the ocean, having fewer opportunities to surf would make him or her worse-off, though to a less significant degree.

In Jonathan's case, he will lose some of the values and preferences that partly constitute his present well-being, so they will no longer count in the same way towards his future well-being. Currently Jonathan deeply values being a loyal friend, but his mother's experience indicates that he will lose his commitment to his friends. At the same time, he will develop new interests (in some cases describable as compulsions, though ego-syntonic) about things that he currently does not care about, such as his collecting trinkets, which will then contribute to his future well-being. Ordinarily, the fact that something would be better for you in the future would give you reasons in the present to pursue it; however, in Jonathan's case it seems that even if he knows that he will develop an interest in collecting trinkets, he doesn't have any reason in the present to collect trinkets, nor to do things that will enable him to collect trinkets in the future. Similarly, we would ordinarily think that if something is irrelevant to your future well-being, you would have less reason in the present to care about it. But in Jonathan's case, even if he knows that he will not value his friendships in the future, it seems that he has just as much reason in the present to cultivate and maintain his relationships with his friends.

Seen in this way, the changes in himself that Jonathan anticipates will break some of the usual links that hold between our future well-being and what we should care about in the present. There is also, I think, an even deeper way that dementia breaks these links between future well-being and present concern, which we can recognize by reflecting on the fact that not all changes in values and preferences are equivalent. Sometimes we welcome such changes, and even take active steps to bring them about. Consider a different case: David has no interest in opera, and in fact would find it excruciating to sit through a two-hour performance. In the present, we may suppose, his well-being would not be enhanced by spending an evening at the opera as opposed to staying at home. However, David's wife enjoys opera a great deal. David admires the way that his wife's passion for opera enriches her life, and also wishes to share in this important part of her life. David decides to cultivate an interest in opera, first by reading about different composers and different operatic styles, listening to shorter pieces at home, and asking his wife and his friends who enjoy opera to explain what they find rewarding in it. In time, David comes to enjoy opera, such that spending an evening at the opera would make him better-off. David's later, opera-loving self is psychologically changed from his earlier self; however, unlike Jonathan, these changes do not break the link between his future well-being and what he should care about in the present.

While David and Jonathan will both undergo a change in their values and preferences, an important difference is that David will be responsible for this change in himself, while Jonathan's change will be the unfortunate product of a disease. While David may not be interested in going to the opera at present, his future interest in going to the opera will be the product of his present choices. (David's case is relatively trivial, but we can also think of cases in which people decide to make more profound changes in their character—such as St. Augustine praying for conversion, or an addict who decides to reject her previous life.) The difference between these cases also suggests that Parfit's account of what matters may need to be amended. In particular, some psychologic continuities may be more important than others for prudential concern. What may be most important are those psychologic traits by which we exercise what Charles Taylor has called "responsibility for self" (Taylor, 1976) and thereby shape our characters and our values. These include memory and intention (particularly when we remember and carry out an earlier resolution), lasting commitments, and judgments about how we should live (Bratman, 2002).

David will have no reason to regret becoming interested in opera, because this change will result from his own commitment to change his future self. Thus, his future interest in opera gives him reason now, in the present, to learn about opera. On the other hand, when Margo and Jonathan contemplate their future selves with dementia, the psychologic changes that they undergo will be ones for which they are not responsible. Margo's future interest in peanut butter and jelly sandwiches and Jonathan's future interest in hoarding trinkets give them no reason, while healthy, to stock up on peanut butter or trinkets.

Even more broadly, because Margo and Jonathan cannot exercise responsibility over their future selves, they have yet another reason to care less about their future well-being. Many of the things we ordinarily do for the sake of our future well-being concern this sort of responsibility for our characters and actions: we avoid bad habits, we learn, we develop new interests, we maintain important commitments. Margo and Jonathan are helpless to influence the characters and actions of their future selves in this way—in Margo's case this may be because she will forget her present intention to change or remain some way, while in Jonathan's case this may be because his commitments will be altered. In both cases, important forms of prudential concern for the future may lose much of their purpose.

