Advance Directives and Personal Identity

Elisabeth Furberg
For my mother.
## Contents

1. Introduction ............................................................................................................. 9  
   Advance directives .................................................................................................. 9  
   The Objection from Personal Identity ................................................................. 10  
   Questions and purpose ......................................................................................... 15  
   Responses to the argument .................................................................................. 15  
   Methodological remarks ....................................................................................... 18  
   Overview of this book ........................................................................................... 22

2. Personal identity: biology over psychology? ......................................................... 25  
   Introduction ........................................................................................................... 25  
   Identity: two rival views ....................................................................................... 26  
   Infants, foetuses and corpses .............................................................................. 28  
   Are we animals? ..................................................................................................... 35  
   Conclusion ............................................................................................................ 40

3. Is identity really what matters? ............................................................................ 43  
   Introduction ........................................................................................................... 43  
   Parfit's views on identity ....................................................................................... 44  
   Fission .................................................................................................................... 46  
   Fission and the advance directive ...................................................................... 49  
   Another suggested implication ............................................................................ 50  
   Conclusion ............................................................................................................ 51

4. Can posthumous events harm us? ...................................................................... 53  
   Introduction ........................................................................................................... 53  
   Making a case for posthumous harm .................................................................... 54  
   Debunking the intuition ....................................................................................... 56  
   The timing problem .............................................................................................. 59  
   The possibility of posthumous harm and the advance directive ...................... 62  
   Conclusion ............................................................................................................ 63

5. The normative relevance of personhood ............................................................. 65  
   Introduction ........................................................................................................... 65  
   Buchanan and Brock’s argument ......................................................................... 65  
   A low (conservative) threshold .......................................................................... 67  
   A quasi-property right ......................................................................................... 70
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kuhse's argument</td>
<td>72</td>
</tr>
<tr>
<td>Are non-persons harmed by death?</td>
<td>73</td>
</tr>
<tr>
<td>Conclusion</td>
<td>75</td>
</tr>
<tr>
<td>6. Consequentialism and advance directives</td>
<td>77</td>
</tr>
<tr>
<td>Introduction</td>
<td>77</td>
</tr>
<tr>
<td>Classic utilitarianism</td>
<td>78</td>
</tr>
<tr>
<td>An alternative theory of value</td>
<td>78</td>
</tr>
<tr>
<td>An (act-) utilitarian defence of advance directives</td>
<td>80</td>
</tr>
<tr>
<td>Some problems with the suggested argument</td>
<td>83</td>
</tr>
<tr>
<td>Conclusion</td>
<td>85</td>
</tr>
<tr>
<td>7. Epilogue</td>
<td>87</td>
</tr>
<tr>
<td>Summary</td>
<td>87</td>
</tr>
<tr>
<td>Final words</td>
<td>89</td>
</tr>
<tr>
<td>Further issues</td>
<td>90</td>
</tr>
<tr>
<td>References</td>
<td>92</td>
</tr>
</tbody>
</table>
Acknowledgements

The road leading up to the finishing of this book has been both long and winding, and it would be impossible for me to mention all the people who have been a support during my journey. Some, however, clearly deserves special mention. Firstly, I would like to express my deepest gratitude to my supervisors Niklas Juth and Torbjörn Tännsjö, mainly for the patience they have shown when reading my texts and for suggesting how to improve on them, but also for cheering me on during the last rather intensive months of work. For supervision and for cheering me on during the less intensive first year I would also like to thank Stellan Welin.

For valuable comments on my texts I would also like to thank all participants in the practical philosophy seminars in the Department of Philosophy at Stockholm University. Some of the more regular participants include Katharina Berndt-Rasmussen, Mats Ingelström, Sofia Jeppsson, Sandra Lindgren, Nicolas Olsson-Yaouzis, Jonas Olson and Olle Torpman. A special word of thanks is due to Jens Johansson who helped me a great deal with the particular questions dealt with in chapters two and four. Naturally, all of the remaining weaknesses and errors must be charged to my own account.

Apart from being grateful to those who have helped me when it comes to philosophical matters, I am also – and no less – grateful to all the people who have supported me by just being good friends and colleagues. Thank you for making the often times all too solitary work of writing a thesis bearable.

In 2010 I was able to spend four rewarding months at the Uehiro Centre for Practical Ethics at Oxford University. I am very grateful to those who made this possible and to all who made me feel welcome while in Oxford. Special thanks to Julian Savulescu who arranged for my visit and to STINT, the Swedish Foundation for International Cooperation in Research and Higher Education, as well as to the Knut and Alice Wallenberg Foundation, for providing me with the necessary funding.

Last but not least, I would like to thank Robert for his never-ending support and for believing in me even at the many times when I have failed to do so myself, and Sigrid and Salme – for being such delightful remainders of what really matters.
1. Introduction

Advance directives

Advance directives, also known as advance healthcare directives, personal directives, advance decisions or living wills, are instructions given by patients – or potential patients – specifying what actions ought to be taken for their health in the event, due to illness or incapacity, of their no longer being capable of making decisions themselves. Even though advance directives can address several different healthcare issues – such as the use of dialysis and breathing machines, tube feeding, blood transfusion, pain relief as well as organ or tissue donation – the term is perhaps most commonly associated with a request of a particular kind, namely for life-sustaining measures to be withheld or discontinued in the event of an incurable and terminal disease. These latter kinds of cases will be our main concern in this book even if many of the arguments presented are also applicable to a wider range of cases.

Over the last decades, there has been a rising tide of opinion in favour of advance directives. Not only is the use of such directives recommended by most medical and advisory bodies, they are also gaining increasing legal recognition in many parts of the world.¹ This development should perhaps come as no surprise given two recent trends in healthcare and in health care ethics. Firstly, medicine is continually increasing its capacity to prolong life without necessarily being able to restore well-being. Patients who most certainly would have died had they not been admitted to hospital, can now be kept alive for long periods of time in comatose or permanently vegetative states. However, a longer life, most people seem to agree, is not necessarily a better life and especially not if the years that could be added to one’s lifespan are expected to be painful, unhappy, undignified, possibly also in combination with being very costly – economically and/or psychologically – to society, healthcare professionals or the patient’s family.

Secondly, there has been a recent trend in healthcare ethics to stress patient autonomy, granting patients the right to decide over large parts of their

¹ See e.g. Robertson (1991), and Davis (2007).
medical care, even when they are only capable of expressing their decisions in advance.

Many ethicists, however, argue that it is a mistake to view advance directives as simple tools for prolonging patient autonomy. This is so, it is often suggested, because there are several significant asymmetries between contemporaneous choice by a competent individual and the issuance of an advance directive to cover future decisions.²

One such asymmetry is the fact that a decision regarding a future state can never be as well informed as a contemporaneous choice: aside from the virtual impossibility of knowing exactly what the future has in store, therapeutic options may change between the time an advance directive was issued and the time at which it is to be implemented.

Another asymmetry is that important formal safeguards that ordinarily tend to restrain imprudent or unreasonable choices are less likely to be present to the same extent in the case of an advance directive: relatives and/or healthcare professionals who might offer their view to help the patient form a more reasonable (and less imprudent) decision might not be there at the time when a potential patient formulates his/her advance directive.

A third asymmetry is that the assumption that a competent person is the best judge of her own interests is weakened when she makes a choice about future contingencies under conditions in which those interests may have changed in radical and unforeseen ways.

Apart from the problems just mentioned, there is one, as we shall see, far more philosophically controversial argument against granting advance directives moral authority. This is an argument that Allen Buchanan, one of the most influential ethicists in the advance directives debate, introduces as “a much more profound and potentially grave threat to the moral authority which remains even if we conclude that, all things considered, the asymmetries cited above do not provide sufficient grounds for limiting that authority”.³ The argument Buchanan refers to I will label the Objection from Personal Identity, and it will be the main subject in this book.

The Objection from Personal Identity

The origins of the Objection from Personal Identity can be found in the writings of Rebecca Dresser.⁴ Together with John Robertson,⁵ Dresser is probably the most influential critic of advance directives in the philosophical de-

² See e.g. Buchanan (1988).
³ Ibid. p. 280. Even though this is the way that Buchanan introduces the Objection from Personal Identity he later argues that it can readily be dismissed. See chapter five.
Irrespective of the identity question, Dresser and Robertson argue that when there is a conflict between doing what the advance directive prescribes and doing what is in a patient’s current best interest, we ought to do the latter. And this is so, they claim, because there is insufficient continuity of interests between the competent author of the directive and the later patient for the former to have authority.

Naturally, not all philosophers have agreed with this conclusion. Ronald Dworkin famously argues, for example, that even when there seems to be a conflict between doing what the advance directives prescribes and acting on what we believe to be in the current patient’s best interest, we ought to abide by the advance directive. To illustrate his argument, Dworkin describes the case of Margo.

When Andrew Firlik was a medical student, he met a fifty-four-year-old Alzheimer’s victim whom he called Margo, and he began to visit her daily in her apartment, where she was cared for by an attendant. The apartment had many locks to keep Margo from slipping out at night and wandering in the park in a nightgown, which she had done before. Margo said she knew who Firlik was each time he arrived, but she never used his name, and he suspected that this was just politeness. She said she was reading mysteries, but Firlik “noticed that her place in the book jumps randomly from day to day; dozens of pages are dog-eared at any given moment... Maybe she feels good just sitting and humming to herself, rocking back and forth slowly, nodding off liberally, occasionally turning to a fresh page.” ... Firlik was confused, he said, by the fact that “despite her illness, or maybe somehow because of it, Margo is undeniably one of the happiest people I have ever known.” He reports, particularly, her pleasure of eating peanut-butter-and-jelly sandwiches. But, he asks, “When a person can no longer accumulate new memories as the old rapidly fail, what remains? Who is Margo?”

Let us now assume that Margo, years earlier, had formulated an advance directive expressing the well-informed decision (or desire) to forego lifesaving treatment under circumstances very similar to the ones described above – ought we abide by it? Dworkin argues that we should, and he does so by appealing to what is usually called “future-oriented”, “prospective” or “precedent” autonomy.  

Those who defend advance directives by reference to precedent autonomy believe that the reasons for abiding by advance directives (even in cases such as Margo’s) are the very same reasons that we have for respecting a person’s autonomy or self-determination. Dworkin thus begins his precedent autonomy account by making a distinction between what he calls the integrity view and the evidentiary view of autonomy.

---

6 See also Dresser & Robertson (1989).


8 Davis (2004).
The evidentiary view is characterised as the belief that we ought to respect a person’s autonomy or right to self-determination because “each person generally knows what is in his own best interests better than anyone else”. This view, Dworkin points out, would not support the case for respecting a patient’s former preferences because “[p]eople are not the best judges of what their own best interests would be under circumstances they have never encountered and in which their preferences and desires may drastically have changed”.

However, Dworkin rejects the evidentiary view in favour of the integrity view. The value of autonomy, in this view, derives not from the well-being it produces for the choosing agent, but rather “from the capacity it protects: the capacity to express one’s own character ... in the life one leads. ... It allows each of us to be responsible for shaping our lives according to our own coherent or incoherent – but, in any case, distinctive – personality”. Dworkin continues:

[If] we accept the integrity view, we will be drawn to the view that Margo’s past wishes must be respected. A competent person making a living will providing for his treatment if he becomes demented is making exactly the kind of judgment that autonomy, on the integrity view, most respects: a judgment about the overall shape of the kind of life he wants to have led.

Although Dworkin recognises that when a person is entrusted to the care of another person the former has what he calls a right to beneficence, he denies that there could be a genuine conflict between acting in accordance with the advance directive and doing what is in demented Margo’s best interest. Although a demented patient may very well retain some experiential interests – i.e. the things that people do just because they like the experience of doing them e.g. cooking or eating out, going to the cinema, doing sports etc. – she cannot have any contemporary opinion about what is in her critical best interest – i.e. interests which, if not satisfied, would cause people to think they were worse off in some way or that their life had been wasted, e.g. interests in having a close relationship, accomplishing a particular task or fulfilling a duty. A demented patient, it seems, simply lacks the “sense of a whole life, a past joined to a future” that is necessary to make an evaluation thereof. However, when the person formulated an advance directive she did have this ability, and to disregard this fact would be no better than to override a competent adult’s judgment about his or her care.

---

10 Ibid. p. 226.
11 Ibid. p. 224.
12 Ibid. p. 226.
13 Ibid. p. 230.
In a specific response to Dworkin’s theory and defence of advance directives, Dresser raises several concerns. One such concern is that Dworkin is unclear when it comes to the question of how his theory is to be translated into policy. For example, Dresser claims, Dworkin states that his ideas apply only to the late stages of Alzheimer’s disease, but makes implementation of Margo’s former wishes contingent on the mere development of the condition. Yet another problem, according to Dresser, is that Dworkin notes, but fails to address, the argument that in the circumstances of dementia, critical interests become less important to the patient and experiential interests more so – a fact that seems to justify a policy against withholding treatment from dementia patients whose lives offer them the sort of pleasures and satisfactions Margo enjoys.

The one concern of Dresser’s that is our main interest here is, however, the one she raises in the following quote:

Dworkin assumes that Margo the dementia patient is the same person who issued the earlier requests to die, despite the psychological alteration that has occurred. ... Another approach to personal identity would challenge this judgment, however. On this view, substantial memory loss and other psychological changes may produce a new person, whose connection to the earlier person could be less strong, indeed, could be no stronger than that between you and me.

This part of Dresser’s argument is thus aimed at Dworkin’s assumption that the author of the advance directive and the later patient are one and the same person. If this assumption is faulty, it seems, his particular defence of the moral authority of the advance directive – in a case such as Margo’s – fails. If we are right to believe that demented pre-demented Margo is a different person from the patient to whom the advance directive supposedly applies, it seems difficult to appeal to autonomy in defence of granting such a directive moral authority, as autonomy is usually considered to be a right to decide only over oneself. Also, if the author of the advance directive and the later patient are two different persons, it seems difficult to claim, as Dworkin does, that the critical interests of pre-demented Margo are also the critical best interests of demented Margo.

In the advance directives debate, however, Dresser’s argument is often considered not as a mere response to Dworkin’s particular defence of abiding by advance directives, but as a more general argument. As such, it is formulated roughly like this:

---

15 Ibid. p. 6.
16 Ibid.
17 Both Dworkin and other proponents of precedent autonomy seem to agree with this limitation of their precedent autonomy accounts.
(1) One person’s advance directive has no moral authority to determine the care of a different (numerically distinct) individual.

(2) Lack of sufficient psychological continuity (due, for example, to advanced dementia) between the person who issued the advance directive and the later incapacitated patient, implies that they are different (numerically distinct) individuals.

(3) Thus, when there is lack of sufficient psychological continuity between the person who issued the advance directive and the later patient, the advance directive has no moral authority.

Although it is difficult to say whether or not this is a correct reformulation of Dresser’s argument, this is (roughly) the way that it has figured in the debate that has followed from her response to Dworkin’s argument. This formulation is thus, more precisely, the argument I will refer to as the Objection from Personal Identity.

Sometimes, however, the fact that the argument is taken out of its original context has caused some unnecessary confusion. For example, it is sometimes argued that it is rather obvious that the Objection from Personal Identity fails because there are cases where it seems obvious that lack of psychological continuity (and, if we accept premise (2), identity) doesn’t threaten the moral authority of an advance directive at all.

Consider for example a case where a patient has suffered such severe neurological damage that he is rendered permanently unconscious and, further, that he had formulated an advance directive suggesting that life-sustaining measures be withheld or discontinued under such circumstances. In this case, it seems, we have even greater reason than we do in a case such as Margo’s to question the assumption that the permanently unconscious patient is the same individual as the person who formulated the advance directive (there can be no psychological continuity between a person and a being with no mental life at all), yet we do not think it particularly problematic to abide by the advance directive. In fact, we don’t seem to think it morally problematic at all. This seemingly shows, then, that the proponent of the Objection from Personal Identity is mistaken in claiming that lacking psychological continuity threatens the moral authority of the advance directive.

