

PDF version of the entry
Advance Directives and Substitute Decision-Making
<http://plato.stanford.edu/archives/sum2009/entries/advance-directives/>
from the SUMMER 2009 EDITION of the

STANFORD ENCYCLOPEDIA OF PHILOSOPHY



Edward N. Zalta Uri Nodelman Colin Allen John Perry
Principal Editor Senior Editor Associate Editor Faculty Sponsor

Editorial Board
<http://plato.stanford.edu/board.html>

Library of Congress Catalog Data
ISSN: 1095-5054

Notice: This PDF version was distributed by request to members of the Friends of the SEP Society and by courtesy to SEP content contributors. It is solely for their fair use. Unauthorized distribution is prohibited. To learn how to join the Friends of the SEP Society and obtain authorized PDF versions of SEP entries, please visit <https://leibniz.stanford.edu/friends/>.

Stanford Encyclopedia of Philosophy
Copyright © 2009 by the publisher
The Metaphysics Research Lab
Center for the Study of Language and Information
Stanford University, Stanford, CA 94305
Advance Directives and Substitute Decision-Making
Copyright © 2009 by the author
Agnieszka Jaworska
All rights reserved.

Copyright policy: <https://leibniz.stanford.edu/friends/info/copyright/>

Advance Directives and Substitute Decision-Making

First published Tue Mar 24, 2009

There is a rough consensus in medical ethics on the requirement of respect for patient autonomy: physicians must ultimately defer to patients' own decisions about the management of their medical care, so long as the patients are deemed to have sufficient mental capacity to make the decisions in question. For patients who lack the relevant decision-making capacity at the time the decision is to be made, a need arises for surrogate decision-making: someone else must be entrusted to decide on their behalf. Patients who formerly possessed the relevant decision-making capacity might have anticipated the loss of capacity and left instructions for how future medical decisions ought to be made. Such instructions are called an advance directive. One type of advance directive simply designates who the surrogate decision-maker should be. A more substantive advance directive, often called a living will, specifies particular principles or considerations meant to guide the surrogate's decisions in various circumstances, for example, "Do not prolong my life if I enter persistent vegetative state," or "I am a fighter: do not discontinue life-sustaining treatment no matter what happens to me."

This general framework opens up a number of ethical issues. I shall set aside here a foundational issue that is a subject of its own encyclopedia article: What are the criteria for decision-making capacity? These must be specified before we can establish, on any given occasion, whether there would be any need at all for decision-making by a third party (with the aid of an advance directive or not). Assuming we have settled, using the appropriate criteria, that surrogate decision-making is indeed called for, the following main issues arise:

Q1. Who should be the surrogate decision-maker?

Q2. On what basis should the surrogate make the decision? What considerations should she take into account? And, more specifically,

Q2a. Should the advance directive be honored?

This article focuses on philosophical contributions to the last two sets of questions.

- 1. The orthodox legal view
- 2. Challenges to the orthodox view regarding the never competent
- 3. Conflicts across time in the formerly competent
 - 3.1 Threshold of authority approach
 - 3.2 Challenge I: Appeal to the forward-looking perspective of decision-making
 - 3.3 Challenge II: Exercise of will as the point of autonomy
 - 3.4 Challenge III: Loss of personal identity
 - 3.5 Challenge IV: Severance of prudential concern
- Bibliography
- Other Internet Resources
- Related Entries

1. The orthodox legal view

In legal contexts, two general standards or approaches to question Q2 have been developed:

The Substituted Judgment standard:

The surrogate's task is to reconstruct what the patient himself would have wanted, in the circumstances at hand, if the patient

had decision-making capacity. Substantive advance directives are here thought of as a helpful mechanism for aiding the application of Substituted Judgment. The moral principle underlying this legal standard is the principle of respect for autonomy, supplemented by the idea that when a patient is not currently capable of making a decision for himself, we can nonetheless respect his autonomy by following or reconstructing, as best we can, the autonomous decision he would have made if he were able. In a subset of cases, a substituted judgment can implement an actual earlier decision of the patient, made in anticipation of the current circumstances; this is known as precedent autonomy.