34.5 Difficulties with critical interests and precedent autonomy

Let us return now to Margo's case, and to the conflict between her previously held values and her present interest in remaining alive in her demented state. Given this conflict, which wishes should her caregivers and clinicians honor: those that Margo had when she was healthy, or those that she currently appears to have?

In a highly influential analysis, Ronald Dworkin has argued that Margo's earlier wishes, when she was healthy, should take priority over her present wishes, now that she is demented (Dworkin, 1993). Dworkin distinguishes between what he calls a person's "experiential interests" and their "critical interests." Experiential interests, as the term implies, are essentially concerned with the quality of our own experience; as examples, Dworkin includes avoiding pain, cooking or eating well, playing sports, watching movies, and even being absorbed in work that one finds challenging. In contrast, critical interests are closely tied to what we judge to be good, or worthwhile, or meaningful, regardless of the experiences that we may derive from pursuing these interests. Examples here include having close friendships (not only for the experiences that friendship brings, but also because these relationships are valued in themselves), raising children, and education for its own sake. Unlike experiential interests, these critical interests can be satisfied or frustrated independently of our experience and knowledge. As an example, a parent's dying wish might be that her children are cared for after she dies, though she will not live to know whether this interest is fulfilled.

In Margo's case, even when severely demented she still can feel pleasure and pain, and has desires that can be satisfied or frustrated. These do contribute to her well-being: Margo's life goes better if she experiences more pleasure and less pain, and when her desires are met rather than left unfulfilled. However, Dworkin argues that these only represent experiential, rather than critical interests. Critical interests, according to Dworkin, require that one be able to find meaning in one's life as a whole, which is precluded in Margo's case by her inability to connect her present to her past and future. Though Margo as a patient cannot have an opinion about her critical interests, Dworkin maintains that she still does have critical interests—specifically, she has the critical interests that she espoused when she was healthy. After all, the fulfillment of critical interests does not depend on the person's experience or awareness that they have been fulfilled, as we have seen in the case of the dying parent who will never know whether her children are cared for. Similarly, healthy people may have critical interests in living independently, maintaining their dignity, or not becoming a burden on their families. These interests can still be fulfilled or frustrated if they become demented, even if they no longer understand these interests or recognize them as their own.

Thus, Dworkin understands our dilemma in caring for Margo as a conflict between her previously-expressed critical interest in being allowed to die, and her contemporaneous experiential interest in remaining alive to continue an existence that she now finds pleasant. In ordinary cases, we hold that critical interests should trump experiential interests; for instance, that we should forego material pleasures if pursuing them would compromise our dignity or values. Dworkin argues that similar reasoning in Margo's case should lead us to honor her previous wishes and allow her to die. Indeed, because Dworkin holds that our critical interests are closely tied to our evaluation of our life as a whole, he argues that critical interests regarding the endings of our lives have special significance. The way that one dies (defiantly, or surrounded by loved ones, or hopeful for an afterlife, or attached to machines) has the potential to resonate with, or to invalidate, the meaning that one has sought to express in one's life as a whole.

Dworkin provides a second argument for giving priority to Margo's wishes when healthy, which is based upon his concept of "precedent autonomy." According to his argument, when a previously competent person loses (either temporarily or permanently) the capacity to exercise autonomy, respect for autonomy demands that we honor his or her previous, autonomously-expressed wishes, even if he or she does not share in them now. To justify this claim, he considers the difference between two cases. In one version a Jehovah's Witness declares that he would not want a blood transfusion under any circumstances, even if necessary to preserve his life. However, he suffers a serious injury and is brought to the hospital, where he finds that he wants desperately to survive and begs for a transfusion despite his earlier declaration. In this case, we should give him the transfusion, even if we know that he will regret this decision when he recovers. The second case begins the same as the first, only in this case the same injury that brings him to the hospital temporarily deranges him. In his deranged state he begs for a transfusion; but in this case, we should not give him the transfusion. Our reason cannot be that he will regret having the transfusion when he recovers, because this may be just as true in the first case. Instead, Dworkin argues, the difference is that in the first case he has the capacity to exercise autonomy, which we must respect by honoring his present preferences regardless of his earlier statements. In the second case he no longer can exercise autonomy, so though he may desire the transfusion, this deranged desire does not have the authority to override his previously-expressed, autonomous desire to avoid transfusion.