Although this argument may seem sound, and we agree with the claim that abiding by an advance directive in a case such as that just described is not morally problematic, it clearly overlooks that Dresser’s argument (and, I will assume, the Objection from Personal Identity) only concerns cases where there is a conflict between doing what the advance directive prescribes and doing what is in a patient’s current best interest. The way of analysing the supposed problem brought on by the suggestion that the author of
the advance directive is numerically distinct from the later patient, then, is in terms of conflicting interests: if we accept the often-made assumption that the author of the advance directive is the same person as the later dementia patient, the conflict that will have to be resolved is one between temporally distinct interests of the same person – let us call this an “intra-personal” conflict. However, if we are right in questioning this assumption, the conflict is rather “inter-personal”, i.e. one between the interests of two distinct individuals. The reason why non-identity between the author of the advance directive and the later patient in the case above isn’t a case where we would question the moral authority of the advance directive, could thus be explained by reference to the fact that there is no conflict of interests to begin with. However, when there is such a conflict – as there seemingly is in the case of Margo – it does seem to matter whether this conflict is intra-personal or inter-personal.

Questions and purpose

As I have already mentioned, the Objection from Personal Identity has prompted a large and rather diverse debate, but one where a vast majority of the debaters have argued that it is an objection we need not take very seriously. The general aim of this book is to critically examine some of these arguments, and to answer the question of whether or not the Objection from Personal Identity really is an argument that we (for some reason or another) have good reason to dismiss.

Responses to the argument

In this section, let us briefly consider the different kinds of responses that the Objection from Personal identity has prompted. Some – but not all – of these arguments will be considered at more depth in the chapters to come. Let us start, however, by recalling the formulated version of the Objection from Personal Identity:

1. One person’s advance directive has no moral authority to determine the care of a different (numerically distinct) individual.

2. Lack of sufficient psychological continuity (due, for example, to advanced dementia) between the person who issued the advance directive and the later incapacitated patient, implies that they are different (numerically distinct) individuals.
Thus, when there is lack of sufficient psychological continuity between the person who issued the advance directive and the later patient, the advance directive has no moral authority.

Even though the advance directives debate is diverse, it could perhaps be said that most of the arguments that have been forwarded against the Objection from Personal Identity roughly fall into one out of two categories: whereas some have dismissed the argument on the basis of premise (1) – i.e. the claim that one person’s advance directive has no moral authority to determine the care of a different (numerically distinct) individual, others have dismissed it because they have found some reason to doubt the truth of premise (2) – i.e. the claim that lack of sufficient psychological continuity between the person who issued the advance directive and the later incapacitated patient, implies that they are different (numerically distinct) individuals. Let us start with some brief words on the latter category.

In the contemporary personal identity debate there are two very general views that have been taken to have some relevance in relation to ethics: psychology-based theories and biology-based theories. Biology-based theories of personal identity claim, roughly, that an individual at one time is identical to an individual at a different time only if they are “biologically continuous” whereas psychology-based theories roughly claim that psychological continuity is a necessary condition for personal identity over time.

According to the latter line of thought, psychological continuity consists in overlapping chains of psychological connectedness. In turn, connectedness consists in direct psychological connections like memories, intentions, beliefs, goals, desires and similarity of character – most, or at least some of which seem threatened by advancing dementia. Thus, if we accept a psychology-based theory of personal identity rather than a biology-based theory, it seems there would be good reason to question the assumption that predemented Margo is the same individual as demented Margo.

Unsurprisingly, then, the second premise of the Objection from Personal Identity has come in for some criticism from adherents to a biological approach to personal identity. The psychology-based theories of personal identity, these critics have claimed, ought to be rejected and if their arguments are convincing, the Objection from Personal Identity is seriously undermined.

There is, however, yet another view on personal identity that has played some part in the advance directives debate, namely that of narrative identity. According to the narrative theorist, it is not (strict) identity or persistence that is germane to our moral and practical concerns, but rather the answer to the Characterisation Question. That is, the answer to the question of which actions, experiences values etc. are properly attributable to a given
person – a question that the narrative theorist usually answers by reference to a correct incorporation into a person’s self-told story of his/her life.18

In the advance directives debate, most ethicists who appeal to a narrative view of identity conclude that this view offers support for the authority of advance directives – even when there is lack of psychological continuity between the author and the later patient.19 Their argument to this affect usually builds on two distinct claims.20 One is, as we have seen, that (strict) identity (persistence) does not matter to the moral authority of advance directives. The second claim is that even if there is lack of psychological continuity between the author of the directive and the later patient, we can defend the directive’s authority by appeals to “surviving interests.”

As we shall see further on in this book, I believe that the narrative theorist is warranted in his/her first claim: it does seem difficult to reconcile the logical form of identity with what matters in relation to our moral and practical concerns. We are also justified, I will later suggest, to assume that appeals can be made to surviving interests. Both of these claims, however, can seemingly be supported without appeals to narrative identity, which is why I will not, in this book, consider this particular view on identity alongside those of the psychological and the biological theorists.

Now, even if we accept the second premise of the Objection from Personal Identity (i.e. that lack of sufficient psychological continuity between the person who issued the advance directive and the later incapacitated patient, implies that they are numerically distinct individuals), there are numerous ways to question the first premise. Let us briefly consider some such alternatives.

As we have seen in relation to the narrative theorist’s argument above, we could, for example, question the assumption that identity (understood as persistence) is a necessary condition to the moral authority of advance directives, and suggest that it is some other relation that matters instead.

Yet another possibility is to accept both that psychological continuity is a necessary condition for personal identity, and that it is identity that matters to our moral and practical concerns, yet argue that identity is not a necessary condition for the moral authority of advance directives because there are possible normative justifications of granting advance directives such authority even if the author and the later patient are numerically distinct.

To mention just a few versions of this possibility, it could be argued that even if the author of the advance directive is not identical to the later dementia patient, she is at least akin to her closest relative, in which case her earlier preferences ought to be granted at least some authority.21 It also seems possi-

20 See e.g. Blustein (1999).
21 See e.g Buford (2008) and Luttrell and Sommerville (1996).
ble to refer to some right of the former person that is not restricted to one’s own person such as a right to one’s body or the right to one’s property. Or one might suggest that there are consequentialist reasons for abiding by advance directives – even in cases where we have good reason to believe that the non-identity claim is true.

What conclusion we ought to draw in relation to advance directives, then, seems to depend on the plausibility of such alternative justifications. Can they provide sufficient support for the claim that in an inter-personal conflict between the interests of the current patient and the interests of the former person, we should give the latter priority?

In this book, I will explore what I believe to be the most important of these possible arguments intended to undermine the Objection from Personal Identity. Before moving on to my main arguments, however, I will make some remarks regarding the method employed throughout this essay.

Methodological remarks

This book is intended as a contribution to the field of applied ethics, which is the area of moral philosophy that discusses concrete or particular moral problems. To be more precise, it also belongs to the area of medical ethics, which discusses moral problems that arise within medicine. The area of applied ethics is often distinguished from two other moral philosophical areas: normative ethics and metaethics. In normative ethics, philosophers discuss general theories of what makes an action right, a state of affairs worthy, and so on. In metaethics, philosophers discuss the meaning of moral terms, for instance, “right”, “wrong”, “good”, “bad” and so on, and whether moral judgments can be true or false or at least, rational or well-founded.

Even if these three fields of moral philosophy or ethics – applied ethics, normative ethics and metaethics – are commonly presented as three distinguishable fields there are no sharp boundaries between them. Applied ethics, for example, often rely on certain metaethical standpoints such as the belief that moral judgments require justification and can be open to criticism (which is an assumption I will make throughout this book). Moreover, standpoints in applied ethics are of relevance to our standpoints in normative theory, and vice versa.

This last idea is captured by the justificatory ideal known as reflective equilibrium, in which our more particular moral judgments (those that are of main interest in applied ethics) and our more general ones (those that are of main interest in normative ethics) cohere in certain ways.23

---

22 Such a suggestion lies at the heart of Buchanan and Brock’s argument, which will be closely examined in chapter five.

23 The term was originally introduced by Rawls (1972).
This ideal, then, opens up the possibility of criticising more general moral principles because they are at odds with our considered, more or less particular moral judgments. One particular effect of this possibility has been the widespread employment in philosophy of so-called thought experiments. These experiments are constructed so as to isolate the features of a situation that one believes to be relevant and to “tease out” certain intuitions. Such experiments are frequently used in the advance directives debate, and will therefore occur frequently throughout this book. Let us therefore take a closer look at how they may be used and how one might criticise them.

One famous thought experiment in medical ethics is Judith Jarvis Thomson’s Violinist.24 Thomson’s thought experiment is aimed at a popular anti-abortion argument that goes something like this. The foetus is an innocent person with a right to life. Abortion is the deliberate and active killing of a foetus and is therefore morally wrong. Now, Thomson asks her readers to ponder the following scenario. A famous violinist falls into a coma and the society of music lovers decides to do whatever they can to save him. From medical records they find out that you and you alone can save the violinist’s life by being hooked up to his circulatory system for nine months. The music lovers break into your home while you are asleep and hook the unconscious (and unknowing) violinist to you. Now, when you wake up you may want to unhook the violinist, but at the prospect of such an attempt you are faced with the following argument forwarded by the music lovers: the violinist is an innocent person with a right to life and unhooking him will result in his death, which means, in effect, that you are deliberately and actively killing him. Therefore, unhooking him is morally wrong.

Naturally, Thomson’s argument is not very realistic, yet it seems to do some work in relation to her intention of showing that the fact of the foetus being innocent and having a right to life does not necessarily mean that one is morally obliged to provide what is needed to sustain its life. If it seems morally legitimate to unhook oneself from the violinist (thereby causing his death), it may also be legitimate to have an abortion.

Thought experiments like the one above can of course be criticised. For example, one might want to argue that Thomson’s experiment leaves out certain relevant features in a pregnancy – for example the special relationship between the woman and the foetus. If such criticism is legitimate one might want to revise the thought experiment to see if one would reach a different outcome. It is important to note also that intuitions themselves are not exempt from criticism or revision. For example one might agree that our immediate intuition in the violinist case is that one would be morally permitted to unhook oneself but that this intuition ought to be reconsidered – perhaps we are morally obliged to save the life of the violinist by remaining hooked to him for the nine months to come. Perhaps one might even argue

---

24 Thomson (1971).
that certain features of the thought experiment obscure the “correct” intuition in this case.

Even if intuitions – both on a more general level and in relation to more particular (real-life or strictly hypothetical) cases – can be revised, they are, in many respects the raw data of moral philosophy. Intuitions that support our moral theories or principles are, in a way, the only “evidence” that we have, and therefore it should come as no surprise that intuitions (and thought experiments) have come to play a large role in the moral philosophical debate. The same thing is true in the other philosophical field that will play a great role in this book: namely that of personal identity.

The philosophical field of personal identity deals with questions that arise about ourselves by virtue of our being persons. These include questions such as: what am I? When did I begin? What is it to be a person? When will I cease to exist? What does it take for the same person to exist at different times? Etc. Even if this is an area of philosophy not usually included in moral philosophy, but rather considered to be part of the philosophical field of metaphysics, the answers to such questions are often believed to have a great deal of relevance in relation to ethics. If we believe, for example, that some particular rights can be ascribed only to persons, we need to know what kind of beings persons are (i.e. what the necessary and sufficient conditions are for being one), we might also want the necessary and sufficient conditions of personhood to at least partly explain why it is that persons have certain rights that other beings do not.

The advance directives debate is one part of applied ethics where questions of identity are often believed to play a predominant role. As we have already seen, it is often assumed that the more particular question of interest is what it takes for a particular person to persist over time. We have also seen already that there are several different kinds of answers to this question. So, the, how can we try to justify a particular answer to this question?

Well, the method here is very similar to that which we use in our ethical enquiries: we see how it matches our intuitions. Would I survive being brainwashed? Would I survive if all of my body parts were successively replaced with mechanical (and non-biological) body-parts? Here, again, our answer seems to depend on our intuitions. And again, these pre-theoretic beliefs or intuitions might have to be revised – perhaps because they cannot be reconciled with what we believe to be the best available theory of personal identity.

For the purposes of this book it seems a few comments are necessary also when it comes to the relationship between ethics and personal identity. Firstly, it seems that some of our arguments for or against a particular view of personal identity derive from strictly moral considerations. Take, for instance, a case where your cerebrum (with your memories, intentions, character traits etc.) is put into my head and vice versa. In an imagined case like this, many people seem to believe that the person ending up in your body
(i.e. the person with what used to be my mental content) would be the one who deserves to be punished for any crimes that I have committed (this person is, after all, the one who has current memories of committing them), and likewise it seems to be this person who, for example, deserves compensation for wrongs that were done to me before the transplant. These kinds of ethical or moral considerations, it is then often assumed, argue in favour of the view that I would actually survive the transplant in your body.

Although ethical intuitions are not the only ones by which we measure the plausibility of a particular view on personal identity (whether or not we survive a particular ordeal, for example, seems to have little to do with ethical considerations), it doesn’t seem to be the case that considerations of personal identity are completely separate from those of ethics – we cannot, it seems, independently formulate a theory of personal identity that we then can “plug into” different normative theories and simply see what the outcome of this operation would be.

A complicating factor is that while it is often assumed in the debate that we can work out or identify the correct theory of personal identity and then apply it wherever needed to the world of ethics, there are (methodological) alternatives. Firstly, one might question the way that the relation between the two philosophical areas is supposed to run from metaphysics to ethics, assuming that the former is somehow ontologically prior. A possible alternative to this view would be to suggest the reverse relation, that is, assuming that ethics is methodologically prior to (informs, constrains or renders practically moot) our metaphysics. One might even suggest that neither of the above mentioned methodological approaches is correct, but rather suggest that we ought to build up both kinds of theory in light of the other, perhaps via a sort of reflective equilibrium.

Another frequent assumption in the debate is that the relation between the (correct) theory of personal identity and the ethical field is univocal, i.e. that there is one relation running from our theory of personal identity to all of our person-related practical concerns. This, however, is not necessarily true. One theory of personal identity might explain, for example, prudence, whereas another might be more in line with our intuitions with regards to desert, or compensation.

Although I will not have the possibility of arguing in favour (or disfavour) of these different possible approaches in this book, it could be worth noting that some of the disagreement we will later consider probably stems from the fact that there are different views concerning the correct methodology, or “fit” between ethics and metaphysics. My own view is that one cannot exclude from the outset that revisions may go in both ways: we may have to revise metaphysical views in light of our moral considerations, but also may
have to revise our moral judgments in light of metaphysical considerations. In this sense, I favour what has been called a wide reflective equilibrium.25

Overview of this book

This book will have the following structure.

Chapter Two. As we have already seen, the metaphysical debate of personal identity has recently been dominated by two incompatible views – the Psychological View, which claims that persons persist by virtue of psychological continuity, and Animalism, which claims that we are essentially animals and that we persist by virtue of our biology. We have also seen that the second premise of the Objection from Personal Identity presupposes the Psychological View.

In this chapter I will thus address some arguments – recently forwarded by philosophers with an Animalist bent of mind – intended to undermine the Psychological View, hence the Objection from Personal Identity. I will argue that these arguments are unconvincing, and that the Psychological View has more to commend it than animalists usually maintain.

Chapter Three. However, there is yet another possible argument with which to charge the proponents of the Objection from Personal Identity. This is to question the assumption that identity matters at all to the moral authority of advance directives. That identity matters in this way, I will argue, can reasonably be denied. However, if identity is not a necessary condition for the moral authority it seems overwhelmingly natural to assume that psychological continuity is. In practice, then, this charge doesn’t seem to matter much: the relation that is a necessary condition to the moral authority of advance directives — be it identity or psychological continuity — seems to be threatened by severe and permanent neurological damage, for example that due to advanced dementia, or so I will argue.

Chapter Four. Another question, often considered to be of relevance to the moral authority of advance directives, is whether or not our interests can “survive” us. In other words, whether or not events that happen after our death can affect the value of our already completed life. If the value of our lives cannot be so affected, it seems this would strengthen the case for the proponent of the Objection from Personal Identity: why, it could reasonably be asked, should we grant the former person’s interests priority (over those of the current patient’s) when doing so is of no value to him or her?

In this chapter, however, I will argue in favour of the view that we can retrospectively affect the value of a former person’s life. It will also be noted, however, that even if appeals to surviving interests enable a case for arguing that advance directives have moral authority, it doesn’t seem to settle the case for authoritative advance directives — it merely allows for the possibility of a genuine conflict between the “surviving” interests of the former person and the patient’s current ones. Whose interest that ought to be given priority, it seems, is till an open question.