The Best Interests standard:

The surrogate is to decide based on what, in general, would be good for the patient. The moral principle underlying this standard is the principle of beneficence. This legal standard has traditionally assumed a quite generic view of interests, asking what a "reasonable" person would want under the circumstances and focusing on general goods such as freedom from pain, comfort, restoration and/or development of the patient's physical and mental capacities. This is because the Best Interests standard has mainly been employed when there is little or no information about the patient's specific values and preferences. However, the concept of best interests is simply the concept of what is best for the person. There is no reason why, in principle, the Best Interests judgment could not be as nuanced and individual as the best theory of well-being dictates.

In practice, the main difference between the two standards is often thought to be this. Substituted Judgment endeavors to reconstruct the subjective point of view of the patient — i.e., the patient's own view of his interests — whenever such reconstruction is a viable possibility. By

contrast, the Best Interests standard allows for a more generic view of interests, without having to rely on the idiosyncratic values and preferences of the patient in question.

The applicability of these standards depends on the context in which the lack of decision-making capacity occurs. Let us distinguish two groups of patients:

Formerly Competent:

Patients who used to have the relevant decision-making capacity, but lost it, for example, due to Alzheimer's disease or other medical problems (or procedures such as surgical anesthesia) undermining normal brain functioning.

Never-been Competent:

Patients who have never had the relevant decision-making capacity, either because the capacity has not yet developed (as in children), or because of a permanent brain deficiency such as severe congenital mental retardation.

The Substituted Judgment standard seems well-suited to the circumstances of the formerly competent patients since, in their case, there are past values or patterns of decision-making that could potentially serve as a basis for the reconstructed decision on the patient's behalf. Furthermore, according to the current orthodoxy, prevalent especially in the law, Substituted Judgment is the preferred solution for formerly competent patients because it promises to preserve respect for autonomy as an overriding moral consideration trumping concerns with beneficence. The picture is this. If, ordinarily, we ought to respect patient autonomy rather than impose our own judgments on patients, we ought to respect autonomy even after the patient has lost decision-making capacity; and we can do so by following or reconstructing, as best we can, the autonomous decision the patient would have made himself when faced

with the current circumstances. In short, in dealing with someone who used to be competent, the widely accepted primacy of respect for autonomy over beneficence calls for Substituted Judgment. And this means that we should use the Substituted Judgment standard whenever possible and fall back on the Best Interests standard only when we lack sufficient information about the patient's prior wishes and values to make Substituted Judgment practicable.

By contrast, for the “never-been competent” patients, the Substituted Judgment standard does not seem applicable (e.g., Cantor 2005): if the patient has never been able to make autonomous decisions in circumstances such as the current one, it seems impossible to reconstruct what the patient's decision would have been. For these patients, the Best Interests standard is the only option.

When combined, these orthodox views generate one unified simple ordering of priority among the several standards and mechanisms for surrogate decision-making, an ordering found in answers to Q2 and Q2a prevailing in the literature (e.g., Brock 1995):

1. Honor a substantive advance directive, as an aid to Substituted Judgment, whenever such directive is available.
2. Absent an advance directive, apply the Substituted Judgment standard based on available information about the patient's past decisions and values.
3. If you cannot apply the Substituted Judgment standard — either because the patient has never been competent or because information about the patient's former wishes and values is unavailable — use the Best Interests standard.

Is this orthodox view correct?

2. Challenges to the orthodox view regarding the never competent

Concerning patients who have never been competent, the orthodox view, as it is typically interpreted, may be misleading in certain cases. By recommending the Best Interests standard as opposed to the Substituted Judgment standard, the orthodox view may help create the impression that, for those who have never had decision-making capacity, only a one-size-fits-all objective assessment of their interests, based on generic goals such as prolonging life or avoiding pain, is available. However, a person may lack decision-making capacity but nonetheless possess the proper starting points of decision-making, so that a surrogate could still reconstruct deeply personal and idiosyncratic choices on the person's behalf. Consider a child or a mildly retarded patient who lacks the capacity to make a sophisticated medical decision because she cannot fully grasp the complex consequences of the available options, or because, if left to her own devices, she would merely choose impulsively. Yet, very meaningful and personally distinctive issues may be at stake for this individual: for instance, alternative treatments may differently impact her relationships with loved ones or differently affect her ability to continue participating in deeply valued activities such as painting or dancing. In such cases, to best serve the interests of the patient, surrogates arguably need to reconstruct the subjective point of view of the patient, and not just fall back on generic choices that "a reasonable person" would make under the circumstances. In short, sometimes — especially in dealing with patients with rich inner lives whose decision-making is nevertheless impaired — the application of the Best Interests standard may look an awful lot like an exercise of Substituted Judgment.

