Journal of Ethics and Social Philosophy
vol. 27, no. 1 · march 2024
https://doi.org/10.26556/jesp.v27i1.2843

REMAINING TRUE TO OURSELVES

DEMENTIA, VALUE CHANGE, AND ENDURING INTERESTS

Andrew Franklin-Hall

It is common to think that when we need to make choices in the best interests of others, we should be guided (at least in part) by their values.[1] But the later stages of dementia may cause people to lose a grip on some of the things that used to be important to them. Should we then consider their former values, even though we do not ordinarily consider what people *used* to care about in trying to discern their best interests? Or should we simply be guided by the person's perspective as it is now (and as it will be in the future)? Consider the following cases:

> *Mr. Amato*: For the last two years, Mr. Amato, who has Alzheimer's disease, has been living in a long-term care facility. He has been reasonably content there, making friends with other residents and enjoying group activities. However, it has recently come to the attention of his children, John and Sally, that their father has been walking around common areas half-disrobed and that he is seldom afforded any real privacy in his room. When they asked the administrators about this, they were told that since Mr. Amato does not seem to mind these things, it is easiest not to worry about them. John and Sally remain disturbed, though, because they know that their father used to consider propriety and privacy as integral to living in a dignified way.

> *Ms. Bell*: Although Ms. Bell has never received a formal diagnosis, she is plainly living with relatively advanced dementia. Nevertheless, with the help of home health aides, she has been able to remain in her house. Recently, her daughter, Claire, noticed certain family heirlooms missing from her mother's shelves. Suspecting the caregivers of theft at first, she eventually discovered that her mother had been giving these things away

1 See, for example, Buchanan and Brock, *Deciding for Others,* 29–36.

> to neighbors and friends (and taking real joy in doing so). Claire had mixed feelings about this. She did not relish inheriting all those objects, but she also knew that it used to mean a lot to her mother to keep them in the family. When Claire tries to remind her mother of this, Ms. Bell tends to make a joke and change the subject.

It is not my aim to prescribe exactly what we ought to do in cases like these. Such decisions will call for considerable judgment, empathy, and creative problem-solving. My concern is with the more fundamental question as to whether the former values of people in circumstances like those just described have any bearing on what is in their best interests—or, put differently, is prudentially valuable for them or contributes to their "good," "welfare," or "well-being." These terms have different shades of meaning in ordinary language, but like many philosophers today, I use them synonymously. I do not know how to define well-being in any more basic terms, but we do seem to use some such notion in thinking about a web of related concepts, like benefit, advantage, harm, and enlightened self-interest. Crucially, as most philosophers today understand the notion, it is not an analytic truth that well-being is equivalent to subjective quality of life. For example, while there is disagreement about whether things external to a person's consciousness can affect her well-being, most philosophers regard that as a coherent claim.

The value of autonomy tends to loom large in discussions about decision-making for patients lacking decision-making competence. Insofar as autonomy or self-determination is conceived of as one thing that is good for a person, it is part of my topic.[2] However, there is also a purely deontic notion of personal sovereignty over one's life that, to keep the discussion manageable, I will not address. Likewise, I set aside here obviously relevant questions about how these decisions might affect the interests of others. This means that I am not offering an account of what we should do, all things considered, in the cases under consideration.

Two approaches to our problem have been especially prominent in the philosophical literature on dementia. On what we might call the *Presentist Model,* what a person used to value simply cannot bear on her current interests. To have an interest in something, they think, one must at least remain capable of caring about it or taking an interest in it. Thus, Rebecca Dresser holds that

> individual incompetent patients' interests are invariably a function of their physical and mental capacities.... A truly patient-centered best interests assessment will incorporate an examination of the particular incompetent patient's interests in light of his individual capacities....

2 See, for example, Griffin, *Well-Being,* ch. 4; and Raz, *The Morality of Freedom,* ch. 14.

> Matters such as dignity, privacy, and bodily integrity arguably are integral to the well-being of the average or reasonable competent person in our culture. But it is nonsense to claim that these matters affect the well-being of many incompetent patients with severely compromised mental abilities.[3]

An alternative view says that the capacity to value, or care, or find meaning in things bestows a certain standing on a person's present perspective, and this makes it inappropriate to appeal to her former values *until* the person no longer has this capacity. At that point, however, a person's former values can matter. I will refer to this as the *Status Model*.

I side with the Status Model in thinking that there are conditions in which a person's past values can bear on his current interests. But I deny that we need to attribute exclusive authority to either the person's present perspective or his past one. Instead, I defend a more piecemeal approach. A person's past values can matter even while he continues to value or care about simpler things now. The critical question, as I see it, is whether a person is answerable for the fact that he no longer holds a certain value that was once dear to him. For that, he must still be able to understand what he used to value and why—he must, in other words, retain the capacity to critically reflect on that past value. If he cannot, we should not think of him as having revised his former value. Rather, that value remains imputable to him as his last authentic verdict about the matter. Call this the *Revision Model*. To defend this view, I will begin with a discussion of a few influential versions of the Status Model, then outline the Revision Model, and finally explain why I think it is more plausible than the Presentist Model.

Often this issue is discussed within medical ethics as a part of a conversation about advance directives, substitute decision-making, and end-of-life care. This makes sense in that the stakes are so high in those cases. But for that same reason, I believe those cases can have a distorting effect on our analysis, for they tend to be represented in stark either/or terms, and they involve questions of life and death, which may raise additional moral issues and pertinent questions of public policy. For this reason, they may distract us from the fundamental question as to whether a person's former values can still have some bearing on his interests. That is why I will be focusing on more mundane cases, like those of Mr. Amato and Ms. Bell. My conclusions should be relevant to the more familiar debates in medical ethics, but they will not settle them.

Finally, I should acknowledge that some have thought that the deterioration of memory and personality can be so profound that the dementia patient is no longer the same person, metaphysically, as she used to be. In that case, it would

3 Dresser, "Life, Death, and Incompetent Patients," 383–85.

seem wrong to impose the former person's values on the present individual.[4] This relies on a controversial theory of personal identity, but pursuing that issue would take us too far afield.[5] I will only note that, even accepting that view, there will be cases where someone with dementia has lost hold of a former value, and yet enough of his memory and personality remain that it would be hard to deny this is still the same person. So, appeal to personal identity will not entirely dispose of the puzzle.[6]

1. Well-Being and the Status Model

I shall take, as my point of departure, the influential views of Ronald Dworkin and Agnieszka Jaworska.[7] Although their accounts differ in their practical implications, they both defend versions of what I am calling the Status Model.[8] I begin with Dworkin's ideas about the objective and subjective dimensions of our so-called critical interests, then outline the general structure of the Status Model, and finally raise some difficulties for it.

1.1. The Objective and Subjective Aspects of Well-Being

According to Dworkin, although we all have "experiential interests," which involve having agreeable experiences and avoiding disagreeable ones, our more important decisions in life are typically guided by our sense of our "critical interests." These critical interests have both an objective and subjective dimension, according to Dworkin.[9] From the agent's point of view, critical interests look largely objective. Most people, he says, take themselves to have "interests that … make their life genuinely better to satisfy, interests they would be

4 Dresser, "Life, Death, and Incompetent Patients" and "Dworkin on Dementia." For criticism of this idea, see Buchanan and Brock, *Deciding for Others*, ch. 3; and DeGrazia, *Human Identity and Bioethics*, ch. 5.

5 Alternative views are defended in McMahan, *The Ethics of Killing;* and DeGrazia, *Human Identity and Bioethics*.

6 Nor do I discuss "time-relative interests" (see McMahan, *The Ethics of Killing*). That past values can matter is consistent with thinking their importance should be discounted the less psychologically connected the person is to his past self.

7 Again, I focus on arguments appealing to well-being, not autonomy.

8 Called the "threshold of authority approach" in Jaworska's entry on "Advance Directives and Substitute Decision-Making" in the *Stanford Encyclopedia of Philosophy*.

9 Dworkin, *Life's Dominion*, 206. My interpretation of critical interests is informed not only by Dworkin's discussion of dementia and end of life in *Life's Dominion*, but also by his more systematic treatment in "Foundations of Liberal Equality," to which we are referred in an endnote in *Life's Dominion* (255n21). (The relevant sections of "Foundations" are also reprinted, with minor amendments, as chapter 6 of *Sovereign Virtue*.)

mistaken, and genuinely worse off, if they did not recognize."[10] Hence, our critical interests involve the kinds of things we commonly find on "objective list" theories of well-being: instances of knowledge, achievement, living up to worthy ideals, maintaining valuable relationships, and so on.[11] These are things that we conceive of ourselves as wanting because they are valuable—things we admire in the lives of others and that we can reflect on in our own lives with pride and satisfaction.

But critical interests also have a subjective side, especially from the observer's point of view. Part of Dworkin's idea is that what makes a person's life go better is the way something fits in with the distinctive character of that particular life.[12] This tends to privilege the values that a person has been living by to this point. But, for a couple of reasons, Dworkin also thinks of a person's current perspective as ordinarily bearing preeminent authority in defining his current critical interests.