Chapter Five. In this chapter I will start by considering an influential argument by philosophers Allen Buchanan and Dan Brock. Buchanan and Brock presuppose both that psychological continuity is a necessary condition for personal identity and that a case can be made for surviving interest, hence the possibility of what I have labelled an inter-personal conflict of interest between the author of the advance directive and the later dementia patient. This conflict, Buchanan and Brock argues, ought to be solved in favour of the former patient, and this is so because lack of psychological continuity sufficient to undermine identity is also very likely to imply that the demented patient is a non-person. And if she is, their argument continues, we could support the moral authority of the advance directive by appeals to a “right of disposal.”

It is very unclear, however, what the suggested “right of disposal” really entails. Even non-persons, it seems, have interests and it is not obvious from Buchanan and Brock’s argument that they themselves believe that these can be overridden in favour of the former person’s interests.

One possibility for Buchanan and Brock, if they want to maintain their belief that advance directives ought to be granted moral authority at least under ordinary circumstances, is to argue that the supposed non-personhood of the dementia patient also implies that she is not harmed by a non-treatment decision. Such an argument has been developed by Helga Kuhse and therefore we will turn to her argument in the latter part of chapter five, where I argue that Kuhse’s argument is flawed because it rests on too narrow an account of why death is harmful to its victim.

Chapter Six. As we saw in the last chapter, Buchanan and Brock defend the moral authority of advance directives by appeals to a “right of disposal”. In this chapter I will consider the less explored possibility of a consequentialist defence. I will argue that even if an established practice where advance directives are respected is likely to have some good consequences, there is also reason to believe that cases such as Margo’s — where there is a conflict between abiding by the advance directive and acting in the current patient’s interest, and where there is lack of psychological continuity between them — are cases that should be exempt if we want an established practice that maximises utility.
Chapter Seven. Here I will recapitulate the arguments made in earlier chapters and make some concluding remarks. In summary, I argue that there is reason to take the Objection from Personal Identity seriously. Contrary to what a vast majority of ethicists in the advance directives debate have argued, I maintain that under certain circumstances at least, we ought to override the advance directive in favour of the interests of the current patient.
2. Personal identity: biology over psychology?

Introduction
As I mentioned in the last chapter there are basically two general views in the contemporary personal identity debate that have been taken to be of some relevance in relation to ethics in general and the advance directives debate in particular: psychology-based theories and biology-based theories. Whereas the proponent of the Objection from Personal Identity – who believes it possible that pre-demented Margo and demented Margo are different persons – seemingly must appeal to some version of the psychology-based theory of personal identity, biology-based theories seem to block the same objection as they suggest that our persistence conditions are biological rather than psychological. If, then, some more particular version of the biology-based theory of personal identity is true, it seems that, the Objection from Personal Identity cannot even get off the ground, as there is simply no reason to question the assumption that pre-demented Margo and the later demented Margo are one and the same person.

David DeGrazia is an influential participant in the advance directives debate who has forwarded an argument against the Objection from Personal Identity along these lines. The main thrust of DeGrazia’s argument is that the adherent of the Objection from Personal Identity mistakenly presupposes that beings like you, me and Margo persist by virtue of our psychology. DeGrazia then continues by presenting some familiar arguments aimed to show why we ought to reject this view. If we ought to favour any essentialist view, DeGrazia argues, we ought to opt for Animalism, i.e. the view that we are essentially animals and persist by virtue of our biology.

In this chapter, I intend to defend the Psychological View (i.e. the view that we – persons – persist by virtue of our psychology) against some recent arguments forwarded by David DeGrazia and Eric Olson, both well-known adherents of the view that we persist by virtue of our biology. These arguments, I will try to show, either fail or present equally troubling implications.

---

27 Olson (1997).
for Animalism, the specific biological approach Olson (like DeGrazia) favours.

Identity: two rival views

On the face of it, it may seem absurd to claim that pre-demented Margo and demented Margo could be two different individuals – just by looking at demented Margo, you might say, we can safely conclude that she is the same individual as pre-demented Margo. Demented Margo might have a few more wrinkles and her hair may have a few more strands of grey, but apart from that the similarities in appearance ought to be enough to convince us. Nevertheless, this “absurdity” seems to be a possibility according to one of the dominating views of what the necessary and sufficient conditions are for a person to persist over time – the Psychological View.

The Psychological View claims that persons persist by virtue of their psychology, and with this view our persistence conditions are usually phrased in terms of psychological continuity: a person at time 1 and a person at a later time 2 are the same person if and only if there is a sufficient level of psychological continuity between them.\(^2\)

Usually, psychology-based theories of personal identity (such as the psychological continuity theory) are taken to fit particularly well with our moral and practical concerns. Imagine, for example, that sometime in the future it will be possible for a computer to record a person’s full mental contents and transfer them to a different brain in a different body. Let us further imagine that a person with a physically wasting disease could be given the following choice by the healthcare professionals at the hospital: “unfortunately, we can do little to prevent your body from deteriorating and, ultimately, dying, but we can record your full mental contents and transfer them to another (vacant) brain and body. Your original body will be destroyed in the process. If you consent to this transfer, you will be sedated and when you wake up you will find yourself in a new body but with exactly the same memories, intentions, character traits etc. that you had just before we put you to sleep.”

Now, the imaginary healthcare personnel clearly assume that our imaginary patient would survive these events – ought the patient to be sceptical of this assumption? Most people, it seems, would agree with the healthcare professionals in this matter: the case described above is indeed a case of survival, perhaps best described as a full body transplant. The person waking up after the “transplant” is the same individual as the one with the deteriorat-

\(^2\) Traditionally, adherents of the Psychological View have also added a “non-branching clause” to ensure that one person cannot be identical with two numerically distinct future persons. I will return to this matter in the next chapter.
ing disease. Our metaphysical intuitions, in this case, thus seem to support the Psychological View.

The belief that the post-transplant person is (numerically) identical to the pre-transplant person would also explain why most of us would be quite willing to agree, for example, that the patient ought rightfully to be worried if the healthcare personnel warned him that the first few days after the “transplant” would be excruciatingly painful – and worried not only in the way that we sometimes worry about other people’s pain, but rather in a self-interested way. Also, we wouldn’t think it strange that the post-transplant person could rightfully be punished for some acts that the pre-transplant person did, nor would we think it any strange to compensate the post-transplant person for losses made by the pre-transplant person.

All this, it seems, adds up to a fairly convincing case for the Psychological View – when the body and the psyche come apart, the person seemingly goes where his psychology goes. The Psychological View, then, seems to capture the practical importance that we attach to identity, and this is probably the main reason why it has proved so popular among identity theorists.

The Psychological View, however, is at odds with a radically different – but increasingly popular – view, namely Animalism. According to Animalists we (persons) are essentially animals, and because we are essentially animals we have the persistence conditions of animals. The persistence conditions of animals, the Animalist further claims, have nothing to do with psychology but are rather biological in kind.

According to Eric Olson – one of the most well-known proponents of Animalism – a general account of identity would go along something like the following lines: an individual (animal) at time 1 and an individual (animal) at a later time 2 is the same individual if and only if the vital biological functions that P has at time 2 are causally continuous, in the appropriate way, with those that P has at time 1. This view would thus support the intuition that we can be pretty sure that pre-demented Margo and demented Margo are not two distinct beings.

One of the problems with Animalism, however, is that it is open to the objection that it fits poorly with our practical concerns. In the transplant case described above, for example, the animalist would have to claim that the above case cannot correctly be described as one of survival – the person with the physically wasting disease cannot survive the destruction of his/her body and the person who exists at the “other side” of the transplant is not the same individual as the pre-transplant person.

29 It should perhaps be noted that this difference is qualitative and not one of degree, sometimes we worry more about another person’s pain than we do about our own.
Now, most animalists recognise that their view is open to this kind of charge, but there are different possible replies here. Whereas some animalists seem inclined to cling to the belief that identity (biological continuity) matters in relation to our moral and practical concerns, others deny that identity has this purported fit with our moral and practical concerns at all. By the latter view, it seems, good reason to embrace a biology-based theory of personal identity is also good reason to reject the belief that identity matters. I will return to this view in the next chapter, but in this chapter I will consider some of the arguments typically forwarded by Animalists against the Psychological View. Are there really overwhelming metaphysical reasons to adopt Animalism and reject the claim that we persist by virtue of our psychology?

Infants, foetuses and corpses

One common argument against the Psychological View is that it implies the absurdity that we were never born. Says DeGrazia:

If we are essentially persons, we cannot ever exist as nonpersons. But newborns are nonpersons, because they lack the complex forms of consciousness that are necessary for personhood. Thus, if we are essentially persons, then we were never born.

This argument of DeGrazia’s is very similar to an argument famously advanced by Eric Olson. The belief that you and I persist by virtue of our psychology, Olson argues, implies that neither you nor I was ever a foetus. This is so because a human foetus less than ten weeks old does not, as far as we know, have a cerebrum developed enough to function as an organ of thought or sensation, and it is obvious that the Psychological View implies that you or I could not be identical with a being that has no mental contents, or capacities at all.

These kinds of arguments have intuitive appeal, for sure. We do speak of ourselves as having been born somewhere at a particular time and as once having been foetuses in our mother’s belly, and if it is an implication of the Psychological View that we are mistaken in our beliefs about this, perhaps it ought to be rejected. I will argue, however, that these kinds of arguments can

---

32 See e.g. DeGrazia (2005).
33 See e.g. Olson (1997).
35 It is not until at least six weeks after fertilisation that a human embryo even starts to develop the rudiments of a cerebrum. See Olson (1997).
be refuted and that the alternative essentialist view that both DeGrazia and Olson favour is likely to face similar difficulties.

One thing to note about the two arguments above is that they seem to have a somewhat different intuitive strength. It seems more counterintuitive to claim that you and I were never born than it would be to say that you and I were never foetuses, just as it would be more counterintuitive to say that you and I were never foetuses than it would be to say that you and I were never a lump of pre-embryotic cells. The argument DeGrazia makes, i.e. that Person Essentialism implies that neither you nor I was ever born is then, in this sense, stronger than the one Olson makes. If it could convincingly be argued that Person Essentialism and/or the Psychological View precludes the possibility of me being identical with my mother’s new-born baby this should at least make the adherent of such a view slightly uneasy.

DeGrazia’s argument, however, relies on the rather controversial claim that a person-essentialist and/or an adherent of the Psychological View must hold the belief that someone who has “the complex forms of consciousness that are necessary for personhood” cannot be identical with a being that doesn’t (e.g. a new born baby). This, however, is not entirely true. There are many different versions of the Psychological View and by some of them, at least, there could quite possibly be a sufficient degree of psychological continuity between someone who has these complex forms of consciousness and a being that doesn’t. For example, a person could start his/her existence as he/she developed the capacity for awareness rather than at the time when he/she developed rational thinking or any other such complex for of consciousness. Therefore it is not entirely correct to say that the Psychological View precludes the possibility of me being identical to the infant that was born on the 2nd of May 1979, even if there might be versions of it that do. But what about the somewhat less counterintuitive claim that I was never a foetus?

Eric Olson readily recognises that there are several possible ways to explicate the Psychological View and that some of these views allow me to be identical with my mother’s new-born baby. To put forward a similar argument then, but one that is supposed to show the implausibility of all accounts of personal identity based on psychology, he claims that none of these views are consistent with my having been a foetus.

However, even Olson seems to be somewhat mistaken. While it is true that the Psychological View precludes the possibility of me being identical to something that doesn’t have any kind of psychological life, it is far from

---

36 As we shall later come to see, Olson’s own view implies that we never were that lump of pre-embryotic cells.

37 As an example, Olson refers to the views of Michael Lockwood (1985), who suggests that I persist as long as my capacity for awareness is preserved.

certain that a foetus doesn’t have what it takes. For example, Olson is keen to point out that a foetus isn’t believed to be minimally sentient before thirty-two weeks after fertilisation, but surely this doesn’t mean that there is no mental life at any stage of a foetal life? Thus, Olson’s argument doesn’t show that the Psychological theorist is barred from saying that you and I were once foetuses – only that he/she is barred from saying that neither you nor I was once a foetus at those very early stages of its development when the prerequisites for any mental life are not yet in place and intuitively, at least, this does seem to make a difference.

However, there is yet another possible reply for the adherent of the Psychological View and, I take it, for the Person Essentialist. This would be to admit that it sounds very strange indeed to say that we were never born or were never foetuses, but that is because there is an ambiguity in the way we speak; the claim that you and I were once foetuses does not necessarily mean that we are identical with that infant or foetus, rather our linguistic intuitions could be explained in something like the following way (surprisingly suggested by Olson himself):

When we learned at mother’s knee that each of us spent time inside the womb before we were born, or that human fetuses and infants develop into and become adult human beings, perhaps we did not learn that each of us is numerically identical with a fetus or an infant, but only that a fetus or infant, as it develops, gives rise to or produces a person. There is nothing absurd about saying that the sparks became a conflagration, or that Slovakia and the Czech Republic were once a single country; it would certainly be wrong to take this as implying that one thing is numerically identical with more than one thing. This shows that there is a sense in which one thing can “become” something numerically different from it, and a sense in which one thing can “once have been” another. The same goes for the claim that a fetus is a potential person. A wooden house is a potential pile of ashes; but that doesn’t imply that any one thing could be first a house and later a pile of ashes. Thus, even if none of us is numerically identical with a fetus, and even if no fetus is ever numerically identical with a person - even if no one thing is ever first a fetus and later a person - it is not clear that this conflicts with anything established by science, or with anything that any enlightened person believes. It is still true that a fetus “develops into” and “becomes” a person in the weak sense that there is a continuous process of self-directed growth that begins with a fetus and ends with a person. Why suppose that our mothers and our biology teachers meant to tell us anything more than that?

This, I think, is indeed a very well-formulated response to the objection that it “sounds wrong” to say that neither you nor I was ever a foetus. If my house were to burn down I might point to the pile of ashes and exclaim: “look at the state of my house!” No one, I think, would believe that I’m thereby expressing the mistaken belief that what was once my house is nu-

39 Olson (1997), pp. 77-78.
merically identical to the pile of ashes now in front of me, and this goes to show that our language is not as precise as philosophers sometimes want it to be. Accordingly, when we say that we were once in our mother’s belly, we don’t necessarily mean to say that we are identical to the foetus in question that was born (at a particular place and a particular time), but something much less determinate. Perhaps we mean to say merely something along the lines that I have developed from that foetus, and if this is the belief that we express, the Psychological View doesn’t make our claims mistaken at all.

Eric Olson who – in spite of his Animalist inclinations – has formulated the very plausible response above, nevertheless insists that this is a mistaken linguistic analysis: By claims such as “I was once born” or “I was once a foetus” we mean nothing less than that we are identical with that infant or foetus. According to Olson, however, we need not rest our case on linguistic evidence since the claim that nothing is ever first a foetus and later a person also faces serious metaphysical problems. These can be captured, roughly, by asking these kinds of questions. If the foetus that was once in my mother’s belly is not numerically identical to me, what has become of it? What happened to it when I came into existence? Is it still with us?40 This challenge, intended as an argument against the Psychological View, is usually known as the Foetus Problem.

DeGrazia, who is admittedly influenced by Olson’s arguments, also believes this to be a serious challenge to the Psychological View (hence to the proponent of the Objection from Personal Identity):

The second reason to reject person essentialism [as well as the Psychological View] is that it leaves unexplained what happened to this predecessor. Did it die? We are not aware of any kind of death that regularly occurs in the first two years of human life .... Perhaps the predecessor merely disappeared, rather than dying. Yet it was unquestionably a human organism, if it existed at all, and organisms die. So how could it go out of existence without dying?41

Now, it does seem strange indeed to say of the foetus that it dies to leave room for its successor. In fact, if we understand death as the irreversible cessation of all vital functions, it is obviously false to say that the foetus dies. There is no irreversible loss of life in the process where a foetus develops the capacities needed to think. Nor is there such a thing as a “corpse”, i.e. some lifeless material object where the living being used to be. Thus, to say that the foetus dies does not only seem highly remarkable, it seems directly mistaken, but fortunately for the psychological theorist, his/her theory implies no such thing.