It is only with regards to patients who do not even possess the starting points of decisions — for example, infants or more severely brain damaged individuals — that the idea of reconstructing the individual's

own point of view as a basis for a decision does not even coherently apply, and the more generic application of the Best Interests standard is called for.

Nonetheless, this is only a challenge to the narrow way in which the Best Interests standard has typically been employed: a more nuanced interpretation of the orthodox view can handle the cases of the never-competent appropriately. The application of Best Interests can, in many instances, procedurally resemble the application of Substituted Judgment because, on any reasonable theory of well-being, a large part of what counts as good for a person is attaining what she values or succeeding in what she cares about. It is thus not surprising that reconstructing the individual's viewpoint is an important part of a nuanced interpretation of Best Interests. Yet, even though in employing the Best Interests standard one usually must take very seriously the subject's own viewpoint, one is not thereby recreating the autonomous choice the person would have made. This is particularly clear for those who have never been competent: one cannot be respecting their autonomy (at least not on the usual understanding of autonomous choice), since they have never had autonomy. Moreover, even in undertaking to respect their "starting points of decision-making," one would not treat these starting points as entirely decisive. An individual who has never been competent may value something that would be terribly destructive to her other values (and be incapable of realizing this), and so, to protect her, the Best Interests standard would have to focus on those other values. So here again the application of the Best Interests standard diverges from what would most plausibly count as a reconstruction of the subject's own autonomous choice. Given that Substituted Judgment is grounded in respect for autonomy, it is thus clear why, according to the orthodox view, Substituted Judgment makes no sense for the never-competent, and why the orthodox view prescribes for them the Best Interests standard, albeit interpreted in a suitably broad way.

As already noted, different views on how to apply the Best Interests standard roughly correspond to different theories of well-being. However, theories of well-being are normally developed with an ordinary fully-capacitated human being in mind, so, when applied to those whose incompetence is due, in part, to substantial deviations from this paradigm, some theories need to be adjusted to accommodate human beings who do not at the time, or ever, possess the paradigm capacities these theories presume (for example, the capacity to experience the pleasures of the intellect, or the capacity to desire). The understanding of well-being and the specifics of applying the Best Interests standard in such cases must be tailored to the details of each particular real-life condition — and to the corresponding levels of mental functioning. Interests of children, including infants, have received some attention in the literature (Buchanan and Brock 1990, ch.5, Schapiro 1999); similar tailor-made analyses are needed for individual mental illnesses and brain deficits.

3. Conflicts across time in the formerly competent

The orthodox view regarding the formerly competent faces deeper challenges. In giving priority to Advance Directives and Substituted Judgment, the orthodox view overlooks the possibility that the earlier competent self and the current incompetent self may have conflicting interests. Advance Directives and Substituted Judgment are best suited for the contexts for which they were first developed in the law — conditions involving loss of consciousness such as persistent vegetative state — where the patient in the current incompetent state cannot have interests potentially different from the interests of the person he used to be. However, loss of decision-making capacity often comes about in less drastic, yet permanent conditions, which can leave the current incompetent patient with what seem to be powerful new interests in his new phase of life. Classic cases of this sort occur in Alzheimer's disease,

other forms of dementia, and stroke. Before the loss of capacity, typically, the patient had numerous interests associated with his rich mental life and with a correspondingly complex set of values. Once mental deterioration progresses, the patient's universe of interests shrinks and new interests may become dominant. Sometimes the two sets of interests can come into conflict. Imagine, for example, a fully competent patient who, in anticipation of developing Alzheimer's disease, espouses a strong conviction, perhaps documented in an advance directive, that she does not wish to have her life prolonged in a demented state. She deeply identifies with her intellect, and thus views life with dementia as terribly degrading. But once she develops dementia, her identification with her intellect drops out as a concern, so she loses the corresponding desire not to prolong her life. In the meantime, she is still capable of simple enjoyments — she likes gardening or listening to music — and perhaps can even carry on meaningful human attachments. Her current, truncated set of interests does seem to favor continued life. Such scenarios raise difficult questions of how the interests of the earlier and current self ought to be balanced in surrogate decision-making. Privileging advance directives and recreating the judgment of the earlier self via substituted judgment are no longer the obvious solutions, given this conflict.