First, Dworkin subscribes to the "endorsement thesis": that nothing can noninstrumentally contribute to our well-being (in the critical sense) unless we endorse its value at some level.[13] "Value," as he says, "cannot be poured into a life from the outside; it must be generated by the person whose life it is."[14] This leads him to assert that a person's critical interests are partly constituted by his convictions.[15] The endorsement thesis does not imply that a person will benefit from something *merely* because he endorses it; the object must actually be valuable. But, presumably, most of what we value is not entirely worthless. So, my life might have been better if I valued family more than achievement, but if I only appreciate achievement to any significant degree, then (holding

10 Dworkin, *Life's Dominion*: 201–2. Compare Dworkin, "Foundations of Liberal Equality," 230, and *Sovereign Virtue*, 245.

11 Dworkin, *Life's Dominion,* 201–2. On the objective-list theory, see Parfit, *Reasons and Persons*, 499–501.

12 Dworkin, *Life's Dominion,* 202, 205–6. Compare Dworkin, "Foundations of Liberal Equality," 249–53, and *Sovereign Virtue*, 257–60.

13 Dworkin, "Foundations of Liberal Equality," 237, 264, and *Sovereign Virtue*, 248–49, 268. In *Life's Dominion,* Dworkin writes: "It is important both that we find a life *good* and that we *find* it good" (206).

14 Dworkin, *Life's Dominion*, 230.

15 To avoid circularity, Dworkin probably should have distinguished between what would be a critical interest for me if I endorsed it ("presumptive critical interests") and what is good for me given what I actually endorse ("confirmed critical interests"). Our convictions are partly constitutive of the latter, not the former. When I deliberate about my own life, I mainly focus on my presumptive critical interests, but benefactors need to take my confirmed critical interests into account.

my values fixed) what will actually contribute to my critical interests is achievement, not family.

The other idea is that there is independent worth in my figuring out for myself what is right for my life and putting that vision into effect. This is at least part of what Dworkin means by "integrity."[16] The good of living with integrity lies not in the content of my values (which may be misguided) but in the way that I am faithful to my understanding of what is valuable.[17]

I am sympathetic to this general picture. I think we do typically assume an objective stance when we reason about how to live.[18] And yet, it also seems that a person's interests (at least from the bystander's perspective) are partly defined by her own values and convictions. Indeed, I am inclined to think that Dworkin is basically right that a person cannot noninstrumentally benefit from a good (at least not substantially) unless it finds some purchase within his evaluative perspective and that there is significant value in living a life according to one's own convictions. But what is most important to my argument going forward is that Dworkin is right to think that our more sophisticated interests have both an objective and subjective side.

1.2. The Status Model

If a person's well-being partly depends on his own ideas about the good life, then we must consider those when trying to promote his interests. But what if a person does not have such views anymore? Dworkin maintains that when a person loses his "sense of life as a whole" and thus no longer has a coherent view of what living a good life consists in, we should look back to the views he formerly held.[19]

16 "Someone has ethical integrity . . . when he lives out of the conviction that . . . no other life he might live would be a plainly better response to the parameters of his ethical situation rightly judged (Dworkin, *Sovereign Virtue*, 270; compare "Foundations of Liberal Equality," 267). In *Life's Dominion*, Dworkin adds the diachronic idea that a life ought to "display a steady, self-defining commitment to a vision of character or achievement that the life as a whole, seen as an integral creative narrative illustrates and expresses (205). I think Dworkin's considered view is that my life has integrity when I live consistently with my current view about what is fitting for my life, given the way I have lived it to this point.

17 In *Life's Dominion*, Dworkin ties the right to autonomy to its protection of the agent's *capacity* for integrity, whether integrity is actually achieved or not (224). But the present claim is that *achieving* integrity actually makes a person's life go better. In fact, Dworkin says that "a life that never achieves . . . integrity cannot be critically better for someone to lead than a life that does" (*Sovereign Virtue*, 270; compare "Foundations of Liberal Equality," 267). See also the remarks in *Life's Dominion* about the value of a life partly consisting in its achieving integrity (206, 224).

18 I take no position on the metaethics of this objective stance.

19 Dworkin, *Life's Dominion*, 230.

To see the structure of the position, let us say that an attitude has *prudential authority* if something is noninstrumentally good for the person, at least in part because the attitude favors that thing. So, according to a simple desire-satisfaction theory, a person's desires have prudential authority. On Dworkin's view, the attitude with prudential authority is a certain kind of judgment about what makes one's life good or successful. By contrast, purely objective theories of well-being deny that any subjective attitude has prudential authority.

Dworkin's suggestion is that, as long as a person continues to express attitudes possessing prudential authority, the person's current perspective has the exclusive standing to pick out what is valuable for her life (insofar as this depends on her subjective attitudes). This is because, as Dworkin says, such a person retains "the ability to act out of a genuine preference or character or conviction or sense of self," and we have an interest in living in accordance with our convictions.[20] Call this the *Status Model*. Expressing attitudes with prudential authority confers on a person's present outlook a standing or status that makes it inappropriate to appeal to that person's former perspective in working out what is best for her.[21] Plainly, you could accept the Status Model but differ with Dworkin over which attitudes possess prudential authority.[22] After all, to say that one needs convictions about what is good for one's life as a whole is a demanding view. If simpler attitudes could bear prudential authority, then the Status Model would imply that a person's present outlook could retain the exclusive standing to define what is best for her until a later stage of cognitive deterioration.

This is the key move Jaworska makes. In her 1999 article, she argues that it is a person's values that possess prudential authority. To value something, she says, just involves being able to give some account of what makes it good or worthwhile and some recognition that such judgments are, in principle, open to revision in the light of better reasons. Valuing, however, need not involve global conceptions about what makes a life successful; it can just consist in local judgments about what seems right for us, here and now. But these local values still inform our sense of ourselves as agents answerable to certain standards and typically affect our sense of pride or shame. Thus, even simple values enable a person to act (as Dworkin might say) from a "sense of self."[23] The implication is that many people living with mid-stage Alzheimer's disease remain valuers, even if they are not capable of the kind of whole-life assessment that Dworkin

20 Dworkin, *Life's Dominion*, 225.

21 This resembles the concept of "full moral standing" in Jaworska, "Caring and Full Moral Standing."

22 You could also demur from Dworkin's commitment to the objectivity of our interests.

23 For independent confirmation of this claim, see the conversations transcribed in Sabat, *The Experience of Alzheimer's Disease*.

privileges. Someone with this kind of impairment should, Jaworska thinks, "be viewed as any other person whose values and commitments change over time and whose currently professed values are taken to bear on what is best for her."[24] Only when a person has lost the ability to value in this modest way should we resort to her former values.

In subsequent work, Jaworska sets the bar even lower.[25] The right threshold for exclusive prudential standing, she claims, is the capacity to *care*, where this involves a significantly intense, sustained, and cohesive pattern of emotional vulnerability and attentiveness to something, which construes the object of concern as important to the agent. Although valuing something typically involves caring about it, caring is a simpler attitude since it need not involve judgments about what is genuinely "good" or "right," nor need it be associated with self-reflective attitudes (like pride or shame). But it is this capacity for sustained emotional attunement that, according to Jaworska, underlies the possibility of a cohesive and enduring self. Therefore, only a self of this sort can have attitudes with the kind of prudential authority that expresses a "sense of self." It is only once a person loses this more basic capacity to care that we should consult his former values since only at that point is there no longer a "deep self" present.[26] For instance, even if Ms. Bell has lost some of her past values, she evidently continues to value or care about making others happy, and that means that her current perspective retains the exclusive authority in defining her interests.

Jaworska makes a strong case that the attitudes with prudential authority need not be as sophisticated as Dworkin suggests. In fact, you might think that even Jaworska's view is too demanding. If the intuition behind the endorsement thesis is that I cannot directly benefit from an objective good unless it resonates with me, that it cannot be alien to me, then why is it not enough that I appreciate the good in the moment? Why does this receptivity have to be a stable disposition (like valuing or caring)? Perhaps the idea is that critical interests are supposed to make our *lives* better, and for this, they must structure some significant portion of our lives. Perhaps only temporally extended attitudes can support the kind of investment in things that makes this possible. If that is how we define critical interests, though, it seems we should acknowledge that we also have more episodic ways of benefiting from objective goods that, nonetheless, are not classifiable as mere experiential interests.

24 Jaworska, "Respecting the Margins of Agency," 112.

25 Jaworska, "Caring and Full Moral Standing." See also Shiffrin, "Autonomy, Beneficence, and the Permanently Demented," 211.

26 Should we, at that point, appeal to the values the person held prior to impairment or to the last set of genuine concerns held while impaired? The former view is more intuitive, but the latter seems more consistent with the logic of the position.

1.3. Problems with the Status Model

That said, I think that some difficulties arise for even the best version of the Status Model. First, although Dworkin and Jaworska tell us *when* they think we ought to resort to a person's former attitudes, they do not really explain why our former values matter in just this way. If a person's past attitudes have any validity, why should they not normally be counted—even for people without impairments?[27] On the other hand, if they do not normally count, then why suppose that they *can* play this reserve role of coming forward just when the person loses the ability to express attitudes with prudential authority?