Sometimes, the psychological theorist might answer, living things perish or cease to exist without there being a (biological) death involved. Few peo-

40 Ibid. p. 79.
ple, for example, would want to deny that a sperm is a living one-cellled or-
ganism (that is, the kind of being that typically perishes by dying). Yet not
many people would be willing to say that the sperm is numerically identical
with the embryo. Most people, I think, would say that the fusion of the
sperm and the ovum is the start of a new organism – the embryo – and the
embryo is identical neither with the sperm nor with the ovum. If this is a
plausible account of what happens at the very beginning of the life of every
human animal, then it seems to provide us with a perfect example of how a
living thing can perish without there being a (biological) death involved.

But perhaps this is making too little of the Animalist argument. Perhaps
the supposed “strangeness” of the claim that a foetus has to cease existing in
order to make room for a person still remains even if we consider the fact
that sometimes living things actually do seem to cease to exist without there
being any biological death involved. So, then, in what would this strangeness
consist?

According to Olson, the supposed peculiarity in the implications of the
Psychological View lies in the claim that each human fetus must perish in
the act of bringing forth a human being. Olson, however, continues: “Why,
we should want to know, should a fetus perish simply because, in the course
of carrying out the program encoded in its genes, it (or rather its successor)
became able to think? This is not the sort of thing that typically causes
death”.42

Having already argued that it is not very remarkable to claim that some-
ting ceases to exist without dying (in the biological sense of the term), we
can now distinguish two further implications of the Psychological View that
Olson seems to be worried by: one is that it implies that something ceases to
exist simply because it becomes able to think, and the other is that it implies
that something ceases to exist by carrying out the program encoded in its
genes. Both of these things, I admit, might seem strange, but as I will try to
show, Olson’s own view (which DeGrazia to a large extent seems to share)
seems open to very similar objections.

According to Animalism, the foetus is the same numerically identical
being as the later, fully grown human being. However, the animal that is me
did not begin its existence at the time where the sperm and the cell gave rise
to the embryo. Rather, Olson holds the view that a human animal comes into
being roughly fourteen to seventeen days after conception.43 This is so be-
cause at this stage the embryo develops into a multicellular organism, mean-
ing roughly that its cells acquire the ability to direct their growth and other
activities in the way required to be part of an organism. The question is then:
if I started to exist fourteen to seventeen days after conception, what ever
became of the pre-multicellular organism that was there (alive and kicking)

43 Ibid. p. 91.
between conception and my coming into being? In other words, couldn’t there also be a “foetus problem” for the Animalist?

First of all we may point out that even the animalist view suggests that something living (the “pre-embryo”) ceases to exist without anything dying so this couldn’t possibly be what he/she considers to be so strange about the implications of the Psychological View (and as we have seen even an animalist who would prefer to claim that the animal life starts at conception would have the same problem as the “successful” sperm, it seems, ceases to exist without dying). So, what about the supposedly strange thing that something perishes simply by “carrying out the program encoded in its genes?” Well, this is also unlikely to be Olson’s main argument. Surely, the development from being a pre-multicellular organism (or pre-embryo) into being a multicellular organism is one that is encoded in the genes of the pre-multicellular organism: at the very least it seems to follow the path of growth and development that it is “supposed to”. Also, biological death itself is something that is encoded in the genes of a living organism and it would be difficult, therefore, to maintain that there is anything strange about a theory that implies that something perishes simply by carrying out the program encoded in its genes.

Let us now turn, then, to Olson’s concern over the implication that something ceases to exist simply because it becomes able to think. Reasonably, this is the Animalist’s bigger worry. Says Olson: “That one necessarily perishes if one loses one’s ability to think may be false, but at least we can understand why someone might think that it was the case. That something should perish by virtue of gaining that ability is absurd”.44

According to the Animalist View it surely isn’t the gaining of any mental capacities that causes the pre-embryo to perish, but if this was the only point that Olson wanted to make it seems we are somehow right back where we started. The Psychological theorist obviously believes that the gaining of, say, a consciousness is something that matters greatly, and that it is something that matters both when it comes to the extinction and the coming into existence of conscious beings like us – the Animalist does not. But does Olson perhaps mean to say that there is something strange in claiming that the gaining of any kind of capacity should cause something to perish? Again, I admit, such a claim does have a slightly strange ring to it: if anything, desuetude usually involves the loss of capabilities, not the gaining of them. Again, however, the Animalist seems to face a similar problem: isn’t it the case that the lump of cells existing between conception and the coming into being of the animal, perishes because it gains certain capacities? And if so, isn’t this problem comparable to the foetus problem?

According to Olson, the answer to the latter question is no. He believes that his Animalist account of how and when we start to exist can escape a

44 Ibid. p. 80.
foetus problem of its own simply because we do not need to suppose that there is any one persisting object between fertilisation and the time when the human animal comes into being. Up until the stage where the cells acquire the ability to direct their growth in the way required to be a multicellular organism we have only individual cells that divide. Just as “when an amoeba divides into two cells, it does not become a two-celled, spatially scattered object. It ceases to exist and is replaced by two different cells".\(^45\)

Olson’s claim that there really is no one thing before the organism that is identical to the animal comes into existence is in itself, I believe, a rather strange one. Surely there is some entity there in the womb before fourteen to seventeen days after conception? In medical terms this entity is called first a zygote (from conception to about three days after) then a morula (the zygote has now divided so that it is made up of 32 cells), then (usually within the first week after conception), a blastocyst. The blastocyst’s outer cells will form an epithelium, enclosing the inner cell mass and after the phase where it adheres to the wall of the uterus (known as implantation) it turns into what is medically referred to as the foetus.

Apart from the strangeness of having to deny that the blastocyst, for example, is one persisting object Olson’s amoeba analogy is misleading. The individual cells of the blastocyst are not like individual amoeba. Whereas an amoeba is shapeless unicellular organism capable of digestion etc., the individual cells of the blastocyst are rather parts of a whole, enclosed by an epithelium. Before implantation, it is true, the blastocyst is capable of dividing (thus enabling the occurrence of twins or triplets), but surely the possibility of something dividing cannot be seen as a reason to say that before (the mere possibility of) division it is not one thing?

Further, even if we accept Olson’s claim that there is no one persisting object before fourteen to seventeen days after conception, this does not automatically solve the Animalist’s problem of showing that his/her account of our coming into existence does not involve anything so strange as saying that something living perishes as it gains a certain capacity. This is so because even with Olson’s view, it seems true to say of each of the (individual) living cells that perish to make room for the organism that they do so by virtue of gaining some capacity.

Perhaps Olson could reject this claim, however, by saying: yes, it is true that the cells all gain some capacity that they didn’t have before, but it is not this fact that causes them to perish – rather it is the fact that they, so to speak, fuse together into a new organism. But I am not sure how well such a rejoinder would work. In an ordinary case where two cells fuse into one new thing there are other things going on such as at least some mixing of matter (such as the fusion of sperm and ovum) but in this case the lump of cells is already there – tightly packed together and enclosed by the epithelium. In

\(^{45}\) Ibid. p. 93.
fact, the only thing that seems to happen between time 1, where according to Olson we only have a mass of living cells stuck together, and time 2, when the individual cells have perished (to leave room for the new organism), is that the individual cells develops new capacities – in Olson’s own words, they now “begin to grow and function in a coordinated manner”.46

Are we animals?

As I hope the arguments above to have shown, the implication that the foetus perishes to make room for its person successor is not much stranger than the Animalist view that a pre-multicellular organism perishes at the beginning of every animal life. However, one might still think that this is an implication we ought to avoid if we can. Some adherents of the Psychological View have therefore suggested that the early foetus continues to exist, but never comes to be a person. The person, in turn, starts to exist at some later time and “overlaps” the animal throughout his or her existence. This view is sometimes referred to as the constitution view and DeGrazia simply dismisses it as just “too strange to believe”.47 I believe, however, that a constitution view has more to commend it than DeGrazia concedes.48

To make the idea of constitution clear, think of a bronze statue. The hunk of bronze from which the statue is made is not, it seems, identical to the statue – destroying its shape by melting will destroy the statue but not the hunk of bronze. Similarly, a number of threads (arranged in a certain way) can constitute a flag while not being identical to it – taking the threads apart would destroy the flag but not the threads. The hunk of bronze, which constitutes the statue, thus has different persistence conditions from the statue; just as the threads that constitute the flag have different persistence conditions from the flag. In the same way then, we may think of the relationship between a person and an animal as the former being constituted by the latter.

One of the strengths of a constitution view – apart from the fact that it can avoid the foetus problem – is that it can also explain the origins of a person’s (non-person) successor – and so it can provide a riposte to yet another one of DeGrazia’s objections to the Psychological View:

[In the case of pre-demented Margo] was her successor conceived (or otherwise biologically brought into being), say, when she was about fifty-five? We are certainly not aware of any such event that occurs at the onset of dementia

46 Ibid. p. 91.
47 DeGrazia (1999), p. 386. DeGrazia’s, however, considers this approach in greater depth in DeGrazia (2005) although he does reject it.
48 For explications of this view see e.g. Shoemaker (1984) and Baker (1999), (2000).
– or even permanent unconsciousness. Did [she] somehow come into being without being conceived (or otherwise biologically brought into being)?

With the constitution view, the obvious answer is that pre-demented Margo’s successor (the animal) was conceived exactly when we think she was – say about sixty years ago, and that she has been in existence ever since – again just as we would have guessed. There are, however, other problems associated with this view – problems Animalists have been keen to point out. One such problem is that it fails to support the very intuitive belief that we (persons) are animals in a strict sense.

Adherents of the coincidence view have traditionally responded to this challenge by claiming that we are animals in one sense of the term, just not in the sense that we are identical with one. We are animals, these philosophers have suggested, rather in the sense that we share our matter with one. When you look into a mirror it is true that you see an animal. This animal is not identical with you, but it is intimately connected to you: the two of you are made of the very same atoms and occupy exactly the same region of space.

Olson uses the following example to illustrate the gist of this argument:

Compare the truck-driver who says at the weigh station, “I weigh 35 tons.” She means, of course, that she is driving a truck that weighs 35 tons. If you want to establish empirically that I weigh 150 pounds (by setting me on the scale, say), you can show only that I weigh 150 pounds in the sense of standing in some relation empirically indistinguishable from identity with something of that weight. If you want to show that I have the property of weighing 150 pounds in the strict and literal sense – that I have that property in the way that I have the property of thinking, for example – you will need a philosophical argument.

Olson argues that this view runs into serious problems. Firstly, it seems utterly strange to say that there are two beings in the chair that you are now sitting in: you (the person) and the animal intimately connected to you. When you decide that you want to put the kettle on for a cup of tea, two beings get up to put the kettle on etc., that is, one more than expected. Secondly, it seems difficult to argue that the animal is not a person, just as you are. This is so at least if personhood is something that a being has by virtue of her mental capabilities, as is suggested, for example, by a Lockean account of personhood. Such an account would suggest that a being is a person if it is a rational, capable of thought, consciousness, and a certain kind of self-awareness. Insofar as we agree with this basic idea, it seems clear that whatever capabilities you have in that regard, the animal which shares all of your

50 See e.g. Shoemaker (1984).
matter and all of your mental processes must have them too. And if so, what is it that makes you a person (with a person’s persistence conditions) and the animal “merely” an animal (with an animal’s persistence conditions)?

Finally, whenever you (the person) think, “I am a person”, or “I want a cup of tea” so, it would seem, does the animal. But supposedly the animal is mistaken (at least regarding the thought “I am a person” – I take it that the animal too could be up for a cup of tea) and if so, how can you know that you’re not the one who’s mistaken?

However, Olson’s arguments against the suggestion that we are animals in the sense that we coincide (rather than being identical with) the animal we see in the mirror seem to constitute a problem to the Animalist stance as well, but from the other end of the spectrum. This problem, formulated by Sidney Shoemaker, is usually referred to as the Corpse Problem and has apparent similarities to the supposed problem for the adherent of the Psychological View in explaining what happens to the foetus when a person comes into existence.52

Before moving on to the thrust of Shoemaker’s argument we need, however, to be clear about the fact that Animalism is not the view that we (persons) are our bodies. If we were, it would be difficult to maintain, as Animalists typically do, that we cease to exist when we die, i.e. when our purely animal functions – metabolism, the capacity to breathe and the like – cease.53 Animalism is the view that we are animals and that we, as such, persist by virtue of biological, not bodily, features. It seems then that just as the adherent of the Psychological View has to explain what happens to the foetus when the person comes into existence, so the Animalist has to explain the origins of a corpse.

Now, let us return to Shoemaker’s argument. In Shoemaker’s own words, it does seem “overwhelmingly natural to say that what after death is a person’s corpse is something that existed, as that person’s body, prior to death – that scars and tattoos on its surface are traces of things that happened to it earlier”.54 But unless Animalists embrace some idea very similar to the coincidence view, it seems they have to deny that this is the case. This is so because if our bodies are not us, yet share our space and our matter, we ought, in accordance with the arguments advanced by Olson, to be persuaded that we are not essentially animals, but bodies. Here is Shoemaker’s argument:

Assuming that the human animal ceases to exist at death, and assuming that where the human animal is there is a body that does not cease to exist but becomes a corpse, the human animal and the body are nonidentical. They will therefore be coincident entities, one of them perhaps constituting the other.

And then ... the objection that Olson raises to the view that persons and human

52 Shoemaker (1999).
53 Olson (1997), pp. 16-17.
animals are coincident entities will apply equally to Olson’s own position. ... If physical properties determine mental properties, then there is the same case for saying that this bodily entity shares the mental properties of the animal (which on Olson’s view is the person) as there is for saying that, on the coincidence view Olson’s argument is directed against, the human animal shares its mental properties with the person. And, of course, if we can find a plausible way of denying that the bodily entity (the corpse-to-be) shares the mental properties of the person, this may provide us with a way of denying that the human animal shares the mental properties of the person.\(^{55}\)

This argument, I think, should cause some stir within the Animalist camp. If we ought not be convinced of the claim that we (persons) coincide with an animal because of the problems connected to this view, how can the Animalist explain the relationship between an animal and its body? And mustn’t the answer here be some sort of coincidence view?

Whether or not it is right to claim – as Shoemaker does – that the Animalist would have to offer an explanation of the relationship between the animal and its body that would be very similar to the one the proponent of the Psychological View offers concerning the relationship between the person and the animal, we might still be reluctant to accept a view implying that we are animals merely in the sense of coinciding with one. In the last part of this subsection I will suggest that there just might be another possibility, namely to suggest that we (persons) are animals yet deny that this fact determines our persistence conditions. In what follows I will simply raise some concerns about the way that Animalists such as Olson assume animal to be a suitable substance concept (whereas person is not) and that it therefore determines the persistence conditions of any being that is an animal. It seems to me that many of the concerns Olson expresses about the possibility of the concept of person being a substance sortal could also be raised with regard to the concept animal.

Now, Olson seems to hold the view that all animals share the same persistence conditions because animal is a substance concept. A substance concept is the most fundamental answer to the question “What is it?”\(^{56}\) And according to Olson, person is not a suitable substance concept because it is not a natural kind.\(^{57}\) Rather, person is a functional kind: to say that a thing is a person is merely to say something about what it can do.\(^{58}\) On the other hand, Olson continues, animal is a natural kind and as such it is also suitable as a substance concept. Thus, if we are animals this is something that determines

\(^{55}\) Ibid.

\(^{56}\) See Wiggins (1967), p. 980.

\(^{57}\) Here too, Olson relies on Wiggins who claim that only natural kinds are suitable substance concepts.

\(^{58}\) Olson (1997), pp.34-35.
our persistence conditions.\footnote{Ibid. pp. 122-123.} Let us take a closer look, first, at the claim that person is a functional kind whereas animal is not.

The concept of person is, according to Olson, comparable to the concept locomotor, a concept that would be applicable to anything that is capable of moving under its own steam. According to Olson it seems absurd to count locomotor as a substance concept or to say that this concepts refers to a natural kind: if locomotor refers to a kind at all, it is surely a functional kind, not a natural one. But why is this?