Much of the philosophical literature on surrogate decision-making has focused on conflicts of this kind. There are subtle differences, though, in how this conflict is conceptualized — more specifically, in how the interests of the earlier self are viewed — sometimes stemming from differences in what is taken as a paradigm example of the conflict. On one view, the relevant interests of the earlier self are autonomy interests: what matters is that the choices of the earlier self be heeded. With this emphasis, the conflict is between the autonomy of the earlier self and the well-being of the current self. On an alternative conception, the interests of the earlier self are well-being interests: what matters is that the earlier self fares well overall. The conflict, then, is between the well-being of the

earlier self and the well-being of the current self. One may also consider both aspects of the conflict as relevant. The arguments below apply to all three interpretations of the conflict.

3.1 Threshold of authority approach

One way to rescue the idea that the former self and its interests ought to have priority is to appeal to the special *authority* of the former self over the current self. The grounds of this authority are cashed out differently in different views, but the basic thought is that the former self's superior capacities give her standing to govern the current self. Once the current self falls below a certain threshold of capacity, her interests in her current state are so marginal as to no longer be authoritative for how she ought to be cared for, and the interests of the earlier self trump.

Several lines of argument have been used to establish the authority of the earlier self over the current self. One is to deny altogether the independence of the current self's interests. On this interpretation, the conflict described above is merely apparent. Once the current self falls below the relevant threshold of capacity, she is incapable of generating her own independent interests, and, despite superficial appearances to the contrary, her fundamental interests are really defined by the earlier self. The interests of the current self are straightforwardly not authoritative since they are merely apparent interests. Further, even were we to accept that the current self has her own independent interests, there are other reasons to see those interests as lacking authority. If one insists on the priority of respect for autonomy over beneficence, or if one views the capacity for autonomy as the essential core of a person, the interests of the earlier self will be seen as having authority over the current self because only the earlier self is capable of autonomy. Ronald Dworkin's analysis combines all of these lines of argument (Dworkin 1993).

Different versions of the threshold approach propose somewhat different thresholds for when the current interests of a formerly competent individual cease to be authoritative. It is usually accepted that the mere loss of decision-making capacity is insufficient (Dworkin 1993, 222-29). Decision-making capacity is context-specific and depends on the complexity of the pertinent information that the decision-maker needs to process. A person may lose the ability to make very complex medical decisions, while still being able to decide perfectly well about simpler everyday matters. Lapses of this nature would not give the surrogate a license to discount the current well-being of the individual in favor of what mattered to him earlier. By contrast, transformations that could leave authority with the past self must involve a more global loss of capacity such that one can no longer generate, in any context, interests of a special, morally weighty type. In crossing this threshold, one ceases to be a being of a certain morally privileged kind: for instance, one ceases to be an autonomous individual, or one turns from a person into a nonperson. If an autonomous individual loses his capacity for autonomy altogether — the thought then goes — he may have some local (possibly merely illusory) interests associated with the non-autonomous self, but his affairs ought to be conducted in accordance with his earlier wishes expressive of his autonomy. Or, in the parallel version, if a person turns into a nonperson, he may have some local (possibly illusory) interests as a nonperson, but his affairs ought to be conducted so as to advance the interests of the person he used to be.

Within this basic framework, several variants are possible, depending on what one takes to be the essential characteristics of a person, or, if one accepts the capacity for autonomy as the essence of personhood, depending on what one takes to be the core aspects of autonomy. Ronald Dworkin's influential work defends the capacity for autonomy as the relevant threshold, with autonomy interpreted as “the ability to act out of genuine preference or character or conviction or a sense of self” (Dworkin

1993, 225). If an individual has lost the capacity for autonomy so understood, this view dictates that her current interests (illusory or not) have no authority over decisions on her behalf, and surrogates ought to cater to her former interests, from before the loss.