Second, suppose that, following Jaworska, we adopt an expansive view of the attitudes that bear prudential authority. Perhaps Ms. Henderson had long valued tithing to her church and intended to continue to give for the remainder of her life. Now, however, her dementia has deprived her of any understanding of that former commitment. The Status Approach says that whether Ms. Henderson still has an interest rooted in that former value depends on whether she still has prudentially authoritative values or concerns for anything else. But why should the moral relevance of this former commitment depend on whether the subject continues to care about other, unrelated matters? Indeed, what if making those donations on behalf of Ms. Henderson would not interfere with anything that presently matters to her? It seems like whether a former value constitutes an enduring interest should turn on the person's relation to that very value.[28]

To see the third issue, return to the case of Ms. Bell. Suppose that, at age eighty-five, she ceased to care about passing on family heirlooms to her children. On Jaworska's caring version of the Status Model, since Ms. Bell continues to care about other things, that former value is irrelevant to her best interests. But suppose that, by the time she turns eighty-eight, she really cannot be said to "care" about anything (in Jaworska's sense). At that point, apparently, she does have an interest in passing on the heirlooms again. But it seems strange to think that the relevance of this value to Ms. Bell's interests should fade away and then return in this manner. Stranger still, when Ms. Bell is eighty-five, we can predict that, while passing on the heirlooms is not currently good for her, it will be good for her later (if she lives long enough). But if acting prudently for Ms. Bell

27 Dworkin might appeal to the importance of a life remaining consistent with the themes that have subsequently structured it, but then we have grounds for overriding a fully competent person's decision to keep him faithful to the sort of life he has already been leading (see Hawkins, "Well-Being, Time, and Dementia," 525–26).

28 A slightly different view would be that past values can matter as long as they do not directly conflict with a person's present values or more robust concerns. I would also reject this approach, however, because I doubt that the slightest current value always trumps the weightiest former value.

means taking her current and (probable) future interests into account, should we not already (while she is eighty-five) think she has a future-oriented interest in passing on the heirlooms? It seems more plausible to either say that this value has always mattered during the period of Ms. Bell's impairment or that it never matters. In this light, Jaworska's version of the Status Model starts to look like an uneasy compromise between those options. These problems do not arise so acutely on Dworkin's version of the Status Model since it is quicker to resort to a person's former values. But this advantage is bought at the unacceptably high price of ignoring the critical interests rooted in Ms. Bell's present perspective.

2. THE REVISION MODEL I: THE PRIORITY OF VALUING

2.1. The Revision Model: An Overview

I call the view I defend the *Revision Model*, though it might more properly be labeled the "Authentic" or "Authoritative Revision Model." Here is the basic idea. Although a range of subjective attitudes of varying complexity can bear prudential authority, a person's values have priority in this respect since these represent the agent's own considered—and thus authentic or authoritative—perspective about his interests. Ordinarily, we should attribute the highest prudential authority to a person's *current* values (not his former ones) because these represent his most up-to-date revision of his earlier views. However, if a person has lost the ability to comprehend some of his former values, then we cannot treat his current perspective as a genuine revision since he is not answerable for the fact that his values have changed in this respect. Having never truly revised those former values, we should recognize their enduring authority in defining his interests insofar as they remain relevant. However, this is consistent with recognizing that the person may, at the same time, have other interests rooted in his current perspective. In this section, I make the case for the priority of values and emphasize the rational element in value revision. In section 3, I address the authority of a person's values at different points in time. And in section 4, I consider the relative importance of different sorts of past values for a dementia patient's current interests, as well as the significance of conflicts between a person's present and past perspectives.

2.2. Prudential Authority and Agential Authority

Any view recognizing the prudential authority of subjective attitudes must explain how to handle the relation between a person's various concurrent attitudes, especially when these conflict with one another. One approach would be to assign weights to attitudes in proportion to their phenomenal intensity. But this is not how we reflect on our own feelings and desires. That I feel a strong urge to do something does not at all settle for me, on reflection, that this desire is important to satisfy, for I might entirely repudiate some of my feelings and desires. If we simply attribute authority to desires or other attitudes based on their felt intensity, then we are treating the person not as an active agent but as a mere site of competing emotions and motivations. To take agency seriously, we must consider the person's own view about the relative importance of her various attitudes (assuming she has such a view). But to do this, we need to characterize the perspective that counts as the agent's authentic standpoint: the point of view from which she authoritatively identifies with some of her attitudes, rejects others, and settles on what weight to attribute to her various priorities. And, if we do not want to posit a little pilot inside the head who is the "true self," we need to identify a person's authentic perspective with certain attitudes, or configurations of attitudes, under certain conditions. Let us say that the attitudes that comprise a person's authentic perspective possess "agential authority."[29] My first claim, then, is that prudential authority follows agential authority. That is, the attitudes that constitute a person's authentic perspective as an agent have the greatest authority in defining what the person's true interests are.[30]

At a verbal level, many philosophers would agree that, for the mature, unimpaired agent, this authentic perspective is framed by what he, on reflection, and in a calm and collected state, values.[31] Valuing, here, should be understood in a dispositional way. I do not lose my values when I am asleep or when they are not present before my mind. To speak of "a collected state" means we are especially interested in what he values when he can reflect on how his different interests relate and compare. So, this suggests that Dworkin is on the right track in thinking our more global values tend to have a certain priority for those of us who have them. The main point of controversy lies not in whether values are important but in how we think about the kind of attitude valuing is.

29 See, for instance, Bratman, "Reflection, Planning, and Temporally Extended Agency" and "A Desire of One's Own." The kind of "authentic perspective" or "agential authority" that Bratman identifies emerges from the influential work of Frankfurt ("Freedom of the Will and the Concept of a Person") and Watson ("Free Agency"). Similarly, Brudney and Lantos speak of living in accordance with one's commitments as the value of "authenticity" ("Agency and Authenticity").

30 Compare Raibley, "Well-Being and the Priority of Values"; and Dorsey, *A Theory of Prudence*, ch. 5.

31 There may be other conditions, such as freedom from manipulation or oppressive socialization, but I leave that aside.

Before expanding on that, though, let me clarify: I do *not* say that nothing can be good for us unless we value it. There are, as I have indicated, simpler attitudes that can open us up to the good things in life and support our investment in them. People with impairments may not be capable of the kind of reflection involved in valuing. And none of us reflects on all our attitudes. What is special about valuing is that it can overrule and constrain the presumptive authority of our other attitudes. I might find myself caring a lot about recognition for my accomplishments, and yet, on consideration, decide that I have been much too focused on this. In this way, my values represent my considered view, insofar as I have one, about what is important in my life and how important it is.[32] So, to the extent that my subjective attitudes have some bearing on what my interests are, my values simply have the highest authority. That said, normal adults typically *have* reflected to some degree on the central ideals, projects, activities, and relationships that structure their lives. Hence, for them, their key interests usually *are* defined by their values since these represent their own perspectives about what is important in their lives.[33]

2.3. The Nature of Valuing

Although many would agree that the agent's authentic perspective is defined by his values, there are different views about what this attitude consists in. The accounts we get from Dworkin and Jaworska are broadly cognitivist in spirit in that judgments of value are conceived to be central to the nature of valuing. Dworkin says that when we are deliberating about how to live, we are trying to form a judgment about "what is really important in life," which is something we think we can be mistaken about.[34] And Jaworska holds that the chief distinction between valuing and mere desiring is this:

> We think it would be a mistake to lose our current values—we hold our values to be correct, or at least correct for us. And this means that we can typically give a rationale for why we consider something valuable or good, usually by situating this value in a larger normative framework.

32 Thus, I disagree with Jaworska when she claims that whatever a person cares about is "internal" to the person—representing (part of) her authentic or true perspective ("Caring and Internality"). Although I agree that concerns can possess agential authority without being reflectively endorsed, I do not think we should say that concerns have agential authority when they are consciously disavowed. For instance, I would not attribute agential authority to ingrained social mores that a person no longer endorses but cannot shake at an emotional level.

33 Compare Tiberius, *Well-Being as Value Fulfillment,* 11–13; Dorsey, *A Theory of Prudence*, ch. 5.

34 Dworkin, *Life's Dominion,* 206.

> Also, since values are the sorts of attitudes that we allow could be correct or incorrect, they are open to criticism and revision.[35]

I think this is basically the right approach.