Well, according to Olson, there are a few particulars about the concept locomotor that makes it suitable as a functional kind rather than a substance concept. One is that the word is tied to what something can do rather than to what it is. A second reason is that this concept seems to include such a variety of different things. Says Olson:

Different locomotors may have little in common besides the fact that they are locomotors ... On the other hand, many locomotors, such as crabs, have a great deal in common with certain non-locomotors, such as barnacles.\footnote{Ibid. p. 34.}

Now, it is probably true of the concept locomotor that it wouldn’t be a suitable substance concept, and perhaps also true that “person” wouldn’t be. But is the concept “animal” really so much better suited?

According to Olson’s own account an animal, or the more general “living organism”, is “anything that has ... “life-giving” features – metabolism, teleology, organised complexity - and whatever further properties necessarily go along with them, such as self-directed growth and development, an internal genetic plan, low internal entropy, and perhaps the capacity for evolution”.\footnote{Ibid. p. 130.} Couldn’t most of the denominators mentioned above be described as things that certain clusters of cells can do? Sometimes, a cluster of cells can metabolise, i.e. it has the ability to “retain (its) characteristic form and structure despite a constant and rapid exchange of matter and energy with (its) surroundings”; sometimes a cluster of cells can govern, from within itself (though influenced by its surroundings), its “size, shape, internal structure and chemical and chemical composition, temperature, and movements”\footnote{Ibid. p. 127.} and so on. If this isn’t an outlandish way of describing the features specific to animals, animal too seems to be a functional kind.\footnote{A similar argument can be found in Shoemaker (1999).}

When it comes to Olson’s second argument - that different locomotors have very little in common besides the fact that they are locomotors – it is even clearer that the concept animal faces similar difficulties (perhaps even greater difficulties) than the concept person. Yes, there are great differences
between an aeroplane or a battleship and a crab, but aren’t there also significant differences between a human animal and a barnacle? Naturally they both share certain features (e.g. those that support the classification of them both being animals), but why would these shared features imply that they also share the same persistence condition any more than the fact that something can move under its own steam would? Of course, one could argue that the features shared by a model aeroplane and a crab are superficial, e.g. since they are based on very different internal structures, whereas the feature shared by a barnacle and a human is not. But I am not sure what such an argument would amount to, except for restating the intuition one has in favour of Animalism and the belief that the concept of animal is a substance concept.

Conclusion

In this chapter I have argued in favour of the Psychological View (i.e. the belief that persons persist by virtue of their psychology) by responding to some recent objections to this view raised by animalists such as DeGrazia and Olson. Most of these arguments, I have tried to show, either fail or present equally difficult challenges to the view that we are essentially animals (with biological persistence conditions).

First, I argued that the Foetus Problem – i.e. the supposed problem for the adherent of the Psychological View to explain what happened to the foetus when the person came into existence – presents a parallel problem to an Animalist such as Olson. This is so because Olson’s biology-based theory seemingly owes a similar explanation of what happened to the organism that was there between conception and the time that it developed into a multicellular organism (an animal). Even though Olson tries to provide such an explanation, claiming that there is no one thing there up until the stage where the cells acquire the ability to direct their growth in the way required to be a multicellular organism, his argument is wanting. The cells of the blastocyst, it seems, are parts of a whole and not just individual cells that divide.

Secondly, I suggested that the constitution view offers a plausible account of the relationship between a person and an animal: just like a bronze statue has different persistence conditions from the lump of bronze of which it is constituted, a person has different persistence condition from the animal with which he/she is associated. Further, it is likely that the Animalist would need to employ some similar explanation of the relationship between the animal and its body.

Finally, I raised some concerns about the way that Animalists such as Olson assume animal to be a suitable substance concept (whereas person is

64 Olson (1997), p. 34.
not) and that it therefore determines the persistence conditions of any being that is an animal.

From what has been said in this chapter we can hopefully conclude, then, that a case can be made in favour of the Psychological View – recent animalist arguments notwithstanding. If the Psychological View is true, we can also reject DeGrazia’s claim that the Objection from Personal Identity fails because it rests on the mistaken assumption that we (persons) persist by virtue of our psychology (rather than our biology). In the next chapter I will consider yet another possible objection to the Objection from Personal Identity, but one that rather concerns the relationship between identity and our moral/practical concerns. This objection essentially claims that the proponent of the Objection from Personal Identity is mistaken in his/her belief that identity matters at all to the moral authority of advance directives.
3. Is identity really what matters?

Introduction

As we saw in the last chapter the Objection from Personal Identity presupposes a psychology-based view of personal identity. One argument that has been raised against the Objection from Personal Identity, then, suggests that psychology-based views of identity are mistaken. I argued against this claim. The Psychological View, it seems, does not face insurmountable problems, at least not problems that its main rival – Animalism – does not. But let us briefly, for the sake of argument, assume that I am wrong, and that Animalism (or some other biology-based theory) is a superior view of identity.

As we have already seen the main problem with a biological approach to identity is that it fits poorly with our moral and practical concerns. For example, being in a persistent and irreversible vegetative state with no conscious experience whatsoever is tantamount in all relevant aspects to being dead. Further, lacking biological continuity where there is psychological continuity doesn’t seem to matter much to us either. Imagine, for example, having all of your body parts successively replaced by mechanical (or bionic) parts yet throughout the whole process retaining full psychological continuity. Surely, this would not be as bad as death: you would retain all of your memories and would be able to carry out your “life plan” and be able to keep all the relations that are important to you. The fact that your current self is not biologically continuous to the person you used to be wouldn’t seem to matter at all.\(^65\)

This problem has, however, prompted two different kinds of response: whereas some animalists, like DeGrazia, try to cling to the belief that personal identity (i.e. biological continuity) is germane to at least some of our moral and practical concerns,\(^66\) others claim rather that the truth of a biologically-based theory of identity provides us with good reason to question the assumption that identity matters in the way we usually believe it to. This, for

\(^{65}\) This example is from Baker (2000).

\(^{66}\) DeGrazia (2005) believes this view defensible partly because he believes that the relevance of farfetched hypothetical cases – like the transplant case described in the beginning of the last chapter – is dubious, partly because in the world as we know it continuation of biological life is a prerequisite of psychological continuity. See also DeGrazia (2003).
example, seems to be Olson’s view. Olson claims that if it is right that “the relations of practical concerns that typically go along with our identity through time are closely connected with psychological continuity ... then the Biological Approach does have an interesting ethical consequence, namely that those practical relations are not necessarily connected with numerical identity”. 67

In this chapter I shall consider an argument intended to make us question the assumption that identity matters in relation to our moral and practical concerns, even if we reject Animalism in favour of the Psychological View. In fact, one of the most famous arguments to this affect has been put forward by a well-known psychological theorist: Derek Parfit.

Parfit’s views on identity

The belief that identity is not what matters in relation to our moral and practical concerns is most often associated with the work of Derek Parfit. 68 Parfit’s views, however, are in many respects a Lockean account of personal identity (although there are – as we shall see – significant departures), and therefore it seems appropriate to start with some brief words on the views of the 17th century philosopher John Locke. 69

In his time, Locke’s views on identity were revolutionary – whereas most 17th century philosophers believed that the identity of persons was secured by virtue of their consisting in the same immaterial substance like, for example, that of an immaterial self or a soul, Locke suggested that a person at time 1 and a person at a later time 2 are the same person rather by virtue of some relation between them. 70 This relation, Locke further suggested, is one of experience-memory: a person at time 1 and a person at a later time 2 are identical if and only if the person at time 2 remembers being the person at time 1 (i.e. remembers his/her thoughts and experiences).

Now, although Parfit is sympathetic to the view that two person-stages are parts of the same person due merely to some relation between them, he believes Locke’s experience-memory criterion of personal identity to be mistaken because it seemingly implies no one could possibly forget the things they once did. According to Parfit this is clearly mistaken: he is the

69 My short account of Locke’s ideas (and those of his contemporaries) leans heavily on Noonan’s (1989).
70 A bodily account would have been out of the question, as an account of personal identity at that time would have to make sense of the Christian doctrines of, for example, human immortality. See Noonan (1989).
same person as he was this morning yet cannot now remember putting on his shirt.\textsuperscript{71}

To remedy this problem, Parfit suggests that we revise Locke’s view and appeal to the concept of “an overlapping chain of experience-memories”.

Let us say that, between X today and Y twenty years ago, there are direct memory connections if X can now remember having some of the experiences that Y had twenty years ago. On Locke’s view, only this makes X and Y one and the same person. But even if there are no such direct memory connections, there may be continuity of memory between X now and Y twenty years ago. This would be so if between X now and Y at that time there has been an overlapping chain of direct memories. In the case of most adults, there would be such a chain. In each day within the last twenty years, most of these people remembered some of their experiences on the previous day. On the revised version of Locke’s view, some present person X is the same as some past person Y if there is between them continuity of memory.\textsuperscript{72}

Parfit also broadens Locke’s view so as to invoke other relations than those between an experience and a later memory. These kinds of relations he terms “direct connections”; i.e. connections like those between a memory and the experience of which it is a memory, but also like those between an intention and the later act in which it is carried out or simply the persistence of a belief, desire etc. Parfit thus defines two general relations: psychological connectedness and psychological continuity. Whereas psychological connectedness is the holding of particular direct psychological connections, psychological continuity is the holding of overlapping chains of strong connectedness.\textsuperscript{73}

Now, in Parfit’s view it is the former of the two relations (i.e. psychological connectedness) that is the more important one – both in theory and in practice – yet he notes that this relation cannot function as a criterion of personal identity. This is so because identity is a transitive relation, meaning that if it is true that A is identical with B and B is identical with C, then A must be identical with C. But it doesn’t follow from the fact that I am now strongly connected to myself three years ago and that the person three years ago is strongly connected to some person six years ago, that I am now strongly connected to the person six years ago. In fact, it is very unlikely that I am, since between me now and myself six years ago there is a significantly smaller number of direct connections than those that hold over any day in the lives of nearly all adults.\textsuperscript{74} Therefore strong connectedness cannot be the

\begin{footnotes}
\item\textsuperscript{71} Parfit (1984), p. 205.
\item\textsuperscript{72} Ibid.
\item\textsuperscript{73} To constitute strong connectedness, Parfit claims, there has to be at least half the number of “direct connections” that hold over every day in the lives of nearly every actual person.
\item\textsuperscript{74} Parfit (1984), p. 206.
\end{footnotes}
criterion of identity. According to Parfit, however, psychological *continuity* can.

After having made the above revisions of Locke’s view, Parfit formulates his criterion of Personal Identity roughly as follows:

*The Psychological Criterion:* X at time 1 is the same person as Y at time 2 if and only if X is uniquely psychologically continuous with Y, where psychological continuity consists in overlapping chains of strong psychological connectedness, itself consisting in significant numbers of direct psychological connections like memories, intentions, beliefs and similarity of character.\(^7\)

The adding of a “uniqueness” or “non-branching” clause is necessary to make the criterion “fit” the logical form of an identity relation, and again, this has to do with the transitivity of identity (meaning that if A is identical with B and B is identical with C, then A must be identical with C). The problem is that without a uniqueness clause, the Psychological Criterion would allow the possibility of one person splitting into two persons who are both psychologically continuous with the past person yet not psychologically continuous with one another (think of a skilful surgeon who is able to split the brain of one person into two equally psychologically continuous halves and then transplant each half to two different bodies), and this possibility would clearly make the criterion unfit as a criterion of identity.

Another important feature of Parfit’s views on personal identity is that he is a reductionist, i.e. he believes that the facts about persons and personal identity consist in nothing but more particular facts about brains, bodies, and series of interrelated mental and physical events (and denies that there is some “further fact” – like, for example, that of a Cartesian ego or a soul). Although I will not discuss Parfit’s reductionism further here, it does play some part in the way he justifies his conclusion that the identity relation is not what matters in survival. Let us turn now, however, to Parfit’s main argument in favour of this conclusion.

**Fission**

In support of his claim that identity is not what matters in survival, Parfit asks his readers to consider a case he calls *My Division*:

My body is fatally injured, as are the brains of my two brothers. My brain is divided, and each half is successfully transplanted into the body of one of my

\(^7\) Ibid. p. 207.
brothers. Each of the resulting people believes that he is me, seems to remember living my life, has my character, and he is in every other way psychologically continuous with me.\(^76\)

Now we would have to ask ourselves, Parfit suggests, would he survive these events, and if so, which of his two brothers’ bodies would now be his?

Well, given that identity is a transitive relation it cannot be the case that Parfit survives as (is identical to) both future persons. If, as in the imaginary case above, Parfit (A) is psychologically continuous with both of his brothers (B and C), B and C have to be identical. This is clearly not the case – B and C are two numerically distinct persons, they can go off and lead separate lives and need not even know of each other’s existence. Nor is it easily claimed that Parfit survives in the body of one of his two brothers rather than the other as any such answer would seem completely arbitrary. Therefore, it seems, an adherent of the Psychological View would have to claim that neither of the two resulting persons is Parfit. He survives in neither of his brothers’ bodies.

This conclusion, however, is problematic. The problem is that if we agree that Parfit would survive in a case where his half-brain was successfully transplanted into just one of his brothers’ bodies, it seems that everything that would be needed for him to survive is also in place in the double transplant. In Parfit’s own words: “It cannot be the nature of my relation to each of the resulting people that, in this case, causes it to fail to be survival. Nothing is missing. What is wrong can only be the duplication.”\(^77\) If we believe, then, that everything that matters (in survival) is there between Parfit before the fission and both future persons it seems we ought to conclude – with Parfit – that identity is not what matters in survival.\(^78\)

Now, Parfit’s argument from fission is rather convincing: when considering the above kind of cases it does seem like fission (with branching sufficient psychological continuity) in all relevant aspects is more like survival than it is like death, and because identity cannot hold between the pre-fission person and the two successors (due to the transitivity of identity) it seems identity is not what matters in survival.

There are, however, alternative reactions to fission cases that maintain the importance of the identity relation. This is the view that persons are four-dimensional (perduing) objects rather than three-dimensional (enduring)

\(^{76}\) Parfit (1984), pp. 254-55.

\(^{77}\) Ibid, p. 261.

\(^{78}\) There seems to be an ambiguity here worth mentioning. It is not entirely clear whether Parfit intends his argument to show that identity is not what matters for survival, or that identity is not what matters in survival. I will, however, take arguments from fission to show that identity is not what matters in survival.
objects. On this view, objects have temporal parts, much as they have spatial parts: just as my right arm is a spatial part of me, so there is a temporal part of me that stretches between my fourth and my fifth birthdays. Three-dimensionalists (such as Parfit), on the other hand, deny that there are such temporal parts and claim that objects exist wholly at every moment of their existence.

On a four-dimensionalist view, then, we can describe the case of fission as one where there are two continuant persons right from the start – one of whom includes person time-slices A and B, and one of whom includes person time-slices A and C. These two persons, naturally, are numerically distinct, and so there is no problem with transitivity; they are like two distinct roads that coincide for a while before branching off in different directions.

This may seem like a possible solution to the problem of fission cases if one wants to maintain the importance of the identity relation, but it has often been pointed out that it seemingly brings some other problems in its wake. One such problem is that the four-dimensionalist view seems to imply that there are two persons located in the same body before the fission. One might respond to this objection, however, by pointing out that it is not entirely true to say that the four-dimensionalist view implies that there are two persons inhibiting the same body before the brain transplant in the case above. This is so because on this view there is really no one person located at any specific time, but only a person time-slice. What the four-dimensionalist is saying, then, is rather that one person time-slice can be part of two distinct (perduing) persons, and once you have accepted that persons are four-dimensional objects this really doesn’t seem to be such a strange claim.

Fortunately, perhaps, there is no need for our present purposes to side either with three-dimensionalism or with four-dimensionalism, because on both views psychological continuity matters in survival. Whereas three-dimensionalists like Parfit believe that it is psychological continuity and not identity that matters in survival, four-dimensionalists usually believe that psychological continuity matters because this is the relation that ties the different stages (time-slices) of the four-dimensional person together – without implying, of course, that one such time-slice is identical with another. That psychological continuity matters is also, one might add, the conclusion drawn by many of the animalist who believe that identity is not what matters. On all of these views, then, it would seem as though lack of sufficient psychological continuity would threaten what matters in survival.

---

79 This terminology is from Lewis (1983), one of the most famous proponents of a four-dimensionalist view.
80 Shoemaker (2012).
Fission and the advance directive

As we have seen, the proponent of the Objection from Personal Identity claims that the moral authority of advance directives is threatened by lack of sufficient psychological continuity, and, further, that it does so because lack of psychological continuity threatens the identity relation between the author of the advance directive and the later patient. The adherent of this objection thus seems to assume that identity is a necessary condition for the moral authority of advance directives.