In Margo's case, Dworkin argues that she does not have the capacity to exercise autonomy, which he understands as the ability to act in ways that express one's own character and sense of self. Therefore, according to Dworkin, we can only respect Margo's autonomy by respecting her precedent autonomy, in honoring the values and preferences that she had expressed when she still could exercise this capacity.

Dworkin's arguments, while influential, have also been controversial. In particular, though the notion of precedent autonomy has been widely accepted in practice, it may seem odd on reflection. In Margo's case, can we really show respect for her autonomy by honoring a value (such as independence) that she can no longer affirm (Davis, 2002)? It might seem, also, that there is an important difference between Margo's case and the case of the temporarily deranged Jehovah's Witness. In the latter case it is plausible to say that he has not entirely lost his devotion to his religious beliefs, only that his derangement has temporarily interfered with his ability to live in accordance with these beliefs. In contrast, in Margo's case these previously held values and beliefs have been lost irrevocably, so it seems implausible to attribute them to her now in her demented state.

A deeper objection to Dworkin's argument, rooted in close observation of patients with Alzheimer's disease, has been raised by Agnieszka Jaworska (1999). Jaworska challenges Dworkin's claim that patients like Margo lack the ability to generate critical interests and to exercise autonomy, arguing instead that they possess these abilities in attenuated forms that still demand our respect. In both cases, Jaworska draws upon clinical cases of Alzheimer's disease to argue that Dworkin places excessive cognitive requirements on the exercise of these abilities.

Regarding Dworkin's distinction between critical and experiential interests, Jaworska observes that many patients with Alzheimer's disease consistently express concerns about matters that go beyond the quality of their experiences. For example, some patients with severe loss of memory and language are still aware of their decline and express regret, indicating that they still value the intellectual abilities that they once had. Patients act in ways that defend previously-held self-conceptions: someone who was previously a celebrated cook might insist on a role in preparing meals, while another patient who has always valued independence may bridle at having her activities monitored by her family. Finally, patients who participate in research sometimes express their sense that the research is *important*, or that their participation gives them the opportunity to make a meaningful contribution, even if they are unable to articulate how.

While Dworkin claims that generating critical interests requires a sense of one's life as a whole, Jaworska argues that critical interests can be rooted in more basic evaluations about what is good or worthwhile, which do not require the cognitive ability to review one's whole life. Her arguments indicate that we should understand these patients with Alzheimer's disease to be expressing true values, which are bound up with their sense of self-worth and matter to them in ways that can be independent of their experience. Jaworska also argues that Dworkin's distinction between mere experiential interests and critical interests is too dichotomous—in the case of many experiential goods such as wine connoisseurship or watching baseball, one often does have a rationale for why they are worth choosing. It might be natural for patients with Alzheimer's disease, even if previously identified with purely intellectual endeavors, to savor and appreciate simpler pursuits such as listening to music or simple art-making because these are the activities still available to them. In such cases, these experiential interests might still reflect deeper underlying values, such as the value of being connected to other people, or of leading a recognizably human existence despite one's disease.

Jaworska goes on to argue for the importance of a limited form of autonomy that these patients may still exhibit. Recall that Dworkin understands autonomy as the ability to act in ways that express one's own character and sense of self. Jaworska admits that patients with Alzheimer's disease may lack the ability to *act* effectively; as they lose their understanding of the world and become increasingly disoriented, they will no longer be able to select appropriate means to achieve the ends given by their critical interests. However, these patients may still be able to express their own character and sense of self, even if they require the assistance of others to act on them. Here again, Jaworska claims that the ability to generate values or critical interests demands our respect, even when patients lose the ability to carry out these values in action.