The main disagreement between Dworkin and Jaworska, recall, is whether we should privilege a person's more global values. On this point, both writers have some piece of the truth. If we are asking which attitudes can possibly possess the sort of prudential authority associated with critical interests, then Jaworska is correct to say that a person's values need not be global. On the other hand, if we are thinking about a person who has a range of values, then his more considered values have more authority. And often our more global values *are* our more considered ones; they represent what we think is most important for us once we have tried to put everything into its proper perspective. This is not invariably true, however. A person might have certain idle notions about the good life, into which he has put nowhere near as much thought as his more immediate values. For this person, his superficial-but-global ideas about the good life should be attributed less prudential authority than the more local values that he takes more seriously.[36]

Some writers agree that valuing is central to both agency and defining a person's interests, but they deny that value judgment is central to the attitude. One might, for example, think that valuing is really a matter of having higher-order desires to maintain and act on certain more immediate desires.[37] I agree that our values often come into view when we reflect on whether we are really "behind" a particular feeling or desire. But the questions the agent is typically asking in such cases are normative or evaluative questions: "Is this really what I ought to be doing?" "Is this a good thing to care about?" "Are my feelings appropriate here?" and so on. Indeed, if these are *not* normative questions, it is hard to see why the mere fact that higher-order desires involve reflections on

35 Jaworska, "Respecting the Margins of Agency," 115.

36 Even if my values possess agential authority for me, they cannot all bear on my well-being (unless I am exceptionally self-centered). Most of us also value things for their own sakes and for other people. So how do we distinguish the welfare-relevant values? This is a version of the "scope problem" that also affects desire-satisfaction theories. Trying to resolve this issue is too large a task to undertake here. But, to mention a few possible responses, one might think that my welfare-relevant attitudes are those that (a) refer to my own life or necessarily presuppose my own existence (Parfit, *Reasons and Persons*, 494–95; Overvold, "Self-Interest and Getting What You Want"), (b) that I conceive of as involving my own good (Dorsey, *A Theory of Prudence*, ch. 5), (c) that are "warm" and "affective" (Heathwood, "Which Desires Are Relevant to Well-Being?"), or (d) that refer to the goods that are relevant to well-being on the correct objective theory.

37 Frankfurt, "Freedom of the Will and the Concept of a Person."

other desires should give them any special authority.[38] This suggests that the perspective of the agent is (at least typically) defined or informed by his view about what, in his circumstances, is good or what he has most reason to do.[39]

Perhaps the most common objection to the cognitivist approach to valuing is that our judgments of value underdetermine what we personally value. As Samuel Scheffler says, "There are . . . many activities that I regard as valuable but which I myself do not value, including, say, folk dancing, bird-watching, and studying Bulgarian history."[40] But the kind of value judgment that is most relevant to the cognitive approach is not a belief about what might be worthwhile for someone, but what is worthwhile for me, given my particular character and circumstances. This kind of judgment will naturally take into account my particular abilities, my tastes, and my existing emotional attachments. But none of these facts about me constitute what I value since I might repudiate any of them or at least downgrade their importance. What I value—in the present sense—is what I think is good, fitting, or worthwhile in my circumstances, taking all of this into consideration.[41]

Might there not still be cases where the reasons are inconclusive? Perhaps I have two very different courses of life open to me, and I come to identify with one even though I have as much reason to favor the other. Or maybe the question is simply how much weight I ought to attribute to each priority. Do such cases not suggest there is a notion of valuing or identification that goes beyond the reasons we have?[42] This is an important theoretical question, but we need not settle it for our purposes. First, even if we grant that valuing can outstrip our value judgments in cases of rational underdetermination, it remains the case that our judgments about what is good or worthwhile play the central role in informing our values. Generally, a person's values do align with her view about what is best or fitting for her life. Second, even if reason is sometimes inconclusive about which values we ought to initially adopt, it is also the case that once our lives have been shaped by one set of values, this changes our circumstances and, thus, also the facts about what is really best for us.[43] For instance, even if there were initially no conclusive reasons for me to personally value philosophy more than history, the fact that my life has been shaped by valuing philosophy

38 Watson, "Free Agency."

39 Wallace, "Caring, Reflexivity, and the Structure of Volition,"; see also Bratman, "A Desire of One's Own."

40 Scheffler, "Valuing," 21.

41 Wallace, "Caring, Reflexivity, and the Structure of Volition."

42 Raz, "Incommensurability and Agency"; Bratman, "A Desire of One's Own."

43 Raz, *The Morality of Freedom*, ch. 14.

for many years generates new reasons to continue valuing it. What is important for the argument going forward is simply that a person is typically unwarranted in revising his values—particularly long-held values regarding central elements of his life—without good reason.

3. THE REVISION MODEL II: THE AUTHORITY OF PAST VALUES

I have argued that, among the subjective attitudes possessing prudential authority, a person's values are the most important. Further, I have claimed that valuing involves beliefs about what is good, fitting, or worthwhile for one's life and that changing one's mind in a justifiable way typically involves some kind of reconsideration or responsiveness to those reasons. Let us now consider how the prudential authority of values works over time and how this relates to cases involving dementia.

3.1. Prudential Authority over Time

The puzzle about the relevance of past attitudes is commonly discussed in connection to desire-satisfaction theories. It might seem that, on such views, the essence of prudence is to accord as much weight to the fulfillment of future desires as present ones. But if desires matter equally whenever they are held, should I not also try to fulfill my past desires?[44] For instance, if we imagine a six-year-old who decided he would like to go on a roller-coaster on his fiftieth birthday, would the person, now fifty, have prudential reason to get on the ride, even though he really does not want to do so now?[45] If past desires continued to matter, then just as we often make present sacrifices for the sake of what we will desire in the future, we would have to accustom ourselves to making sacrifices now for what we used to want. We would also have more reason to fulfill desires we had entertained for longer. This means it would be imprudent not to keep track of how long you had wanted something. Some find these implications very counterintuitive and infer that the desire-satisfaction theory of well-being is untenable. True, some of the problems dissipate once we remember that desires are usually conditional on certain factual assumptions (which may change), or on their own persistence, or on getting some enjoyment from their fulfillment.[46] But we can at least imagine cases that do not involve those conditions. For instance, Derek Parfit tells us that, when young, he strongly wanted to be a poet one day. This desire, he assures us, was not conditional in

44 Brandt, *The Good and the Right*, 249–50; Parfit, *Reasons and Persons,* 150–53.

45 Brandt, *The Good and the Right*, 249.

46 Parfit, *Reasons and Persons*, 151; Bykvist, "The Moral Relevance of Past Preferences."

any of the above ways. "Does my past desire," he asks incredulously, "give me a reason to try to write poems now, though I now have no desire to do so?"[47]

But what happens when we shift our focus away from mere desires, to a person's values? As Parfit observes, when desires involve value judgments, things look different from the first-person perspective. If your fundamental values change, it seems you must take yourself to have an improved view of the reasons that apply to you. In other words, you are committed to believing that your current values are more justified than your former ones. So, just as you do not think your former, now discarded beliefs have some enduring validity, likewise your former values seem to have no prudential weight from your current perspective.[48]

Things are different, however, from the observer's point of view. Since the observer does not necessarily share either the agent's past or present perspectives, he is not committed to supposing that the agent's present values have more warrant than the old ones. So, why should he recognize the agent's present values as possessing special prudential authority? The answer, I suggest, is this. We have already seen that a person's values have higher authority in defining what a person's good consists in. But if we are going to take a person's values seriously *as values*—as beliefs about what is best for her—we have to accept her current perspective as a revision of her former one. If we simply toss her former views on the scales with her current ones, we will not be treating these as genuine values at all. Without the possibility of value revision, the normativity of valuing is undercut. But, as I argued above, it is that normativity that lends values their special authority to speak for the agent.

You may object to this reasoning that, just as a person's current values involve a commitment to rejecting her past values, a person's past values involve a like commitment to rejecting her incompatible future values. Why should the observer treat the agent's current retrospective rejection of her past values any differently than the agent's former prospective rejection of her current values?

Well, you might think that the observer's default assumption should be that an agent's later values are better supported by reasons than her earlier values

47 Parfit, *Reasons and Persons*, 157. Some writers say that satisfying past desires *does* promote well-being (e.g., Baber, "Ex Ante and Post Hoc Satisfaction"; Sarch, "Desire Satisfactionism and Time"; Dorsey, *A Theory of Prudence*, ch. 8). But suppose a person, now permanently unconscious, long wanted his life extended no matter what, but had a change of heart last year. These views surprisingly say that (all else equal) the older, longer-held desires count for more than his more recent desires.

48 What about values I predict I will hold in the future? If I expect I will later be better situated to make an accurate value judgment than I am now, I should bring myself to value that thing presently. If I expect my future judgment to be worse, I should disregard it.

because she has had more opportunity to gather evidence and experience. This may be true, but it seems like this presumption will be undermined whenever there is independent reason to believe that the person's judgment has worsened. If this were the only reason to favor a person's present perspective, then bystanders *should* sometimes ignore a person's present values in favor of her former, more reasonable ones. This would not be so very surprising if we were just talking about temporary judgment shifts due to weakness of will. But, here, our focus is on stable changes in a competent agent's values.

A better explanation of why a person's present values should count from the third-person perspective concerns the temporal nature of agency. We care about a person's values because we think it is important to know how a person has made up her mind. But making up one's mind is something done in time, taking off from previously held conclusions or unreflective attitudes. My claim is that you just cannot pay due regard to what a person *does* in making up her mind about her values unless you see her later judgments as supplanting her earlier ones. To look at values in a tenseless way obscures the person's active role in forming and revising them.