As we have now seen, there might be good reason to question this assumption. If we build on Parfit’s My Division and add to the case the formulation by pre-fission Parfit of an advance directive, there seems to be two post-fission persons (each of whom believes that he is Parfit, seems to remember living his life, has his character, and is in every other way psychologically continuous with him) would not lead us to conclude that the advance directive would lose all moral authority. Rather, it seems we would believe that the advance directive is applicable to both of these persons (at least if we believed it to be authoritative if the psychological continuity relation was non-branching). 81

As we have also seen, however, a four-dimensionalist could agree with this normative conclusion while still being able to maintain that fission cases fail to show that identity is not a sine qua non of the moral authority of advance directives. This is so because according to the four-dimensionalist, the identity-relation holds between pre-fission Parfit (who formulated the advance directive) and both post-fission persons (to whom we believe the advance directive to apply). Thus, whether or not fission cases show that the proponent of the Objection from Personal Identity is mistaken in assuming that identity to be a sine qua non to the moral authority of advance directives, seems to depend on whether or not we ought to adopt a three-dimensionalist or a four-dimensionalist view of persons.

I suggested above, however, that there is really no reason – for our present purposes – to side with one of these views. And this is so because, on either view, lack of psychological continuity threatens the relation that matters (be it the identity relation or not). This, it seems, is what is important to the core of the Objection from Personal Identity. Even if three-dimensionalism is true and fission cases were to convince us that personal identity is not a necessary condition to the moral authority of advance directives, it seems these cases do nothing to show that psychological continuity is not a necessary precondition. In light of the objection that she mistakenly assumes identity to be a necessary condition for the moral authority of advance directives, the proponent of Objection from Personal Identity could thus either point to the fact that if we adopt a four-dimensionalist view of persons her assumption might

81 For a similar argument, see Buford (2008).
still be correct, or she could simply reformulate her argument slightly. She might say: "I may have been mistaken in my belief that identity is a sine qua non of the moral authority of advance directives. Nevertheless, there is nothing about fission cases to suggest that psychological continuity is not a necessary precondition, and therefore my claim that lack of psychological continuity threatens the moral authority of the advance directive seems to be correct, and this is what really matters."

Another suggested implication

As we have seen, Parfit believes the consideration of fission cases ought to convince us that identity is not what matters. What matters instead, according to Parfit, is Relation R, i.e. psychological continuity and/or connectedness with the right kind of cause. 82 Now, because identity is an all-or-nothing relation (either two person stages are identical or they are not), whereas psychological connectedness is a matter of degree, Parfit suggests that there is yet another implication that follows from the acceptance of his views. This implication is that we ought to treat moral and practical matters as matters of degree rather than as an all-or-nothing. When it comes to matters of desert, for example, we ought to accept that "when some convict is now less closely connected to himself at the time of his crime, he deserves less punishment [and] if the connections are very weak, he may deserve none."83 Or, when it comes to matters of commitments, that "the weakening of connections would reduce the strength of a commitment."84

In relation to the use of advance directives, then, it is quite possible that the conclusion that it is not identity, but rather psychological continuity that matters, could imply that the more psychological continuity there is between the author of the advance directive and the later patient, the greater the moral authority of the advance directive, and vice versa. This would be a change then, if we pre-theoretically believed that an advance directive would have full moral authority if there were sufficient psychological continuity for identity and no moral authority if there was an insufficient level.

Buchanan and Brock, however, refute Parfit’s suggestion and argue that, even if we accept that it is psychological continuity – and not (strict) identity – that matters to our moral and practical concerns, there could still be pragmatic reasons for employing a threshold level and not treating these matters as matters of degree. This argument will, however, be more closely examined in chapter five.

82 When it comes to his use of the term Relation R, Parfit is rather ambiguous. For a discussion of Parfit’s use, see Belzer (1996).
84 Ibid.
Conclusion

In this chapter we have seen that there are good reasons to question the assumption – seemingly made by the proponent of the Objection from Personal Identity – that identity is a sine qua non of the moral authority of advance directives. This is so, at least, if we are unwilling to side with four-dimensionalists, who believe that persons are four-dimensional (perduing) objects. We have also seen, however, that fission cases do not show that psychological continuity is not a necessary precondition of the moral authority of advance directives. And the claim that psychological continuity is a necessary condition, I have suggested, is the more important claim made by the proponent of the Objection from Personal Identity. A proponent of this objection could easily modify her argument and say that lack of psychological continuity still matters, and further, that it does so in the very same way that we believed loss of identity to matter.

There is, however, yet another possibility if one wants to argue that identity is not a necessary precondition of the moral authority of advance directives, namely that of appealing to normative theory. Because there are normative views that could support such moral authority even if we accept that the author of the directive and the later patient are numerically distinct beings, it could be argued that the proponent of the Objection from Personal Identity is mistaken in her assumption. As David Shoemaker points out, the intuition that a person’s advance directive ought to be granted moral authority in a case such as Margo’s, could be defended even if it would turn out that demented Margo is numerically distinct from pre-demented Margo, for example by appeals to “surviving interests” or “substituted judgments.”

Such attempts to defend the moral authority of advance directives will be considered in chapters five and six. In the next chapter, however, I will consider another possible implication of the non-identity claim, namely, that if the author of the advance directive and the later patient are numerically distinct the author must have gone out of existence, and if so nothing that we do can retroactively affect the value of this person’s already completed “life”.

---

85 Shoemaker (2010).
4. Can posthumous events harm us?

Introduction

In chapter two I argued in favour of the Psychological View, i.e. in favour of the claim that we (persons) persist by virtue of our psychology. If the Psychological View is correct the proponent of the Objection from Personal Identity could be justified in questioning the assumption that the author of the advance directive and the later patient are the same (numerically identical) individual. In chapter three I offered yet another defence of the Objection from Personal Identity by arguing that even if it is psychological continuity (and not identity) that matters in relation to our moral and practical concerns, this is not a very serious objection to this argument. In other words, even if it is not the case that pre-demented Margo and demented Margo are numerically distinct individuals, lack of sufficient psychological continuity between them matters in very much the same way that lack of identity supposedly did. Thus, we have to work out how to resolve what is in all morally relevant aspects just like an inter-personal conflict of interests, rather than a conflict between temporally distinct interests of the same person. So, how do we solve this conflict?

Now, the topic of the chapters to come will be arguments intended to support the belief that, in such a conflict, we ought to give priority to the former persons interests, i.e. grant the advance directive moral authority. In this chapter, however, I will consider an argument in favour of granting the demented patient’s interests priority in such a conflict – i.e. an argument that supposedly weakens the case for granting advance directives moral authority. This argument takes as its point of departure the fact that the non-identity claim seemingly implies that the author of the advance directive has ceased to exist (or at least because it is just as if she had gone out of existence). And if this is the case, the argument continues, nothing that we do will retroactively affect the value of this person’s already completed “life”.

Although this argument admittedly has initial appeal, it hinges on the questionable assumption that death leaves its victim immune from any fur-

---

ther harm (or good) – a view that we may refer to as the immunity thesis. In this chapter I will suggest that we ought to reject this assumption. I will thus argue in favour of the view that the value of a person’s life can be affected retroactively and, further, that this (sometimes) gives us at least a pro tanto reason to abide by a person’s advance directive, even if the non-identity claim implies that the person in question has ceased to exist.

Making a case for posthumous harm

The immunity thesis is often associated with a well-known argument stated by Epicurus. According to this argument, neither death, nor anything that occurs after death can harm those who die – and this is so simply because those who die are not made to suffer as a result of either. Here are the words of Epicurus:

>Become accustomed to the belief that death is nothing to us. For all good and evil consists in sensation, but death is deprivation of sensation. So death, the most terrifying of ills, is nothing to us, since as long as we exist death is not with us; but when death comes, then we do not exist.

Now, although I will later argue that posthumous events may harm us, I believe a denial of this possibility may sound initially plausible. The claim that death cannot harm us, however, seems rather absurd. According to most people, death is bad for the person who dies, and therefore it seems any account of harm that suggests it cannot be, must somehow be mistaken.

Because most of us have such strong intuitions when it comes to the badness of death, several philosophers have wanted to reject Epicurus’ claim. One way of doing so is to adopt the view that an event harms us if it prevents us from having something that is good for us. Let us call this a deprivation view of harm. According to this account we could easily explain why we often believe death to be bad for the person who dies by reference to the fact that it (often) deprives its victim of a valuable future. Further, it can quite plausibly explain why we believe death not to be a misfortune in cases where it is obvious that the life lost would have been one filled with nothing but suffering and/or misery. In such cases death does not seem to deprive the person in question of anything good, but rather seems to relieve him of something that would have made him worse off.

However, a preclusion view of harm also underlies many of the arguments made by philosophers who want to defend the further belief that post-

---

87 In choosing this label I follow Luper (2004).
88 See Epicurus Letter to Menoeceus.
89 See Luper (2004).
humous events can harm us. Just like death, these philosophers argue, posthumous events seem capable of preventing us from attaining things that would be good for us. These goods are then usually spelled out in terms of satisfaction of our “surviving interest”.

Many of our interests, it seems, are thwarted by death itself. This is, for example, the case with interests we might have that we do something. If I have an interest, say, in writing a great novel or being the first Swede on the moon, this interest seems to be thwarted by my death, i.e. my death itself precludes the possibility of me writing a great novel or travelling to the moon. And if death itself thwarts my interests, nothing that happens after my death can compound the damage. These kinds of interests are not, then, interests that can “survive” me.

Not all our interests, however, are thwarted by our death. For example, let us say that I have invested much of my adult life in making sure that my children will lead happy lives (say by trying to instil certain values in them etc.). Now, this interest of mine surely wouldn’t be thwarted by my death – it would be thwarted only by the event that my children’s lives are unhappy. Of course, some event that would render my children unhappy could happen within my lifetime, but it could also happen after my death and either way, the unhappiness of my children would harm me or make me worse off. Or at least so a friend of the possibility of posthumous harm would claim.

Now, some might intuitively find the latter part of this argument to be very mysterious. Sure, one might say, not all of our interests are such that they are thwarted by our death, yet there must be some mistake here: assuming that death is the total and final end of a person, how can a person possibly be made worse off (or indeed be harmed) posthumously?

Joel Feinberg, a well-known proponent of the belief that posthumous events could harm the dead, offers the following argument to convince the sceptic that a person can be harmed even by posthumous events:

If someone spreads a libellous description of me among a group whose good opinion I covet and cherish, altogether without my knowledge, I have been injured in virtue of the harm done my interest in a good reputation, even though I never learn what has happened. That is because I have an interest, so I believe, in having a good reputation as such, in addition to my interest in avoiding hurt feelings, embarrassment, and economic injury. And that interest can be seriously harmed without my ever learning of it.

Feinberg continues his line of argument as follows:

Suppose [now] that after my death, an enemy cleverly forges documents to “prove” very convincingly that I was a philanderer, an adulterer, and a plagia-

---

rist, and communicates this “information” to the general public that includes my widow, children and former colleagues and friends. Can there be any doubt that I have been harmed by such libels? The “self-centred” interest I had at my death in the continued high regard of my fellows, in this example, was not thwarted by my death itself, but by events that occurred afterward.\textsuperscript{92}

Feinberg’s argument is convincing: some things do seem to harm us irrespective of whether or not we ever come to know of them (or in some other ways experience the harm – for example by missing out on good things we otherwise would have had), and if this is the case it seems fairly natural to assume that such harms can befall us even posthumously. Assuming that I never know about the slander, it doesn’t seem to matter whether I will die in five minutes or died five minutes after the slander took place to the question of whether or not I was harmed by it.

\section*{Debunking the intuition}

Some philosophers have argued, however, that Feinberg’s argument fails because, on closer inspection, we ought to abandon the intuition we have with respect to the first scenario where a person is supposedly, and unknowingly, harmed by the slander conveyed to his colleagues and kin. This is, for example, the argumentative strategy employed by Ernest Partridge.

According to Partridge, most of us would make the judgment that having his reputation spoiled even if he never becomes aware of the slander would indeed harm the person in the first scenario. In fact, it seems we would make the same judgment even if we were to add the assumption (probably also made, implicitly, by Feinberg) that the slander doesn’t cause our supposed victim to lose out on any other goods he otherwise would have had, such as job opportunities and or rewarding friendships. Partridge argues, however, that this judgment is mistaken and that the mistake can be explained by reference to the fact that we cannot be \textit{absolutely certain} that these rumours have no effect on the victim of our story. In Partridge’s own words, there seems to be some “illegitimate subjective leakage past our conditions”: we cannot really imagine a case where we can be absolutely certain that the rumours have no affect on the (living) victim of our story, and the mere possibility of him being affected is what makes us reluctant to drop the intuition that he is harmed. In the case where the victim is dead when the slander takes place, however, there is no risk of “feedback” and therefore we are more prone to say (and rightly so) that no harm is done to him.

\textsuperscript{92} Ibid.
In support of his argument Partridge points to the fact that our “harm” intuition seems to fade with the likelihood that the rumours would have an effect on the victim’s life.

[To whom are these slanders told? To professional colleagues in the next state? To some hill folk in Appalachia? To some unlettered natives of inland Borneo? The alleged harm would seem to fade with the likelihood of feedback and, significantly, to the degree that the “victim” is an identifiable person to those who deride him. In short, I just do not believe that we are subjectively capable of accepting the “unaffected-harm paradigm” totally and without qualifications as it applies to the living. We cannot, I suggest, because our intuitions suggest that the victim cannot be isolated from harm leak back in.93]

Now, Partridge seems to be correct in his observation that we would be less prone to claim that the person is harmed by the malicious rumours the longer the distance is between the person and the group of people they are told to. I will suggest, however, that this observation does very little to support Partridge’s view over Feinberg’s.

It seems to me that one could easily rebut Partridge’s claim by saying that it is true that our response to the different cases changes with the distance, but this is not because we are incapable of accepting that the person is in no way affected by the rumours in the first imagined scenario, but rather because it is very likely that the interest the person has in not being slandered among his kin is so much greater than his interest in not being slandered among, say, a group of natives of inland Borneo. We simply make the (often times correct) assumption that the longer the distance the less strong the interest and therefore the lesser the harm. If it was the case that the victim of our story – for one reason or another – cared much more strongly about what the natives of inland Borneo thought of him than he cared about the thoughts of the people closer to him (say his colleagues), perhaps simply because the former were dearer to his heart, then I believe we would be very likely to think that slander told to those natives would be a worse harm – and this regardless of the (perhaps minimal) possibility of it having an effect on his well-being.

Partridge argument notwithstanding it is thus far from obvious that we are mistaken when we judge that harm is done to the victim of our story with regards to the first scenario. Partridge’s argument, however, is admittedly one that he makes “reluctantly” because he believes that “Feinberg has succeeded so well in binding the case of the unaffected living and the dead that no logical wedge can be inserted such that the former [person] can be said to be “harmed” and the latter cannot.”94 This however, seems to be mistaken. There are, it seems to me, possible – and at least initially plausible – ways to

94 Ibid.
argue that the person in the first scenario is unknowingly harmed, yet deny that there is any harm done in the second.

One possible way to argue that there is harm done in the first scenario while still holding onto the belief that there can be no such harm in the second, is to suggest that the thwarting of a person’s interest only harms him or her if it occurs at a time when the person in question (still) has that particular interest. Whereas it is quite clear that the person depicted in the first of Feinberg’s scenarios still has the interest at the time of its being thwarted (although it is thwarted unbeknown to him) this is not the case in the second scenario and therefore it makes no sense to speak of any posthumous harm being done.

Now, it does seem initially plausible to argue that a person can only be harmed by having his interests thwarted at time 1 if the relevant interest is among his current interests at time 1. The initial plausibility of this claim can probably partly be explained by the fact that we rarely believe that we have reason to act on our past interests. Take, for example, Parfit’s famous argument that his childhood ambition to become a poet doesn’t seem to give him any reasons, now that he has abandoned it, for writing poetry.95 This is probably true – once we have voluntarily given up on an interest and/or desire there is no reason to try to satisfy it and the thwarting of an interest that we have abandoned cannot reasonably be claimed to make us worse off. It is less clear, however, that the same thing can be said of interests that are not given up but that we may still lose before they are either thwarted or satisfied.