It is, however, important to notice that the capacity for autonomy, as interpreted by Dworkin, comprises two distinct abilities: (1) the ability to espouse a “genuine preference or character or conviction or a sense of self” — what may be called, for short, the ability to value — and (2) the ability to act out of one's sense of conviction, that is, the ability to enact one's values in the complex circumstances of the real world. In many brain disorders these two abilities come apart. For example, a patient in the middle stages of Alzheimer's disease may retain genuine values — she may hold on to family ties or to the conviction that helping others is good — and yet, due to a rapid deterioration of short-term memory, she may be perpetually confused and unable to figure out how to enact these values in the concrete circumstances of her life. The set of values such a patient retains would typically be a curtailment of the original set, introducing the potential for conflict between the interests of the earlier and current self. For example, earlier, the person may have valued independence above all else, and so was adamantly against having her life prolonged if she developed Alzheimer's disease. Now, in moderate stages of Alzheimer's, she has lost her commitment to independence, but still values emotional connections to family members, and thus has a strong interest in continuing to live. On Dworkin's approach to decisions on this individual's behalf, her current interests are not allowed to override her earlier interests because she has lost her standing as an autonomous agent: due to her confusion, she is unable to act on her commitment to the family ties or on any other values — she is unable to run her life by her own lights, that is, to govern herself. However, on an alternative view (Jaworska 1999), what matters most for autonomy and personhood are the starting points of autonomous decision-making: the genuine values that

the person still holds. So long as an individual is capable of valuing, she remains a being of a morally privileged type, and interests stemming from her values have the authority to dictate how the individual ought to be treated. The person need not be able to enact her values on her own — it is part of the surrogate's role to assist with this task. In short, on this alternative view, the capacity to value marks the morally crucial threshold above which the current interests of a formerly competent individual remain authoritative for the surrogate's decisions and the conflicting interests of the earlier self can be set aside.

The two views I have just discussed share the underlying idea of a threshold of capacity beyond which an individual's current interests lose authority. This idea has been challenged in several ways.

3.2 Challenge I: Appeal to the forward-looking perspective of decision-making

The most straightforward challenge emphasizes that decision-making inherently involves a present- and future-oriented perspective: the surrogate must make the best decision for the patient in front of him about how to manage this patient's life from now on. The patient may have had different interests in the past, but how can these be relevant to current decisions, which can only affect the present and future but not the past? This approach may accept it as unfortunate that the patient's past interests were left unfulfilled, but insists that this unfortunate fact cannot be remedied, and that there is no use in catering to bygone interests in current decision-making (Dresser 1986).

An advocate of the threshold view, such as Dworkin, would emphasize two points in response:

First, past interests can often be satisfied in the present. Dworkin distinguishes between what he calls “experiential” and “critical” interests

(Dworkin 1993, 201-08). Experiential interests are, roughly, interests in having desirable felt experiences, such as enjoyment (and in avoiding undesirable experiences, such as boredom). These interests are indeed tied to the present: there is no point in trying to satisfy one's past experiential interest in a specific enjoyment (for instance, in playing with dolls), if one at present has no hope of still deriving enjoyment from what one used to enjoy in the past. By contrast, critical interests are not tied to the experience of their satisfaction; these are interests in having what one values or cares about become a reality, such as a parent's interest in the success and prosperity of his child or a sailor's interest in preserving his beautiful wooden boat. According to Dworkin, such interests can be meaningfully satisfied even if they belong in the past: for example, even after the sailor dies, it makes sense to preserve the boat he cared about and do so for his sake. Similarly, according to Dworkin, it makes sense to satisfy a formerly competent person's critical interests, such as the interest in avoiding the indignity of dementia, for her sake, even if she has ceased to understand those critical interests now.

Second, on a view like Dworkin's, the past critical interests of an individual who formerly possessed the capacity for autonomy are, in a crucial sense, still her interests in the present, even if she can no longer take an interest in them. This is an essential element of the claim that the patient's earlier autonomous self has authority over her current non-autonomous self. The thought is this. For any person, the interests she has autonomously defined for herself are her most important interests. And this is so even for an individual who has lost her capacity for autonomy or her personhood: so long as the individual survives the loss as numerically the same entity, her interests stemming from autonomy (or the subset of them that can still be satisfied) remain her most important interests, even if she can't espouse them now, and they are, in this sense, "past." Thus, Dworkin offers a powerful rationale for why satisfying "past" interests can still matter, and matter very deeply, in the present.