3.2. Answerability for Value Change

Appreciating why we should generally recognize a person's present values as authoritative throws into relief the exceptional cases in which it makes sense to continue to recognize the authority of a person's former values. Usually, we treat a person's current values as revisions of her past values. When ordinary agents without significant impairments discard values, the fact that they no longer hold the same views is something that they are directly answerable for since they can sufficiently understand and appreciate the considerations on both sides of the matter. This "answerability" is a species of responsibility: it makes sense to ask a person who is answerable why she changed her mind, and this makes her eligible for judgments as to whether she exercised her judgment well or poorly.[49] A person can be directly answerable for his change of values, notice, even if he did not discard his old convictions in a deliberate way. Preoccupied with other things, a person may simply let his former ideals and commitments fade. But this person is nonetheless answerable for the change in view if he remained capable of rational reflection on the views he was drifting away from; it would still make sense to ask him why his mind has changed. In that case, we should still treat his current perspective as authoritative.

49 My thinking about answerability is influenced by David Shoemaker, *Responsibility from the Margins*, though I am adapting the idea for my own purposes. I discuss the same idea under the heading of the "Answerability Approach" in my "Binding the Self."

By contrast, when someone loses the psychological capacity to understand and appreciate why she held her former values, she is not directly answerable for the change of view. The change in what she cares about is not up to her, even if she remains capable of valuing or caring about other things.[50] And so, the reason we usually have for disregarding a person's past values—that this is part of recognizing the authority of judgment—no longer holds. We cannot sensibly regard her current outlook as a revision of her past one. That is why, contrary to Jaworska's suggestion, losing one's values as a direct result of impairment is not comparable to changing one's mind. As Dennis McKerlie observes, the dementia patient in this circumstance "did not change her mind, the disease changed her mind. It stripped her of the ability to even understand her old point of view."[51] Therefore, we have reason to think of those former values as still valid for her, still legitimately her own.[52] They are, as I will say, "enduring interests."

My central claim, then, is that people living with moderate-to-severe dementia sometimes have these enduring interests, even when they remain capable of valuing or caring about other things. Crucially, just because a person has enduring interests does not imply that we should ignore their current values or concerns. Whereas the Status Model asks us to think about prudential authority as belonging in a holistic manner to either a person's former or current perspective taken in its entirety, I am urging a piecemeal approach, according to which we identify the individual attitudes that can still properly be imputed to the person as authentically her own.

50 Bearing responsibility for the event or process that changed one's view is not the key issue. If a person suffers a serious cognitive impairment due to a reckless accident, the person bears some responsibility for the injury. But if the impairment changes his values, he is not answerable for discarding his old convictions since it would not make sense to ask him to explain why he has changed his mind. See also the example of "Hollywood amnesia" below.

51 McKerlie, *Justice between the Young and the Old*, 186. I am trying to offer a fuller explanation as to why this difference matters and make the case that recognizing enduring interests does not preclude attending to a person's present evaluative perspective.

52 In "The Moral Relevance of Past Preferences," Krister Bykvist reasons that, just as it would be "unfair" to count *A*'s preferences involving *B*'s life unless *A*'s preferences are in harmony with *B*'s, so too it would be unfair to count the desire of *A*-at-t_1 for *A*-at-t_2 if it conflicts with the present-oriented desires of *A*-at-t_2. This avoids present-for-past sacrifices but counts the former desires of the now deceased or comatose. (I am not sure if he thinks present desires exclude past ones only if the person remains competent.) Bykvist and I reach some similar conclusions, but he thinks desires we held for longer matter more. This has the counterintuitive implication I mentioned in note 47 above. More generally, Bykvist analogizes our problem to interpersonal cases, whereas I start with the first-person perspective of the agent deliberating about whether he should revise his former values.

3.3. *Further Remarks*

Let me now clarify and expand on a couple of points. First, just what does it mean to say that a person must be able to "understand and appreciate" his former values? The clearest cases will involve losing the ability to grasp the concepts involved in one's former values. Jaworska gives the example of a Mr. O'Connor, who had been "a deeply religious man for whom thoughts of taking his own life or of withholding lifesaving measures for whatever reason were completely unacceptable." However, when his Alzheimer's disease undermined his "capacity for complex reasoning, most of his religious beliefs gradually faded away." Now, grieving for the loss of his wife, he says he is ready to die.[53] If Mr. O'Connor really cannot understand his former religious convictions anymore, then he cannot be responsible for the fact that he no longer holds them. In that case, we should give them some weight in thinking about his best interests.[54] The same would be true if the person retained some understanding of the value but had lost the ability to apply it to his own life.[55] Furthermore, deliberating about what is important to us also draws on our emotional capacities. If a person only has a very superficial grasp of his former values—somewhat in the way that a child might faintly grasp some of the things that adults deeply care about—then again, the person lacks the capacities for taking responsibility for a change in view.[56] For example, perhaps Mr. Amato still retains some command over the basic concept of privacy, but his dementia has affected his ability to feel any of his former concern about it.

It is true that a person without what we would ordinarily regard as an impairment may also say that he cannot understand his former perspective anymore. Suppose that Don formerly wanted to remain a bachelor his entire life, but having found himself with a wife and a child (not in that order), he is now profoundly satisfied with his current life and regards his former perspective as quite alien. Still, I think cases like this are distinguishable from those which undermine responsibility for a change in values. Even if Don no longer

53 Jaworska, "Respecting the Margins of Agency," 107.

54 This does not necessarily mean that his former convictions should have the final word. See below for my discussion of conflicts between a person's former perspective and current one.

55 Similarly, a person could retain his convictions in a rigid way while losing his ability to reflect on their grounds. This, too, could deprive the person of the ability to adapt and revise his beliefs in light of new experience. In this circumstance, it may be appropriate to attribute less weight to the person's present convictions if it seems that the person's unimpaired former self would have applied the values in a different or more nuanced way.

56 Understanding and emotional appreciation are also frequently regarded as requirements for decision-making competence (see Buchanan and Brock, *Deciding for Others*, 23–25).

feels the pull of his former values, he can comprehend what it was about living without familial responsibilities that once appealed to him, and he can recognize the relative significance someone (like his former self) might attribute to living that way. I grant that the difference between Don and someone like Mr. Amato may be a matter of degree, but that is generally true when we talk about the capacities that underlie responsibility. That there will be some hard cases (which may call for caution) does not show that we cannot often differentiate factors that do, and do not, undermine answerability.

Perhaps I should also say that I do not think it is essential to answerability that someone personally remembers having held certain values in the past. Suppose a person suffers from the sort of "Hollywood amnesia" where she remembers nothing about her past prior to an accident but is otherwise unimpaired. On the present account, this person's former values would *not* bear on her current interests so long as she remains capable of reexamining the sorts of reasons that grounded her former values (particularly if someone informed her of what used to be important to her). True, she was not responsible for her initial change of view, but she is answerable for the values she adopts henceforth.

Someone might accept that a person is harmed when he loses the ability to value what used to be important to him but think that, nonetheless, the interest disappears when the value does.[57] That is a possible view, but I am more persuaded by the intuitions on the other side. When we think about what is important for making our own lives go well, we probably suppose that some of these things could still matter even if we should lose the ability to care about them. Certainly, we would be willing to make sacrifices now to improve the prospects of fulfilling our values in such cases. Likewise, it does not seem extravagant to believe that one could set out to harm someone with dementia by undermining a project he had previously invested himself in but could no longer remember or appreciate—for instance, by destroying the manuscript of his last book.[58]

Given this last example, it may seem that the issue of enduring interests is no different from that of posthumous interests. That is not quite true, though there is some overlap. Some find it incredible that anything could affect a person's well-being if he no longer exists. But there is no "missing subject" in our dementia cases. Moreover, not only is the subject very much alive, but the subject is still the kind of being who can have interests grounded in his current perspective. For this reason, dementia cases can involve apparent conflicts

57 Compare Luper, "Posthumous Harm."

58 Someone might accept that destroying the manuscript is contrary to the individual's interests but deny that this affects her *well-being*. I will not quibble about the word. If you think that there are non-welfare interests, then I am happy to say that I am talking about those.

between the earlier and later perspectives, which again makes them unlike posthumous-interest cases.

4. THE REVISION MODEL III: THE IMPORTANCE OF WHAT WE USED TO CARE ABOUT

According to the Revision Model, a person's former values can still impinge on what is good for her, but only if she is not answerable for having discarded those values. How much, though, do these former values really matter in practice?

First, it is evident that people's former values, properly interpreted, do not always apply to their later lives. Some of the things we value are valued because of the way they make us feel. If those activities no longer have the same effect, then the former values will no longer be relevant.[59] There are many activities, similarly, that we only value doing on condition that we will continue to competently perform them. Again, a person might want her life to contain certain goods (like having a career or children) without feeling they must be present in every part of her life. And some values might be specific to particular life stages: Jane might have hated being highly dependent on others in her sixties but accept this as tolerable in her nineties. In other cases, a person has quite definite views about what she wants for her later life, but these are based on predictions or expectations that are not born out. People may, for instance, be overly pessimistic—or overly optimistic—about what life with dementia will be like.[60] Value judgments based on incorrect predictions should be accordingly discounted.