Another possible reason why it seems initially plausible to argue that a person can be harmed by having his interests thwarted at time 1 only if the relevant interest is among his current interests at time 1, could perhaps be the fact that some of our interests are “conditional on their own persistence”.96 This is the case, for example, with interests we have because we think we would enjoy the satisfaction of having them gratified. For instance, if I have an interest (or desire) to have a glass of wine on Friday night, it is probably conditional on me still being up for a glass of wine on Friday night. Thus, if I no longer want a glass of wine on Friday night I am not likely to enjoy it and then not having one does not make me worse off. But again, not all of our interests are like that. I may have an interest in some future event regardless of whether or not I will be there to enjoy it – such as an interest in the well-being or happiness of my children, or in the future flourishing of some project. It just doesn’t seem true to say of all of our interests that their thwarting cannot harm us unless they are thwarted at a time when they are among our current interests.

95 Ibid. p. 157.
There is, however, yet another challenge to the view that the thwarting of the above kinds of interests amounts to genuine cases of posthumous harm. This challenge is known as the timing problem and concerns the dating of the bad (or good) things for which posthumous events are responsible.

The timing problem

The timing problem stems from the fact that seemingly, things that harm us harm us at a point in time. If I miss the train to the airport, for example, this event is bad for me at those times when I would have enjoyed myself abroad had it not been for me missing the train. But when, it might reasonably be asked, could an event occurring after my death possibly harm me? If no reasonable answer can be given to this question, it is often assumed, there can be no such thing as posthumous harm and the immunity thesis ought to be accepted.

There are, however, several possible responses to the timing problem, some of which are more plausible than others. One of the more persuasive of them is one formulated by George Pitcher – yet another well-known proponent of the view that posthumous events can harm us. Pitcher’s suggestion is that posthumous events harm us before we die and this is so because even when the person was alive it was either true or false of him that his interests were thwarted.

The sense in which an ante-mortem person is harmed by an unfortunate event after his death is this: the occurrence of the event makes it true that during the time before the person’s death, he was harmed – harmed in that the unfortunate event was going to happen.

So, when a post-mortem occurrence makes it true of someone that his interests were going to be thwarted, it makes the same thing true of the ante-mortem person, and therein lies the harm. For example: if my last wishes will not be respected after I’m dead, this is also true of me when I have these wishes, and this fact is what is harmful to me (then) although what makes it true will occur somewhere in the future.

According to Pitcher himself, the main obstacle to accepting the thesis that an ante-mortem person can be harmed after death is the disturbing suspicion that it would imply backward causation, i.e. the idea that it is possible for an effect to temporally precede its cause. If this were the case, the notion of posthumous harm would indeed be rather mysterious, as it would suggest

---

97 For a thorough discussion of these different responses, see Luper (2009).
quite the reverse temporal order of causes and effects than the one we are familiar with.\textsuperscript{99}

Pitcher continues, however, by strictly rejecting that his view has this implication. This objection, he argues, rests on the assumption that being harmed has to be viewed as a kind of alteration in one’s metaphysical state. To illustrate why we ought not view harm to be such an alteration Pitcher tells the story of Mr Black and his son Jack.

Suppose that Mr Black’s son Jack is killed in an airplane crash many miles away. Given that his son’s welfare is one of Black’s strongest interests, the son’s death harms Black (is a great misfortune for him). There should be no temptation to think that this harming of Black requires instantaneous causation at a distance - the plane crash sending out infinitely rapid waves of horror, as it were, diminishing Black’s metaphysical condition. If that idea is absurd, so is the idea that if the son’s death should occur after Mr. Black’s, and should thus harm the ante-mortem Black after his death, it must do so by a process of backward causation.\textsuperscript{100}

This response to the objection is seemingly a result of the fact that Pitcher (like Feinberg) understands harm in terms of “invaded interests”, and naturally, one’s interests can be blocked or thwarted without this having any affect whatsoever on one’s metaphysical state.

Some philosophers, however, have found other reasons for being reluctant to accept Pitcher’s response to the timing problem. Joan C. Callahan, for example, argues that Pitcher’s response fails because it seems to prove too much:

For what now shall we say of a person who will later perform an action productive of posthumous harm? It seems that we must say that he is, long before doing something, already responsible for placing the [ante-mortem person] in a harmed state. Even worse, it can surely turn out on this view that our would-be agent is responsible for harming another even before he is born. ... For those of us who do not ascribe to some doctrine of original sin, this is a difficult implication to swallow.\textsuperscript{101}

This objection, it seems to me, is problematic because it confuses the claim that if it is true that something will happen it is true even now, with the claim that something is necessarily true.

Usually we take someone to be (morally) responsible for some act only after he/she has performed it. This is so because up to the point when the act is performed we like to believe that the person has some choice about whether or not to perform the act in question. The problem with Callahan’s

\textsuperscript{99} See e.g. Anthony Flew (1954, 1956, 1956-7) who has argued that backward causation involves contradiction in terms, and Max Black’s (1956) bilking argument.

\textsuperscript{100} Pitcher (1984), p. 186.

\textsuperscript{101} Callahan (1987), p. 345.
argument above is that it seems to rely on the false belief that Pitcher’s suggestion precludes the possibility of such a “free” choice. To say that *if it is true of some person X that he harms Y then it is true of X that he harms Y even before the event occurs,* is not to say that X doesn’t have a choice in the matter or that he is guilty of performing an act before he performs it. That Y is harmed by X’s actions is true *only if* X actually harms Y, but if he does, it is true of Y that he is harmed by X also before this happens. And the same goes, of course, for the event where Y is not so harmed, and which way it goes is still (we can assume) up to X right until the point where he chooses to act, or not to act, so as to harm Y. There is thus no reason to believe that Pitcher’s account implies that someone is guilty of harming some other person before he chooses to act so as to harm him.

As we have already seen, the timing problem stems from the fact that most things that harm us harm us at a point in time. However, the argument that there could be no such thing as posthumous harm (because there is no plausible account of when such harm would befall its victim) rests on the far more controversial assumption that *anything* that is bad for a person is bad for her at a time. This assumption, naturally, could be questioned. Perhaps there are harms that are such that they do not harm us at a point in time, but that are *atemporally bad for their victim,* and perhaps posthumous events can cause this kind of harm.

Naturally, it might be objected to this view that the suggestion that posthumous harms are atemporally bad for their victim while most other harms are bad at a point in time, makes them so different from other kinds of harms that we ought to question that they really are harms, but I see no reason why we ought to agree with this claim. There are, it seems, many different ways in which a person could be made better or worse off (surely there are clear differences between, for example, stubbing one’s toe and – perhaps even unknowingly – missing out on a job opportunity) but the fact that there are great differences need not in any way lead us to conclude that one of the two cannot really be a harm. Why should it? We could still say of both stubbing one’s toe and missing out on a job opportunity that a person would have been better off had either one of these events never happened – and therefore both kinds of events constitute harms. So, unless we agree that all things that harm us must do so at a point in time, the timing problem doesn’t seem to be a problem to the adherent of the belief that posthumous events can harm us.

So far we can conclude, then, that at least some of the supposed mysteries with Pitcher’s response to the timing problem are not very mysterious after all, and that even if we accept the assumption that all things that harm us harm us at a point in time (an assumption I believe we have good reason to

---

102 For a defence of the view that death is such a harm, see Johansson (forthcoming). See also Luper (2009).
question) there is a case to be made for the belief that posthumous events can harm us and that we could reasonable reject the immunity thesis.

The possibility of posthumous harm and the advance directive

If we reject the immunity thesis, as I have suggested that we should do, it seems the non-identity claim does not imply that we give priority to the demented patient’s interest – at least not on the basis that nothing that we do can retroactively affect the value of the author’s already completed life. However, as we have also seen in this chapter, it is far from true that all of our interests “survive” us. This is for example the case with interests that stem from what one might call “personally defined projects” (me writing a great novel or being the first Swede on the moon), it is also the case with those of our interests that are, in Parfit’s words, “contingent on their own persistence”, i.e. interests that we wouldn’t have if we didn’t think that we would enjoy the satisfaction of them.

Perhaps, then, we could suggest a revised argument that doesn’t rely on the immunity thesis but suggests rather that many of the kinds of interest that are actually expressed in an advance directive are the kind of interests that cannot survive us, and therefore these interests ought to be discounted when they are in conflict with the patient’s current interests.

There are perhaps good grounds for claiming this. For example, my interest in not becoming so severely demented that I no longer recognise my family members, or cannot take care of myself without assistance might very well be interests that are contingent on me still caring about remembering my family members or having an aversion against having to be taken care of, and if this is the case, a former persons life wouldn’t actually be made worse off by having those interests thwarted after he has lost those interests (either by death, loss of personal identity or simply through advancing dementia).

But again, there are other kinds of interests as well that might play a role in determining how one formulates an advance directive. Let us say, for example, that a person formulates an advance directive suggesting that we let his “successor” die under certain circumstances if he contracts pneumonia. The main reason for this is that his children be spared the financial burden that keeping his “successor” alive in a hospital would entail. The thwarting of this kind of interest (the well-being of one’s children) is one, I assume, that could be harmfully thwarted after one’s death, and perhaps this kind of advance directive therefore would provide us with at least a pro-tanto reason for abiding by it.
Both the question of what kind of interests can “survive” us, as well as what kinds of interests really underlie an advance directive, are complicated issues. But perhaps it suffices to say that some such interest would not survive a person’s demise but that there are others that could. It should also be noted, however, that the fact that some such interests survive says nothing about the normative significance they have or what weight they ought to be assigned in a conflict with the interests of the current patient.

Conclusion

In this chapter I have considered an argument supposed to support the belief that in a conflict between the interests of (numerically distinct) pre-demented Margo and demented Margo, we ought to give priority to the interests of the demented patient because pre-demented Margo has ceased to exist and therefore nothing that we do can retroactively affect the value of this person’s already completed “life.”

This argument, I suggested, hinges on the immunity thesis, i.e. the assumption that persons cannot be harmed by posthumous events. My main objective in this chapter, then, was to show that the immunity thesis can reasonably be rejected. Some of our interests, I have suggested, are such that their thwarting can harm us or make us worse off even if they are thwarted after we have ceased to exist. Therefore, the non-identity claim (if true) does not give us reason to simply disregard the interests of the former person. Nevertheless, my arguments to the effect that we could have a pro-tanto reason to abide by an advance directive because of the harm we might otherwise do to the former person, does not in itself solve the suggested conflict between the former person and the current patient. The question now is rather whose interests we ought to give priority: the former interests of pre-demented Margo or the interests of demented Margo? In the next chapter I intend to consider some suggestions that have been made towards the conclusion that, in view of the fact that the demented patient is not likely to be a person, we ought to give priority to the past person’s interest – an argument that I will regard with some scepticism.
5. The normative relevance of personhood

Introduction

In the last chapter I argued that we ought to reject the immunity thesis, i.e. the belief that death leaves its victim immune from any further harm (or good). I concluded, however, by pointing out that even if the immunity thesis fails, this doesn’t make so much of a case for the moral authority of advance directives as one might perhaps be led to think. This is so because even if events that occur after we have ceased to exist can make us worse off or indeed harm us, it is far from obvious that a person’s surviving interests ought to be granted authority in relation to the current interests of a demented patient. If we want to arrive at this conclusion, it seems, we need further arguments.

One attempt to provide such an argument, i.e. an argument in favour of granting advance directives moral authority even in cases where there is an inter-personal conflict between the interests of the former person and the later patient, has been put forward by Allen Buchanan and Dan Brock in their influential book *Deciding for Others. The Ethics of Surrogate Decision Making*. Their argument – which makes use of the supposed fact that the demented patient is a non-person – will be our main interest in this chapter. As it turns out, it is very unclear whether or not their normative framework really provides the solid defence of the moral authority of advance directives (in the cases under consideration) that they seem to argue.

Yet another argument which makes use of the supposed fact that the demented patient is a non-person is one suggested by Helga Kuhse. Kuhse suggests that a demented patient like Margo is not harmed by a non-treatment decision. This argument, I will suggest, fails because it rests on too narrow an account of harm.

Buchanan and Brock’s argument

Before moving on to Buchanan and Brock’s main argument, it is perhaps important to say something about where they stand in relation to the issues I have so far discussed in this book. First, Buchanan and Brock agree with the
assumption that “whatever the correct theory of personal identity turns out to
be, it will include the claim that psychological continuity is (at least) a neces-
sary condition for personal identity.” Further, they agree that we ought
to accept that some of our interests are such that they survive us and that
therefore “we must refrain from passing final judgment upon how good or
successful the person’s life was until the relevant interests are thwarted or satisfied”. They also agree with the suggestion, made earlier in this book,
that in a case in which brain damage is so severe that personal identity is
lost, there are now two distinct sets of interests: the experiential interests of
the current patient and the surviving interests of the person who has ceased
to be. Accordingly, if it is true that pre-demented Margo and demented
Margo are numerically distinct, there at least could be a genuine inter-
personal conflict between the interests of pre-demented and demented Mar-
go. So, then, how do Buchanan and Brock suggest that we resolve this con-

Well, according to Buchanan and Brock we ought to distinguish, first,
between two kinds of cases: the first kind of case is one where the author of
the advance directive and the later patient are distinct and where the later
patient is a person; the second kind of case is one where, again, the two are
numerically distinct but where the later patient is a non-person. According to
Buchanan and Brock, the first kind of case would indeed threaten the moral
authority of the advance directive. Say Buchanan and Brock:

if faced with a choice between following one person’s wishes in an advance
directive and saving another person’s life we would have to disregard the ad-
vance directive. To do otherwise would be to give one person a wholly illegiti-
mate, nonconsensual power of life and death over another person”.

The second kind of case, however, i.e. where the later patient is a non-

person, is different. When the demented patient is a non-person, they sug-
gest, we can provide a justification for granting the advance directive moral
authority by appeals to “a right of disposal” or a “quasi-property right”. Later
on in this chapter, I will point to some problems with this suggestion. For
now, however, I will leave the second kind of scenario to one side and focus
on Buchanan and Brock’s argument with regards to the first kind of case, i.e.
where the author of the advance directive and the later patient are distinct
and where the later patient is a person.

103 Buchanan and Brock (1990), p. 155.
104 Ibid. p. 163.
105 Ibid. p. 164.
106 Ibid. p. 168.
A low (conservative) threshold

According to Buchanan and Brock the problem with the Objection from Personal Identity is that it presupposes the first kind of case: it presupposes that “the individual who remains after neurological damage has destroyed the person who issues the advance directive is a (different) person”.\(^{107}\) This presumption, Buchanan and Brock continue, is problematic because these kinds of cases are very unlikely to occur. This is so, at least, if we adopt a low (or conservative) threshold level of psychological continuity to count as sufficient for preserving personal identity – something Buchanan and Brock believe we have good reason to do.

If the degree of psychological continuity necessary for preservation of personal identity is set rather low or, conversely, if the degree of psychological continuity compatible with the preservation of personal identity is set rather high, then there will be very few if any real world cases in which we would be justified in concluding that neurological damage has destroyed one person but left a living, different person.\(^{108}\)

Why then, according to Buchanan and Brock, ought we assume a low (or conservative) threshold-level?