3.3 Challenge II: Exercise of will as the point of autonomy

The versions of the threshold view that see the capacity for autonomy as the relevant threshold can be challenged by approaches that cast the requirements of the capacity for autonomy as being so minimal that *any* individual capable of generating independent interests in his deteriorated state counts as autonomous. On such approaches, conflicts between earlier interests grounded in autonomy and later interests no longer so grounded become impossible, and the claim of authority of the earlier autonomous self over the current non-autonomous self loses its bite: the threshold of autonomy is so low as to cease to mark any contestable difference in authority. Seana Shiffrin's response to Dworkin can be interpreted as a view of this sort (Shiffrin 2004). Shiffrin sees a key point of autonomy in the ability to exercise one's own will: the ability to control one's experiences through the enactment of one's own choice. Shiffrin emphasizes that so long as an individual has this ability, its exercise calls for protection, and this is a crucial part of what we protect when we respect autonomy. On this picture, so long as an individual is able to make choices, have preferences, exhibit a will, etc., there is a rationale for catering to his current interests, and so his current interests have authority to override interests espoused in the past.

The proponent of the threshold view may, in response, acknowledge the importance of the ability to control one's experience through acts of will, but still insist that a more robust capacity for autonomy — for example, a capacity that involves expression of values and not just mere preferences — has moral importance of an altogether different order. This difference can then support the position that, in cases of conflicts between an earlier self capable of such robust autonomy and a current self merely capable of exercises of will, the earlier self retains authority and her interests ought to be heeded.

3.4 Challenge III: Loss of personal identity

According to the threshold views, the earlier self has authority to determine the overall interests of the patient because the current self has lost crucial abilities that would allow it to ground these overall interests anew. This picture assumes that the earlier and current self are stages in the life of one entity, so that, despite the talk of local interests associated with each life-stage, there is an underlying continuity of interests between the two. But this is a very substantial assumption, and it has been contested by appeal to an influential account of the metaphysics of personal identity over time, the psychological continuity account. Roughly, the idea is that, in the wake of a drastic transformation of one's psychology such as Alzheimer's disease, one does not survive as numerically the same individual, so whatever interests one's predecessor in one's body may have had are not a suitable basis for decisions on behalf of the new individual who has emerged after the transformation (Dresser 1986). The lack of identity between the earlier and current self undercuts the authority of the former over the latter.

This approach works best in cases in which we can assume that the new entity emerging after the psychological transformation is still a person: the interests of the earlier self cannot dictate how the current self ought to be treated because it would be a clear violation of the rights of persons to allow one person to usurp the affairs of another. (Some may doubt whether loss of numerical identity without loss of personhood is even possible in any real-life cases of dementia or brain damage, but the theoretical point still holds.) What, though, if the psychological deterioration is indeed severe enough to strip the resulting entity of the capacities of a person?

Some might see the loss of personhood as a particularly clear-cut sign of a change in numerical identity: if the current self is not even a person,

surely the current self cannot be the same person as the earlier self. However, as David DeGrazia has emphasized, this line of reasoning rests on an undefended (and controversial) assumption that we are essentially persons (DeGrazia 1999). For if we are not essentially persons — but, rather, for example, conscious minds of some other, less complex kind — an individual may very well lose the properties of a person without any threat to his numerical survival.

Nonetheless, even if we are not essentially persons, on the psychological view of our identity, we are essentially defined by our psychological properties. If these properties change drastically enough, the old individual ceases to exist and a new individual comes into existence. And the transformation of a person into a nonperson does seem to be a drastic psychological transformation. Thus, even if DeGrazia is right that loss of numerical identity does not automatically follow from loss of personhood, it is certainly possible, and perhaps even likely, on the psychological view of our identity, that a transformation of a person into a nonperson would involve such a profound psychological alteration as to result in a numerically new being. How should we adjudicate conflicts between the earlier and the current individual in such cases?