If, moreover, we think that well-being is mainly a matter of being receptive to what is objectively valuable, we will want to be convinced that what a person valued really is of some independent worth. Suppose Friedrich despised the elderly and thought that, when he himself reached old age, he would have no worth and his simple concerns should be disregarded. We might find that value judgment so unreasonable that we discount it or even entirely set it aside.

59 Wendy Mitchell, living with early-stage dementia, writes of her former self: "We would not get on now, you and I. . . . We like different things. You live the work and bustle of a busy city, whereas some days I lose hours just looking out of a window at the view" (*Somebody I Used to Know*, 9). But, even if Mitchell could not understand and appreciate her former value for city dwelling, that former value would not be relevant anymore (on the present account) if it was connected to the way the city used to make her feel. (I owe this reference to Walsh, "Cognitive Transformation, Dementia, and the Moral Weight of Advance Directives," 60.)

60 As Dresser observes in "Dworkin on Dementia."

When *are* a dementia patient's former values likely to be of lasting relevance? One category involves a person's goals or projects. Now, generally, a person must accomplish her own goals if they are to count as achievements. But, since accomplishments can be vulnerable to being undermined, we can have an enduring interest in their protection. In other cases, a goal has been *all but* accomplished, and it is only necessary for someone else to take the final step or fulfill her part so that past labors are not in vain.[61] A second category involves the maintenance of certain personal connections. Someone may, for instance, think it is important that grandchildren continue to visit, even if one day they seem like strangers to him. A third category involves cases in which a person sees something as important to a larger group, community, or tradition which she identifies with. Thus, Ms. Bell saw herself as a part of a family and thought it was important for certain heirlooms, which represented that family, to be passed down from one generation to the next. A fourth category relates to certain ways in which a person wants to be treated. While it is true that a person might cease to care about things like privacy, dignity, and bodily integrity, it may well have been previously important to the person that these things continue to be respected throughout her life. A fifth category relates to personal ideals. For these to be of lasting importance, they cannot be like ordinary moral standards, which do not seem to apply to a person to the extent that she is no longer accountable for her conduct. But not all ideals are like this. For example, someone might think it is important to remain faithful to his wedding vows, even if he should one day forget that he is married. Or, again, Mr. Amato might have formerly thought it important that others keep him clothed in public, even if he should cease caring about this, since walking around disheveled and half naked is not consistent with his conception of dignity. Of course, I acknowledge that some people would not want us to make decisions for them when they have impairments in the future in line with what they currently care about. "Just let me be who I am at that time," they may insist. But this value might itself be understood to be an enduring interest that should guide our interactions with them in the future.

As I have said, just because a dementia patient's former values have enduring relevance for her current interests, it does not follow that we should ignore her current concerns. These can still anchor critical interests. This is a key difference between the Revision Model and the Status Model. A person who cannot do many of the things he once enjoyed might discover simpler goods he had previously neglected. Suppose dementia deprives a man of his capacity to care about his past scholarly projects, but he comes to develop a deeper appreciation

61 Compare Portmore, "Welfare, Achievement, and Self-Sacrifice."

for simply being with loved ones.[62] Acting in this man's best interests might involve both seeking to bring his projects to completion and ensuring he has opportunities to be with his family.

That is an easy case, in a way, since it does not force any trade-offs between a person's former values and what matters to him now. Part of the appeal of the Status Model is that it avoids those trade-offs altogether. But it does so at the cost of simply ignoring values and concerns rooted in either the person's past or present perspective. Moreover, it is not as if the Status Model can entirely avoid making other trade-offs. For one thing, the decision-maker will still face trade-offs between the interests rooted in a person's former values and his experiential interests.[63] Further, whether we are privileging the person's past or present perspective, there will still be trade-offs we have to make between different values internal to that perspective.

If the problem seems more intractable when trying to balance a person's past and present values, that may be due to the way that we deliberate when making decisions for others. In hard cases, at least with someone we know well, we do not identify the person's values individually and then assign them various weights. Rather, we imaginatively project ourselves into the other's place and ask ourselves how she would feel about things given her whole constellation of values. In essence, we run a mental simulation from the other person's perspective. But how can we reason from another's point of view if we must take into consideration values and concerns that are rooted in different perspectives?

We should not exaggerate the problem. A similar issue arises when we must make decisions for children with both immediate and long-term consequences, which requires balancing interests rooted in the child's current and future perspectives. But most people think that, however much uncertainty it may involve, this task of balancing the child's current and future interests can be done. Indeed, it would be strange to accept, in this context, something like the Status Model and say that we should not consider the values the child is likely to have in the future as long as she values some things now.

But I want to suggest that there *is* a perspective we can imaginatively construct that helps us reason about the dementia patient's interests. We begin with the person as she was just before the onset of her impairments. Then

62 Christine Bryden, writing about her experience with dementia, observes: "As my cognitive ability fades, I have felt a greater sense of emotional connection within the community, and an increasing relationship with the divine" (*Will I Still Be Me?* 80).

63 Dworkin apparently assumes that a person's critical interests are always more important than her experiential interests—but that is implausible. The smallest improvements in a person's critical interests do not matter more than (for instance) avoiding intense and prolonged pain. See Shiffrin, "Autonomy, Beneficence, and the Permanently Demented," 209.

we imagine the unimpaired person with us, observing her current situation, empathizing with her current priorities, and we imagine asking her what she would want for her life, given what she knows now, if she knew that she was about to take the place of her impaired self.[64] This may seem to grant a certain priority to the person's past perspective since we are asking how this perspective would change after encountering her present self. But it is hard to see what alternative there is since only the person without impairments can consider both perspectives. In any case, by adopting this approach, we recognize that there is at least one respect in which the person with the impairment is in a better epistemic situation since she is acquainted with what it is actually like to live with dementia—something her former self probably could not adequately appreciate. Ordinarily, when a person gains significant new experiences, that is an occasion for her to revisit and possibly revise her values. But, in this case, the new experience is precisely what prevents her from taking the necessary broader point of view. So, that is why we ask whether the person as she was before the impairment would have changed her mind about anything if she could encounter her current self and foresee taking her place.[65]

5. THE PRESENTIST MODEL

So far, I have been discussing positions that accept that a person's former attitudes can sometimes bear on his best interests. But what about the *Presentist Model*, which holds that only a person's present attitudes can ever affect his present well-being?[66] How compelling is this approach?

Well, you might accept the Presentist Model because you endorse a version of the "experience requirement": that nothing can affect a person's well-being at

64 This echoes (but departs from) Railton, "Facts and Values."

65 In "Cognitive Transformation, Dementia, and the Moral Weight of Advance Directives," Emily Walsh argues that we cannot simply assume that changes in a dementia patients' preferences are due to the degradation of their faculties; they may be due to having undergone the "transformative experience" of living with dementia. I agree that we must consider how the dementia patient's experiences might have changed her mind if she were capable of reflecting on both her former views and her current ones. I disagree with Walsh, however, when she suggests that we can never make a reasonable inference as to what caused a change in perspective. That pessimism seems especially implausible when we look beyond the medical context to cases in which a person has apparently ceased to care about loved ones or past projects. There is much more one could say about transformative experiences and dementia, but the issues are too intricate for me to take them up here.

66 Of course, a person's future wants or values can matter insofar as they bear on her future well-being, but a person's past attitudes cannot matter because we cannot (on this view) affect our past well-being.

a time unless that thing is consciously experienced as good or bad by the person at that same time.[67] Such a view does not *identify* well-being with the quality of the experience (as hedonism does); it only regards that experience as a necessary condition on something's affecting a person's interests.[68] So, even if it is generally good for me to get what I value, it will not benefit me if I never find out that my value was fulfilled. Taking a symmetrical view about harms implies the truth of the old adage, "What you don't know can't hurt you" (at least, not directly).

Those who are skeptical of the experience requirement might appeal to a case from Thomas Nagel, which (following some later retellings) we can dub the "Deceived Businessman." Nagel says that if you think that what you do not know cannot affect your well-being, then "even if a man is betrayed by his friends, ridiculed behind his back, and despised by people who treat him kindly to his face, none of it can be counted as a misfortune for him."[69] If that seems like the wrong answer, then we should reject the experience requirement.

But that is not the only possible ground for the Presentist Model. One might instead invoke an idea I have also drawn upon: that a person cannot directly benefit from something unless it resonates with him in the right way (at some level). Crucially, this simple formulation of the "resonance requirement" leaves matters of timing vague. *When* must a person have the right pro-attitudes toward something if it is going to directly benefit him? Jennifer Hawkins, in an important article, defends a concurrentist interpretation: something cannot directly benefit a person at a given time unless he responds positively to it at that same time or *would* do so were he aware of it. (She accepts a symmetrical view for direct harms.)[70] The implication is that if a person does not care about something anymore, then that thing can no longer bear on her interests. However, since the principle only requires that a person would respond to something in a certain way if he were aware of it, this approach can allow that the Deceived Businessman has been made worse off. Still, Hawkins keeps the hypothetical on a short leash: we are not supposed to idealize the subject's responses in any way beyond imagining her aware of that which she is ignorant.