This analysis leaves us with the difficult question of how to handle cases in which patients make choices that are not in their own interests (for instance, when a patient must be prevented from wandering, or wastes money on extravagant purchases). Jaworska emphasizes that these patients possess autonomy in only a limited form, and that this limited form of autonomy requires a different kind of respect. While respect for the autonomy of healthy people requires us to allow them to make their own choices and decisions, patients with Alzheimer's disease may be unable to select actions that appropriately promote their values and critical interests. In these cases, caregivers and clinicians must try to interpret the values and critical interests that are expressed by the patient's wishes (however ill-formed), and find more effective ways of enabling the patient to promote them. Jaworska offers the example of a patient who resists being bathed, perhaps as a means of exercising her own agency, but who also values visits from her grandchildren. If her grandchildren begin to limit their visits because they think she smells bad and looks disheveled, then her caregivers might reasonably conclude that her own values and critical interests are best served by insisting on bathing her, thereby imposing an option that promotes her overall autonomy.

Returning to Margo, Jaworska argues that even a prior advance directive should not be regarded as binding in her case. There is an important difference between executing an advance directive in anticipation of dementia, and executing an advance directive in anticipation of permanent loss of consciousness. In the latter case, the patient no longer can generate values or critical interests, and therefore does not possess even a limited form of autonomy. However, if Margo is capable of generating contemporaneous critical interests, then these interests should not be overridden by critical interests that she previously espoused. Because Jaworska still regards Margo as autonomous, though in a mitigated way, she regards the change in Margo as analogous to other cases in which someone undergoes a change in values.

However, Jaworska's sensitive remarks on the difficulties of interpreting patients' values and critical interests do suggest a role for advance directives and previously-expressed wishes, even if we do not regard them as absolutely binding. First, while Margo herself is quite consistent in her preferences, there are other patients whose preferences are too inconsistent to be interpreted as expressing any underlying pattern of values and critical interests. In such cases, we probably should not regard them as capable of even this mitigated form of autonomy, so their prior critical interests should take priority. Second, some of Margo's contemporaneous critical interests may come in conflict with each other, and she may lack the intellectual resources to recognize this conflict and resolve it. (As in the case of the patient who resists baths but also enjoys visits from her grandchildren.) In such cases of ambiguity, we might interpret these earlier statements has having an evidentiary role, providing context about the critical interests that patients had previously espoused, which may help us to give appropriate weight to their current critical interests.

34.6 Conclusions

Dementing illnesses such as Alzheimer's disease and frontotemporal dementia strike at cognitive capacities that are central to our identities and our sense of self. As such, they raise important practical and philosophic problems, and demonstrate that for patients, families and caregivers, deep philosophic questions regarding autonomy, personality and well-being are not merely of academic concern. Because the problems raised by these illnesses can be so profound, clinicians' own thinking about these topics may be helpfully informed by a long history of rigorous philosophic debate.

Some of the most difficult ethical problems in dementia arise in intermediate stages of disease—for instance, weighing the ethical significance of contemporaneous interests that conflict with an advance directive that a patient had completed while still healthy. One reason that these intermediate stages are especially troublesome may be that many philosophic concepts such as personal identity, autonomy and moral responsibility are framed in all-or-none terms; for instance, someone living now either is or is not numerically identical with another person, and either possesses or does not possess autonomy. One theme of the present chapter is that these all-or-none assumptions easily break down in the case of dementia, which is of course a progressive degenerative condition in which the capabilities presupposed by these philosophic concepts are lost gradually. In the case of personal identity and concern for one's future, thinking about intermediate stages of dementia may lead us to Parfit's counterintuitive conclusion that personal identity is not what matters. In the case of autonomy, Jaworska's arguments recommend a special form of respect that is afforded to agents who retain the capacity for a mitigated form of autonomy, even when full-blown autonomy is absent.