On one view, if a person turns into a new individual in the later stages of dementia, this by itself undercuts the authority of the earlier person over her successor, regardless of whether the successor is a person or not. After all, why should an altogether different individual dictate how the current self is to be treated? However, more nuanced positions can also be found in the literature. Buchanan and Brock (1990) see the authority of the earlier self in cases of loss of numerical identity as crucially dependent on whether the current self is still a person. They accept that if the current self is a person, it would be a violation of her rights as a person to allow another individual to commandeer her affairs. However, if the current self is no longer a person, he lacks the same rights. And, as

Buchanan and Brock see it, the earlier self has “something like a property right... to determine what happens to [his] nonperson successor” (166). That is, if one ceases to exist by turning into a nonperson, one retains a quasi-property right to control the resulting nonperson, presumably in much the same way that, when one ceases to exist by turning into a corpse, one has a quasi-property right to control the resulting corpse. Hence, on this approach, even if the earlier and current self are distinct individuals, the earlier self has the authority to determine what happens to the current self, so long as the current self has been stripped of personhood. In this way, the idea of a threshold of capacity beyond which the earlier self gains authority to dictate the current self's affairs is resurrected, despite the assumption that the earlier and current self are not the same individual. But, this time, the basis of the authority is different: it is not grounded in the continuity of overall interests between the two selves, but rather in the earlier self's quasi-property right. Note, though, that the claim that quasi-property rights could extend to rights over successors who are nonetheless conscious beings is controversial and requires further defense.

3.5 Challenge IV: Severance of prudential concern

It is possible to retain the intuitive idea that the weakness of psychological connections between the two selves undermines the authority of the earlier self over the current self without accepting the metaphysical view that the earlier and current self are numerically distinct entities. Suppose we maintain that even the most drastic mental deterioration is not equivalent to death — that the same individual persists through the ravages of Alzheimer's disease. We can still question the continuity of interest between the earlier and current self by examining the concern the earlier and current self would appropriately have for one another (McMahan 2002).

Ordinarily, each of us has a very unique type of concern for our own past and future selves: it matters to you, in a very special way, what happens to *you*, what experiences *you* undergo, how *you* act, etc., now, in the future, and in the past. Call this special species of concern *prudential concern*. We normally assume that prudential concern strictly tracks personal identity: one has prudential concern only for oneself and one is always concerned with oneself in this way. By contrast, Jeff McMahan has argued that (at least within the bounds of numerical identity) prudential concern also ought to track the degree of psychological ties: prudential concern of two selves at different stages of life for one another ought to weaken in proportion to the weakness of the psychological connection between them (McMahan 2002, 69-82). In the context of the severe psychological transformations caused by a disease like Alzheimer's, this means that the appropriate level of prudential concern of the earlier and current self for one another would be rather slight. The two selves are not bound by a sufficiently strong common prudential interest, and so the person's earlier interests, no matter how important, do not transfer as particularly important interests of her current psychologically disconnected self. Any potential conflicts between the interests of the two selves are now akin to conflicts between two entities with altogether independent interests.

David DeGrazia has tried to counter this picture by suggesting that, in addition to the factors discussed by McMahan, the earlier self's appropriate degree of prudential concern for the current self can be boosted by the earlier self's autonomously formed self-narrative: if the earlier self identified with the current self, in the sense of seeing the current self as a true stage in the unfolding complex narrative of her life, strong prudential concern is warranted (DeGrazia 2005, 196). Oddly, this view makes the rational level of prudential concern for the future partly a matter of choice of the earlier self. Yet, unlike our ordinary concerns for specific plans, projects, other people, etc., which are up to us, prudential

concern is a requirement of rationality and should not be a matter of choice. Just as we cannot rationally have prudential concern for someone else simply because we happen to incorporate (perhaps delusionally) their life into our self-narrative, similarly, if we are not otherwise warranted to have prudential concern for our own future self, we cannot change this simply by virtue of how we happen to construct our self-narrative.

Return, then, to McMahan's picture. If the interests of the earlier and current self are genuinely prudentially estranged and independent, how should we resolve conflicts between them?

When the current self is still a person, her rights as a person call for allowing her interests to control her treatment and against letting the interests of the earlier self interfere; any weakness in the current self's prudential connection to the earlier self merely reinforces this stance. But how should we balance the interests of the two prudentially estranged selves when the current self is not a person?

McMahan himself suggests that, in a conflict between an earlier person and a current nonperson, the interests of the earlier person ought to trump (McMahan 2002, 502-03). He emphasizes that the earlier self is a "higher self," "rational and autonomous," and that her interests are associated with the dominant, more substantial and lengthy part of life, integrated — through strong prudential connections among its various stages — into one unified life-segment. But part of McMahan's reasoning is also that the interests of the current nonperson are, as types of interests, not very substantial. Here McMahan relies on the specifics of the example of conflict he happens to analyze (a version of the "preference to die" example we saw earlier), together with a controversial claim that a severely demented patient would not have a strong interest in continuing to live.