We can make this view more concrete by considering two versions of a story.

> *Forgiveness 1*: Jay, struggling with alcohol, abandoned his family twenty-five years ago. Feeling remorseful, though, he recently wrote his

67 See Griffin, *Well-Being;* and Sumner, *Welfare, Happiness, and Ethics*.

68 Thus, the experience requirement is not so vulnerable to the most widely accepted problem for hedonism: that most of us would not consider a pleasant life plugged into an "experience machine" to be a good one (Nozick, *Anarchy, State, and Utopia*, 42–45).

69 Nagel, "Death," 4. See also Kagan, "Me and My Life."

70 Hawkins, "Well-Being, Time, and Dementia." She dubs this the "non-alienness principle."

> daughter, Kate, a long apology. He was disappointed when he did not receive a response, but he felt that if only Kate would forgive him in her heart, he would have achieved something in making amends. When Kate first received the letter, it only made her angry again. But soon thereafter, she discovered that her own teenage son was dealing with a drug problem. This softened her feelings toward her father. But with all that was going on, she did not reach out to him right away. Then she received word that he had died.

No doubt it would have been better for Jay if he had learned that Kate forgave him. But what mattered most to him was his daughter's feelings toward him, not his own knowledge of those feelings. So, although the experience requirement would say that Kate's change of heart could not have been good for Jay, Hawkins would take the opposite view since he would have been pleased if he *had* known how she felt. Now consider a second version of the case:

> *Forgiveness 2*: The story begins the same way, but now Kate does call her father as soon as she has the change of heart. However, Jay is so ill at the time he lets the call go to voicemail. He hears Kate say that she forgives him, and he understands what she is saying, but he is feeling so wretched that now the call just annoys him. Before he can recover, he dies.[71]

In this version, since Jay is aware of Kate's feelings but is unmoved by them, the concurrent resonance requirement says that Jay does not benefit.

I think we should reject this interpretation of the resonance requirement. When we imagine how Jay would have responded in Forgiveness 1 to Kate's change of heart, I think we implicitly represent him as sober, not fatigued, not depressed, not distracted with pain, and so on. Why? Because we want to know what Jay's true opinion is, and we assume that those states distort one's authentic perspective. But if that is right, we should also think about Forgiveness 2 differently. If Jay felt so bad that he was not "fully himself" when he learned that Kate forgave him, we should consider instead how he would have responded if he were feeling better.

Someone might deny that we should appeal to Jay's "authentic" attitudes even in the first case; we should only consider how Jay would actually respond given the condition he was presently in. If Jay was ill when Kate had the change of heart, and he would not have been pleased if he had known about it, then we should say that her new feelings are not good for him *at that moment.* But, oddly, this implies that if Jay cycles through a series of moods, then his well-being will

71 This resembles the Princess Lovely example in Hawkins, "Well-Being, Time, and Dementia," 529–30.

wax and wane depending on whether he would then welcome Kate's forgiveness. My sense is that if something can be intrinsically good for someone without it entering his awareness, then the more plausible view is that the welfare effect depends on his stabler underlying evaluative perspective, not his transient moods. In that case, it seems like it is also a person's authentic attitudes that matter when he *is* aware of the event.

We could, however, revise the concurrent resonance requirement so that it refers to how a person would respond to something if she knew about it *and if she were judging it in accordance with her authentic evaluative perspective*. If we assume that a person's authentic or authoritative perspective is grounded in her current psychological dispositions, it would still follow that her former values can have no bearing on her current interests. In my opinion, this is the most plausible basis for the Presentist Model.

Nevertheless, once we accept that it is a person's *authentic* perspective that matters, it is but a short step to the Revision Model, which says that a dementia patient's authentic standpoint may be grounded (in part) in the values she used to hold and never genuinely revised.[72] Moreover, I would suggest that the Revision Model is closer to everyday ways of thinking. If your mother, once devoted to her grandchildren, lost her ability to remember or recognize them, you probably would not declare, "She has obviously changed her mind about how important her grandchildren are to her." No, you would more likely say, "Sadly, she just is not herself in that respect anymore." That, anyway, is what I would want loved ones to say about *me* in that circumstance.

6. Final Thoughts

Let me retrace the outline of the Revision Model. When making decisions for a person without significant permanent impairments, we properly focus on his current values since these represent the outcome of his deliberations to this point about what is important in his life. But if, due to an impairment like dementia, a person cannot understand and appreciate the considerations underlying some of his former values, then the fact that he no longer holds these is not really up to him. Since the person never really revised his values, we should presumptively regard those values as still a part of his last authoritative

72 If well-being can be higher or lower at different times, then at what point in the person's life is his well-being promoted if his former value is realized now. Is it now or is it in the past? I would say that the person is better off once both the relevant attitude and the state-of-the-world have obtained. So, if a former value is fulfilled now, then the person benefits now. Likewise, if he comes to value something now that happened in the past, he also benefits now. See Lin, "Asymmetrism about Desire Satisfactionism and Time."

verdict about his own good. If those values directly bear on his current interests yet seem to conflict with his current values or concerns, we should ask ourselves what the formerly unimpaired person would have wanted for himself if he were here with us now, sympathizing with his current self and that self's distinct perspective, and anticipating taking the place of his current self. A virtue of this approach, I have said, is that it explains why a person's past values may be relevant in cases involving impairment but ordinarily are not.

Someone may object that this account is inconsistent with treating those living with dementia as persons worthy of respect and concern because it emphasizes what people with dementia are losing, rather than who they are now. As such, it might seem stigmatizing or depersonalizing. What people living with dementia want and need is not that we become the caretakers for their past values and projects but that we take their current perspectives seriously. Consider, for instance, the words of Christine Bryden, who has been living with dementia for many years: "I am who I am now, and meaning is what I can find in this present moment. My narrative results from finding meaning in life and developing a sense of identity in the present moment, not based on events in the past."[73] This is a serious concern. The move toward a "person-centered" approach has been the most notable trend in the ethics of caring for people living with dementia over the last couple of decades.[74] If the present account is inconsistent with that approach, that is probably a fatal strike against it.

So, let me respond. First, I hope it goes without saying that I have been assuming throughout that the individual with dementia remains *a person* in the sense of deserving our full moral consideration.[75] My question has simply been whether any of the person's interests might be rooted in her former evaluative perspective. This should be no more depersonalizing than asking whether a person has an interest in fulfilling his future values or desires. Further, I trust that I will not be interpreted as saying that anyone who is diagnosed with dementia (much less diagnosed with a disease that causes dementia) will be unable to authoritatively change their minds about what is important to them. Someone may have significant cognitive impairments, including difficulty remembering parts of their past and in articulating what matters to them now, and yet sufficiently retain the rational and emotional abilities that make

73 Bryden, *Will I Still Be Me?* 99. I am grateful to an anonymous reviewer for recommending I address this issue and for directing me to Bryden's work.

74 See Kitwood, *Dementia Reconsidered*; Sabat, *The Experience of Alzheimer's Disease*.

75 Rejecting the Status Model does not imply that people living with dementia lack moral status. Thus, I demur from the way that Jaworska defines "full moral standing" as inconsistent with recognizing the current prudential value of a person's former values (in "Caring and Full Moral Standing").

them fully answerable for discarding their former values. In all but the most severe cases, to reasonably infer that a person has enduring interests will require having communicated with the individual about her past perspective. Simply observing that the person has certain impairments is not enough.

I also want to reiterate that my main aim has been to offer a theory about the interests that people have—not a theory about the conditions under which we ought (all things considered) to try to promote those interests. It seems to me that most of the concerns about stigmatization are concerned with how we treat people, or the ungrounded assumptions we make about them, not with the sorts of interests they have. Indeed, any act of justified paternalism must (at the minimum) do more good than harm. Just because a person has certain enduring interests, it does not follow that interventions to promote those interests are always justified, particularly if those might inflict significant harms on the person.

More to the point, I consider it a virtue of the Revision Model that, as compared to the Status Model, it does not justify the relevance of a person's former values on the grounds that the patient can no longer act from a "sense of self." On the contrary, I have tried to stress the ways that the current perspective of the person with enduring interests also continues to matter on my approach. Therefore, the Revision Model is very far from regarding the person living with dementia as an "empty shell" of her former self.[76]

In fact, I think that attending to a person's former values *is* often a good way of showing concern and respect for the present person. Consider, first, the individual who is in the process of losing her memory of the values that long structured her life but who remains capable of recalling them with the aid from others. I think it quite plausible that helping this person retain a connection to her past concerns is intrinsically good for her. But why would we think that if we do not suppose that the person retains an interest in values she never truly discards? Why not just think she is in the midst of changing her mind and go with the flow?

By the same token, I think that many people in the early stages of dementia who are anxious about losing important parts of their self-conception would be comforted in knowing that, if they should eventually lose a grip on some of these core values, their loved ones will remember, honor, and (if necessary) attend to them in their stead. In this way, a person's biographical identity can be upheld through relations with others, even while one is losing the ability to preserve it oneself.