One problem that may require greater philosophic attention is whether it is significant that, when someone's personality or values change in the course of dementia, these changes occur due to a pathologic process. Personalities and values are always changing over time—in some cases in response to new experiences and changed life circumstances, in other cases in response to other people's persuasion, and in many cases due to biologic processes such as normal aging. (For instance, an extensive literature on the "positivity effect" indicates that older adults devote greater cognitive resources to positive than to negative emotional information, which may reflect a greater importance placed on emotionally meaningful goals and close personal relationships (Mather & Carstensen, 2005).) Intuitively, however, there is an important moral difference between changes in one's personality and values that occur as a response to changed life circumstances, or even as a result of normal aging, and changes that are due to a neurodegenerative disease such as frontotemporal dementia. One rationale for this difference, suggested in the argument of the section on dementia and concern for one's own future, above, may be that such illnesses may sever the link between present intentions and future changes in one's personality and values, such that they no longer allow us to take responsibility for our own selves. In turn, this difference may provide a new way of characterizing the tragedy of the loss of self that accompanies dementia—not as the prospect of literally becoming a new person, as suggested in Dresser's analysis, but becoming a person for whose personality and values one can no longer take responsibility.

References

Bratman ME (2002). Hierarchy, circularity, and double reduction. In: Buss S, Overton L (Eds.), *Contours of Agency: Essays on Themes from Harry Frankfurt*. MIT Press, Cambridge, Mass., pp.65-85.

Buchanan A (1988). Advance directives and the personal identity problem. *Philos Public Aff* 17: 277-302.

Chiong W (2005). Brain death without definitions. *Hastings Center Rep* 35: 20-30.

Chiong W (2011). The self: From philosophy to cognitive neuroscience. *Neurocase* 17: 190-200.

Dresser R (1986). Life, death, and incompetent patients: Conceptual infirmities and hidden values in the law. *Arizona Law Review* 28: 373-405.

Dresser R (1995). Dworkin on dementia: elegant theory, questionable policy. *Hastings Cent Rep* 25: 32-38.

Dworkin R (1993). *Life's Dominion.* Alfred A. Knopf, New York.

Firlik AD (1991). A piece of my mind. Margo's logo. *JAMA* 265: 201.

Hawkins AK, Ho A, Hayden MR (2011). Lessons from predictive testing for Huntington disease: 25 years on. *J Med Genet* 48: 649-50.

Jaworska A (1999). Respecting the margins of agency: Alzheimer's patients and the capacity to value. *Philos Public Aff* 28: 105-138.

Kaldjian LC, Shinkunas LA, Bern-Klug M, et al (2010). Dementia, goals of care, and personhood: a study of surrogate decision makers' beliefs and values. *Am J Hosp Palliat Care* 27: 387-97.

Karlawish JH (2011). Addressing the ethical, policy, and social challenges of preclinical Alzheimer disease. *Neurology* 77: 1487-93.

Mather M, Carstensen LL. (2005). Aging and motivated cognition: the positivity effect in attention and memory. *Trends Cogn Sci* 9: 496:502.

McMahan J (2002). *The Ethics of Killing: Killing at the Margins of Life.* Oxford University Press, Oxford.

MetLife Foundation (2006). *MetLife foundation Alzheimer's survey: What America thinks.* MetLife Foundation, New York.

Miller BL, Seeley WW, Mychack P, et al (2001). Neuroanatomy of the self: evidence from patients with frontotemporal dementia. *Neurology* 57: 817-21.

Nussbaum MC, Sen A (1993), Eds. *The Quality of Life.* Clarendon Press, Oxford.

Parfit D (1984). *Reasons and Persons*. Oxford University Press, Oxford.

Sartori G, Snitz BE, Sorcinelli L, et al (2004). Remote memory in advanced Alzheimer's disease. *Arch Clin Neuropsychol* 19: 779-89.

Sperling RA, Aisen PS, Beckett LA, et al (2011). Toward defining the preclinical stages of Alzheimer's disease: Recommendations from the National Institute on Aging and the Alzheimer's Association workgroup. *Alzheimers Dement* 7:280-92.

Taylor C (1976). Responsibility for self. In: Rorty AO (Eds.), *The Identities of Persons*. University of California, Berkeley, pp.281-99.