While McMahan may be right that strong interests of an earlier person

trump comparatively trivial interests of a current nonperson, his answer covers only a subset of possible conflicts between the earlier and current self. It is much more difficult to arbitrate between the two selves when the interests of the current nonperson are also substantial. I have argued (Jaworska unpublished) that when the interests of the earlier person are relatively minor, substantial interests of the current nonperson ought to trump. Thus, for example, if the earlier self had only a relatively weak preference to die — let's say she simply didn't want to further complicate relationships with family members for whom she happened not to care all that much — the current self's substantial interest in continuing to live ought to prevail. More controversially, I also argued that highly vital interests of the current nonperson ought to trump even some non-trivial, rather serious interests of the earlier person. Thus, in the standard version of the "preference to die" scenario, what is at stake for the earlier self is a rather serious interest in maintaining the integrity of her life narrative. However, this interest does not reach the level of high vitality, because there is only so much damage that a period of senility at the end of a life can do to an otherwise successful life-narrative. By contrast, the current self's interest in her very survival is more highly vital. It is also, unlike the earlier self's interest, an active interest — an ongoing subject of interests has reason to be strongly prudentially invested in it. These factors combined lend, in this scenario, priority to the interest of the current self.

Bibliography

- Brock, D., 1995, "Death and Dying: Euthanasia and Sustaining Life: Ethical Issues," in *Encyclopedia of Bioethics* (Volume 1), W. Reich (ed.), New York: Simon and Schuster, 2nd edition, pp. 563-72.
- Buchanan, A. E. and Brock, D. W., 1990, *Deciding for Others: The Ethics of Surrogate Decision-Making*, Cambridge: Cambridge University Press.

- Cantor, N., 2005, “The Bane of Surrogate Decision-Making: Defining the Best Interests of Never-Competent Persons,” *The Journal of Legal Medicine*, 26(2): 155-205.
- DeGrazia, D., 1999, “Advance Directives, Dementia, and ‘the Someone Else Problem’,” *Bioethics*, 13(5): 373-91.
- DeGrazia, D., 2005, *Human Identity and Bioethics*, Cambridge: Cambridge University Press.
- Dresser, R., 1986, “Life, Death, and Incompetent Patients: Conceptual Infirmities and Hidden Values in the Law,” *Arizona Law Review*, 28(3): 373-405.
- Dworkin, R., 1993, *Life's Dominion: An Argument about Abortion, Euthanasia, and Individual Freedom*, New York: Knopf.
- Jaworska, A., 1999, “Respecting the Margins of Agency: Alzheimer's Patients and the Capacity to Value,” *Philosophy and Public Affairs*, 28(2): 105-138.
- Jaworska, A., unpublished, “Vanishing Persons and the Authority of the Former Self: Dilemmas in Alzheimer's Disease.”
- McMahan, J., 2002, *The Ethics of Killing: Problems at the Margins of Life*, Oxford: Oxford University Press.
- Schapiro, T., 1999, “What is a Child?” *Ethics*, 109(4): 715-738.
- Shiffrin, S. V., 2004, “Advance Directives, Beneficence, and the Permanently Demented.” in *Dworkin and His Critics with Replies by Dworkin*, J. Burley (ed.), Oxford: Blackwell, pp. 195-217.

Other Internet Resources

- Sample Advance Directive, American Academy of Family Physicians
- Papers on Advanced Directives, in philpapers.org.
- The Limited Wisdom of Advance Directives, Chapter 2 of *Taking Care: Ethical Caregiving in Our Aging Society*, The President's Council on Bioethics Report.

Related Entries

autonomy: personal | beneficence, principle of | decision-making capacity | legal obligation and authority | paternalism | personal identity | personal identity: and ethics | respect | well-being

Acknowledgements

Special thanks to Jennifer Hawkins and Govind Persad for very helpful comments on earlier drafts of this entry. Thanks also to Caleb Perl for research assistance.

Copyright © 2009 by the author
Agnieszka Jaworska