So, I do not think that the approach defended here is guilty of depersonalizing or stigmatizing people living with dementia. Of course, I cannot *refute* the

76 A concern voiced by Bryden, *Will I Still Be Me?*

view that respect for personhood demands that we attend solely to the present perspective of the person before us. That might be a fundamental intuition for some. My own approach has been guided by an attempt to combine two ideas. The first is that, because dementia is something that happens to people who have built their lives around certain goods, the interests rooted in those goods can endure even when the person loses the ability to appreciate them. The second is that, even when people do have these enduring interests, that does not mean that their current perspective on what matters to them is unimportant. Whether the Revision Model is compelling is, of course, up to the reader to decide. But I think it deserves a hearing, as it is rooted in some common—and I would venture quite humane—ways of thinking about the interests of people living with advanced dementia.[77]

University of Toronto
andrew.franklin.hall@utoronto.ca

REFERENCES

Baber, H. E. "Ex Ante and Post Hoc Satisfaction." In *Time and Identity*, edited by Joseph Keim Campbell, Michael O'Rourke, and Harry S. Silverstein, 249–67. Cambridge, MA: MIT Press, 2010.

Brandt, Richard B. *A Theory of the Good and the Right*. Oxford: Clarendon Press, 1979.

Bratman, Michael E. "A Desire of One's Own." *Journal of Philosophy* 100, no. 5 (2003): 221–42.

———. "Reflection, Planning, and Temporarily Extended Agency." *Philosophical Review* 109, no. 1 (January 2000): 35–61.

Brudney, Daniel, and John Lantos. "Agency and Authenticity: Which Value Grounds Patient Choice?" *Theoretical Medicine and Bioethics* 32, no. 4 (August 2011): 217–27.

Bryden, Christine. *Will I Still Be Me?* London: Jessica Kingsley Publishers, 2018.

77 Work on this paper was supported by a postdoctoral fellowship funded by the Templeton Foundation ("Love and Human Agency") as well as by an Insight Grant from the Social Sciences and Humanities Research Council. I am grateful to much helpful feedback I have received over the years from audiences at Binghamton University, the University of Toronto, the Centre for Ethics at the University of Toronto, the New York University Ethics Forum, and the 2022 "Ethics, Agency, and Personhood in Dementia" conference in Copenhagen organized by the University of Southern Demark. I also want to individually thank Laura Franklin-Hall, Thomas Hurka, L. W. Sumner, Sergio Tenenbaum, Kevin Vallier, Daniel Viehoff, two anonymous referees for this journal, and especially Agnieszka Jaworska.

Buchanan, Allen E., and Dan W. Brock. *Deciding for Others: The Ethics of Surrogate Decision Making*. Cambridge: Cambridge University Press, 1989.
Bykvist, Krister. "The Moral Relevance of Past Preferences." In *Time and Ethics: Essays at the Intersection*, edited by Heather Dyke, 115–36. Dordrecht: Kluwer, 2003.
DeGrazia, David. *Human Identity and Bioethics.* Cambridge: Cambridge University Press, 2005.
Dorsey, Dale. *A Theory of Prudence.* Oxford: Oxford University Press, 2021.
Dresser, Rebecca. "Dworkin on Dementia: Elegant Theory, Questionable Policy." *Hastings Center Report* 25, no. 6 (November–December 1995): 32–38.
———. "Life, Death, and Incompetent Patients: Conceptual Infirmities and Hidden Values in the Law." *Arizona Law Review* 28, no. 3 (1986): 373–405.
Dworkin, Ronald. "Foundations of Liberal Equality." In *Equal Freedom: Selected Tanner Lectures on Human Values*, edited by Stephen Darwall, 190–306. Ann Arbor: University of Michigan Press, 1995. Originally published in *The Tanner Lectures on Human Values*. Salt Lake City: University of Utah Press, 1990.
———. *Life's Dominion: An Argument about Abortion, Euthanasia, and Individual Liberty.* New York: Knopf, 1993.
———. *Sovereign Virtue: The Theory and Practice of Equality.* Cambridge, MA: Harvard University Press, 2000.
Frankfurt, Harry G. "Freedom of the Will and the Concept of a Person." *Journal of Philosophy* 68, no. 1 (January 1971): 5–20.
Franklin-Hall, Andrew. "Binding the Self: The Ethics of Ulysses Contracts." *Ethics* 34, no. 1 (October 2023): 57–88.
Griffin, James. *Well-Being: Its Meaning, Measurement, and Moral Importance.* Oxford: Clarendon Press, 1986.
Hawkins, Jennifer. "Well-Being, Time, and Dementia." *Ethics* 124, no. 3 (April 2014): 507–42.
Heathwood, Chris. "Which Desires Are Relevant to Well-Being?" *Noûs* 53, no. 3 (September 2019): 664–88.
Jaworska, Agnieszka. "Advance Directives and Substitute Decision-Making." *Stanford Encyclopedia of Philosophy* (Summer 2017). https://plato.stanford.edu/archives/sum2017/entries/advance-directives/.
———. "Caring and Full Moral Standing." *Ethics* 117, no. 3 (April 2007): 460–97.
———. "Caring and Internality." *Philosophy and Phenomenological Research* 74, no. 3 (May 2007): 529–68.
———. "Respecting the Margins of Agency: Alzheimer's Patients and the Capacity to Value." *Philosophy and Public Affairs* 28, no. 2 (Spring 1999):

105–38.
Kagan, Shelly. "Me and My Life." *Proceedings of the Aristotelian Society* 94, no. 1 (June 1994): 309–24.
Kitwood, T. M. *Dementia Reconsidered: The Person Comes First*. Buckingham, UK: Open University Press, 1997.
Lin, Eden. "Asymmetrism about Desire Satisfactionism and Time." In *Oxford Studies in Normative Ethics*, vol. 7, edited by Mark C. Timmons, 161–83. Oxford: Oxford University Press, 2017.
Luper, Steven. "Posthumous Harm." *American Philosophical Quarterly* 41, no. 1 (January 2004): 63–72.
McKerlie, Dennis. *Justice between the Young and the Old*. New York: Oxford University Press, 2013.
McMahan, Jeff. *The Ethics of Killing: Problems at the Margins of Life*. New York: Oxford University Press, 2002.
Mitchell, Wendy. *Somebody I Used to Know: A Memoir*. London: Bloomsbury Publishing, 2018.
Nagel, Thomas. "Death." In *Mortal Questions*, 1–10. Cambridge: Cambridge University Press, 1979.
Nozick, Robert. *Anarchy, State, and Utopia*. New York: Basic Books, 1974.
Overvold, Mark Carl. "Self-Interest and Getting What You Want." In *The Limits of Utilitarianism*, edited by Harlan B. Miller and William H. Williams, 186–94. Minneapolis: University of Minnesota Press, 1982.
Parfit, Derek. *Reasons and Persons*. Oxford: Clarendon Press, 1984.
Portmore, Douglas W. "Welfare, Achievement, and Self-Sacrifice." *Journal of Ethics and Social Philosophy* 2, no. 2 (September 2007): 1–29.
Raibley, Jason. "Well-Being and the Priority of Values." *Social Theory and Practice* 36, no. 4 (October 2010): 593–620.
Railton, Peter. "Facts and Values." *Philosophical Topics* 14, no. 2 (Fall 1986): 5–31.
Raz, Joseph. "Incommensurability and Agency." In *Engaging Reason: On the Theory of Value and Action*, 46–66. Oxford: Oxford University Press, 1999.
———. *The Morality of Freedom*. Oxford: Clarendon Press, 1986.
Sabat, Steven R. *The Experience of Alzheimer's Disease: Life through a Tangled Veil*. Oxford: Blackwell, 2001.
Sarch, Alexander. "Desire Satisfactionism and Time." *Utilitas* 25, no. 2 (June 2013): 221–45.
Scheffler, Samuel. "Valuing." In *Equality and Tradition: Questions of Value in Moral and Political Theory*, 15–40. New York: Oxford University Press, 2010.
Shiffrin, Seana Valentine. "Autonomy, Beneficence, and the Permanently Demented." In *Dworkin and His Critics: With Replies by Dworkin*, edited by Justine Burley, 193–217. Oxford: Blackwell Publishing, 2004.

Shoemaker, David. *Responsibility from the Margins*. Oxford: Oxford University Press, 2015.
Sumner, L. W. *Welfare, Happiness, and Ethics.* Oxford: Clarendon Press, 1996.
Tiberius, Valerie. *Well-Being as Value Fulfillment: How We Can Help Each Other to Live Well*. Oxford: Oxford University Press, 2018.
Wallace, R. Jay. "Caring, Reflexivity, and the Structure of Volition." In *Normativity and the Will: Selected Papers on Moral Psychology and Practical Reason*, 190–211. Oxford: Clarendon Press, 2006.
Walsh, Emily. "Cognitive Transformation, Dementia, and the Moral Weight of Advance Directives." *American Journal of Bioethics* 20, no. 8 (2020): 54–64.
Watson, Gary. "Free Agency." *Journal of Philosophy* 72, no. 8 (April 1975): 205–20.