Well, as we have already seen, Buchanan and Brock agree with the proponent of the Objection from Personal Identity that whatever the correct theory of personal identity turns out to be it will include psychological continuity as a necessary condition for personal identity. But this view does not, they continue, in itself answer the question of how much psychological continuity that is necessary for the preservation of personal identity. How much that is sufficient, they suggest, can be decided much in the same way that we decide – to use their own example – what amount of decision-making capacity constitutes competence. The similarities are here being described as being “a matter of choice, not a decision uniquely determined by the facts of the case.”\(^{109}\)

Buchanan and Brock then suggest that “a healthy and reasonable conservatism weighs heavily in favour of setting the threshold rather low”\(^{110}\) and that any attempts to raise the threshold (from the low level we are currently employing) would be very costly both morally and financially. Such a shift, they claim, would:

result in the “births” of large numbers of “new persons” who would, as it were, spring full-blown into the world and who would not, strictly speaking, be the sons, daughters, husbands, wives, or friends of anyone. Such “new per-

\(^{107}\) Ibid. p. 158.
\(^{108}\) Ibid. p. 159.
\(^{109}\) Ibid. p. 158.
\(^{110}\) Ibid. p. 174.
sons’ would have no financial assets (nor debts), nor would any individual or family be legally responsible for them. ... The price of setting the threshold for psychological continuity high is that doing so enormously complicates and magnifies the problem of intergenerational justice.\textsuperscript{111}

Now, there are some problems with this part of Buchanan and Brock’s argument. Let us first consider the claim that the level of psychological continuity necessary for the preservation of personal identity is one of choice, rather than one uniquely determined by “the facts of the case”. This claim, it seems to me, is not entirely true. While it is true that the Psychological View, i.e. the view that we persist by virtue of our psychology, doesn’t in itself answer the question of how much psychological continuity that is necessary for the preservation of personal identity, it seems very unlikely that an adherent of this view would agree with Buchanan and Brock’s further claim that this therefore is a matter of choice. Rather he or she is likely to maintain that there is a matter of fact here.\textsuperscript{112} Naturally, the answer to the question of how much psychological continuity is sufficient for personal identity might be very difficult to find, but this doesn’t mean that there isn’t one.

As we saw in chapter three, however, it is true that some adherents of the Psychological View (as well as some adherents of the biological approach) have come to the conclusion that what matters (in survival) is not personal identity but rather psychological continuity in itself. In the light of this view Buchanan and Brock’s suggestion seems more reasonable. If it is not identity that matters, but rather the psychological continuity relation in itself, the relation that matters is one that comes in degrees. If this is so, we could either take the route suggested by Parfit, i.e. acknowledge this and treat moral and practical matters as matters of degree rather than as an all-or-nothing issue, or we could do what Buchanan and Brock suggest and choose to employ a threshold level that seems to fit our already established social practice and institutions. Either way, it seems to me, it is only if we have given up on the view that it is identity that matters, that it is reasonable to claim that what threshold-level we ought to employ is one of choice.

However, it does seem very reasonable to accept the view that it is not identity, but psychological continuity in itself, that is germane to our moral and/or practical concerns (especially if one wants to hold on to the view that persons are three-dimensional (enduring) objects), and if this is what we believe there is also some truth in Buchanan and Brock’s claim that we can still (Parfit’s suggestion notwithstanding) choose to employ one threshold-level in preference to another for pragmatic reasons. Having accepted Buchanan and Brock’s argument thus far, however, it seems one can still ques-

\textsuperscript{111} Ibid. p. 177.

\textsuperscript{112} This “matter of fact”, I take it, could be either metaphysical or normative, but either way a proponent of the Psychological View would be likely to say that there is one.
tion their argument in favour of a low (conservative) threshold level over a higher one. Let us next consider those arguments.

As we have seen, Buchanan and Brock’s main argument in favour of a low threshold is that raising the threshold would be morally and financially costly. The problem with Buchanan and Brock’s argument is that their examples of social practices where attempts to raise the threshold would be costly (when it comes to being the son or daughter of someone, as well as legal aspects concerning financial assets and liabilities) are all examples where any suggested level of psychological continuity seems to run into some problems – at least if we want to stay in line with what our social practices and institutions. According to our social practices, it seems, the cases that Buchanan and Brock mention tend more to be cases where we don’t believe psychological continuity to matter at all. Here, it seems, it is truer to say that we are making judgment based solely on something like biological continuity.

This becomes even clearer, it seems, if we employ the threshold Buchanan and Brock favour in relation to these kinds of matters. For example, let us consider a clear-cut case where there is no psychological continuity between a person and his/her successor – like that of a patient who has slipped into a PVS. According to Buchanan and Brock’s own suggestions no psychological continuity would surely be too little to preserve identity. In the above mentioned case, then, we would be forced to say that the PVS patient is no longer the son or daughter of anyone and we wouldn’t be entitled, for example, to use the previous person’s financial assets to pay for the current patient’s care etc. This doesn’t seem to be in accord with our social practices and institutions either, but the reason for this is probably not that there is something wrong with the suggested threshold level of psychological continuity, but rather that when it comes to these kinds of practical matters we tend not to care about psychological continuity at all. When it comes to assets and debts as well as “formal” relationships such as being the son or daughter of someone, we tend to think that psychological continuity matters very little. The examples cited by Buchanan and Brock are therefore not suitable ones, it seems, of why we ought to employ a conservative level of psychological continuity.

One further possibility is that, although we agree with Buchanan and Brock that we need not treat all moral and practical concerns as matters of degree, we reject their assumption that the same level of psychological continuity is employed in relation to all possible different kinds of cases. Why, for example, would the same threshold-level be relevant for having formulated an authoritative advance directive, or being the same “person” as the one who committed a crime years ago, as that we would have for being entitled to that former person’s financial assets, and being called by the same name etc.? In fact, Buchanan and Brock’s own analogy with decision-making capacities and competence seems to open up the possibility of reject-
ing this assumption. Competence, it seems, is not usually considered to be an all-or-nothing question in the sense that either you are competent to make all kinds of decisions or you are not competent to make any. Quite the contrary – the level of decision-making capacity constituting competence is often believed to be context specific. To take an example that seems close at hand: a demented patient might very well be incompetent with regard to complex medical decisions, yet be fully competent to decide what he/she wants for dinner. By way of analogy then, perhaps we ought to set the identity-preserving threshold very low (or even say that psychological continuity is redundant) when it comes to matters like assets, debts and the like, while raising the threshold – or, quite possibly, holding onto the threshold we are in fact employing – when it comes to questions such as deserts, moral authority and compensation.

A quasi-property right

As we have now seen, Buchanan and Brock believe that there is good reason to question the assumption that a distinct successor would be a person, and that this, in turn, seriously undermines the Objection from Personal Identity. I have suggested above, however, that their main argument in favour of the claim that a successor is very unlikely to be a person has some problems and that this is so because it fails to convince us that we ought to adopt a low threshold level of psychological continuity as sufficient for personal identity (or indeed as sufficient for what matters).

This, naturally, is not to say that we never would be justified in making the judgment that the successor and current patient is a non-person. Therefore, let us next turn to Buchanan and Brock’s view that this matters morally. Why is it, according to Buchanan and Brock, that we need not be overly worried about lack of personal identity in cases where the author’s successor isn’t a person?

Well, according to Buchanan and Brock the fact that the demented patient is very unlikely to be a person matters because as long as she isn’t, we could support the moral authority of the advance directive by reference to a “right of disposal” or a “quasi-property right”. Thus, the reasons we have for respecting a person’s advance directive (at least when it comes to the cases under consideration here, where we assume that the author of it and the later patient are distinct beings) are very similar to those we have for respecting a person’s regular will: we should do this simply because a person’s material possessions as well as her body are hers, and this grants her the authority to decide (albeit within certain limitations) over what should be done with them.

Although a quasi property right, according to Buchanan and Brock, provides the advance directive with less authority than a right to self-
determination does, they argue that this right suffices to justify the moral authority of advance directives. Say Buchanan and Brock:

If the right to determine what happens to one’s nonperson successor is viewed as a quasi-property right it would, like one’s ordinary property rights, be dispositive in ordinary circumstances. It could only justifiably be overridden to secure some very important good or to avoid some very important harm.\textsuperscript{113}

Sometimes, it has been argued that an “ownership model” such as that suggested by Buchanan and Brock is simply too controversial to be accepted, and that this is so because most people would explicitly reject the moral legitimacy of one person owning another human being – person or not.\textsuperscript{114} This kind of critique would surely be justified if Buchanan and Brock’s suggestion entitled the “owner” of demented Margo to the same kind of freedoms that a person is usually taken to have over her material property. That is, if we took their argument to be that the supposed fact of demented Margo being a non-person would imply that we do not owe her any moral consideration at all.

On closer inspection, however, this doesn’t seem to be Buchanan and Brock’s view. Quite to the contrary, they emphasise that “lack of personhood does not imply lack of moral status altogether. The fact that a being can experience pleasure and pain may itself impose significant limitations on how we may act towards it.”\textsuperscript{115} But if we can owe demented Margo moral consideration even if she is a non-person, one might wonder how far pre-demented Margo’s supposed “quasi-property right” really stretches? If it is limited by every appeal one might have when it comes to demented Margo’s well-being or “best interest,” Buchanan and Brock’s defence of the advance directive doesn’t seem to be as strong as they make it out to be – suggesting only that we ought to abide by the advance directive when there is no conflict between doing what is in demented Margo’s best interest and doing what the advance directive prescribes.

There is, however, a further component in Buchanan and Brock’s argument that suggests yet another implication of the supposed fact that the demented patient is a non-person. This is the suggestion that non-persons have radically truncated interests, and as such:

our obligations toward that being are quite limited. Following an advance directive to achieve a painless termination of life-support for a being with such radically truncated interests need not involve the violation of any obligations toward that being.\textsuperscript{116}

\textsuperscript{113} Ibid. p. 166.
\textsuperscript{114} Buford (2008), p. 429.
\textsuperscript{115} Buchanan and Brock (1990), p. 160.
\textsuperscript{116} Ibid. p. 168.
The argument here seems roughly to be that if a being is a non-person, he/she has no interest in being kept alive. If this were true, it would at least suggest that we ought to abide by an advance directive, counselling us to let the demented patient die without compromising the best interest of the demented patient, because there just seems to be no conflict: we can do what we owe to the former person without violating any of the moral obligations we have towards the demented patient. The problem with this argument is, however, that one needs further arguments as to why we ought to assume that non-persons have no interest in continued life. To my knowledge, Buchanan and Brock provide no such argument but in the next part of this chapter I will consider – and reject – such an argument propounded by Helga Kuhse.117

Kuhse’s argument

Like Buchanan and Brock, Kuhse starts her line of argument by accepting – at least for the sake of argument – that psychological continuity is a necessary condition for personal identity, and thus that the author of the advance directive and the later patient (due to lacking sufficient psychological continuity) could be two distinct individuals. Kuhse does not, however, agree with the view that some of the author’s interests – such as an interest in the financial or emotional well-being of her loved ones – could “survive” the former person. Rather, she argues, “the value of advance directive lies ... in the comfort and assurance those of us who are persons derive from the knowledge that our wishes will be honoured after we have ceased to exist.”118 However, Kuhse continues, not even saying that advance directives serve the interests of (existing) persons is sufficient to justify their use. In order for the moral authority of advance directives to be upheld, she suggests, it is necessary also to question the assumption that a patient such as Margo will be harmed by a non-treatment decision. This, then, is the task she sets herself.

According to Kuhse, it is instructive to compare a severely demented patient like Margo with a non-human animal. Neither a non-human animal nor demented Margo, Kuhse argues, has the capacities we usually associate with personhood, e.g. self-consciousness, rationality and purposive agency. More importantly, neither the animal nor demented Margo seems to have a conception of themself as existing over time, and this specific ability, Kuhse then suggests, is a necessary condition for being harmed by death. This is so, according to Kuhse, because “only a being capable of understanding that it

has a prospect of future existence can have a desire to go on living, and only
a continuing self – or “person” – can have an interest in continued life.”
Although I agree that the comparison between a severely demented patient
and an animal is in some ways instructive, I will later challenge the suggestion that neither is harmed by death, but first, let us more closely consider what Kuhse believes to follow from this argument with regard to the moral authority of advance directives.

Kuhse emphasises that to say that an individual lacks an interest in future life (and hence in life-sustaining treatment) is not to say that she has also lost an interest in the kinds of experience she will have while alive, and therefore, her defence of granting the advance directive moral force only covers directives suggesting that we forego life-sustaining treatment. It is far from clear, however, that we ought to grant other kinds of advance directives moral authority. For example, Kuhse suggests, matters are different when it comes to the refusal of palliative care.

Given ... that severely demented patients retain a strong interest in receiving pain and symptom control, it is far from clear that this basic interest should be trumped by the interests persons might have in knowing that the values and beliefs they hold dear (for example, a belief in the redemptive value of suffering) will find expression after their lives as persons have ceased.

Kuhse’s defence of advance directives, it thus seems, is not really one suggesting that we give priority to someone else’s interests over the demented patient’s interest, but rather one to suggest that cases most often considered to be problematic because there seems to be a conflict between doing what the advance directive suggests and doing what is in the demented patient’s interests, really holds no genuine conflict. In cases such as Margo’s we can simply act on an advance directive saying that we should forego life-sustaining treatments without violating any of the moral obligations we have towards the demented patient. So, is the argument Kuhse offers towards this conclusion convincing?

Are non-persons harmed by death?

Kuhse’s argument seemingly rests on two assumptions. First, it assumes that the ability to see oneself as existing over time is a necessary condition for personhood; second, it assumes that the same ability is a necessary condition for being harmed by death. The first of these assumptions, I believe, is not very controversial. Although there are some problems with defining personhood – due mainly to the strong conceptual link between being a person and

120 Ibid. p. 360.
having full moral status – there does seem to be some consensus regarding the claim that for a being to be a person it has to have the capacity to see itself as existing over time. According to John Locke, for example, “person” stands for “a thinking intelligent being, that has reason and reflection, and can consider itself as itself, the same thinking thing, in different times and places,” and several current philosophers have kept within this Lockean tradition. The second assumption Kuhse makes, however, seems far more controversial. So, what are her arguments in favour of the view that only beings with this ability are harmed by death?

Well, the view on which Kuhse means to build her argument is one developed by Michael Tooley. In short, Tooley argues that one can have a right to something only if the violation of that right would frustrate some corresponding desire or interest. Thus, a being can only have a right to life if death frustrates the corresponding desire to go on living. Kuhse, however, who wants to argue not merely that demented Margo does not have a right to life, but that she is not even harmed by death, formulates the argument as follows:

the wrongness of an action is related to the extent to which the action prevents some interests, desires, or preferences from being fulfilled. This basic principle explains both why it is wrong, other things being equal, to inflict pain, and why it is wrong, other things being equal, to kill a being with the desire to go on living. Any being capable of feeling pain can have a desire that the pain stop, but only a being capable of understanding that it has a prospect of future existence can have a desire to go on living, and only a continuing self – or “person” – can have an interest in continued life.

Now, Kuhse’s argument seems to rest on what has sometimes been referred to as a desire account of the badness of death. According to this account, death harms its victim only insofar as this being has a corresponding desire to continue to live. If this is the correct explanation of why death usually is a great harm to its victim, it does seem to follow that a being that is incapable of having such a desire (supposedly because he/she/it lacks the ability to see him/her/it-self as existing over time) cannot be harmed by death at all. Hence demented Margo’s non-personhood implies that she is not harmed by a fatal non-treatment decision.

This account of death’s badness is, however, problematic. For example, we do not usually believe that the death of a suicidal teenager doesn’t harm him or her just because he/she does not desire to go on living, nor do we usually

---

122 Kuhse mentions, for example, Buchanan (1988), Singer (1993), Warren (1973) and Rachels (1987).
125 See e.g. Marquis (1989).
consider death not to harm an infant or a young child who has not yet developed the necessary capacities for having the corresponding desire in continued life. All of the above, it seems, are harmed by death and the reason is simply that death – we usually assume – deprives them of a valuable future. Margo too, it seems, is harmed by death in a similar way. The fact that death deprives her of the fulfilment of future interests, desires and preferences (the desire, say, to have yet another peanut-butter-and-jelly sandwich) is quite enough – on this view – to conclude that a premature death would harm her.

Kuhse’s argument thus seems to rest on too narrow an account of why death is harmful to its victim. It is important to note, however, that even on the deprivation view of harm it is quite reasonable to assume that demented Margo – being a non-person with rather limited time left to live – is not harmed by death to the same extent as a healthy human being with the prospect of many years of life ahead. This is so because the harm caused by death seems to vary with the amount of good that is lost. We can thus assume that the loss made by demented Margo by losing out on her “happily demented” future is rather limited. Her future will not include the things that a person would normally value, such as the finishing of one’s long-term projects, rewarding long-term friendships etc., but nevertheless, the fact that her future is expected to contain more good than bad does seem to suggest, Kuhse’s argument notwithstanding, that a premature death would harm her, and it is thus in her best interest to continue living – person or not.

Conclusion

In this chapter I have considered some arguments which aim to “preserve” the moral authority of advance directives even in the cases under consideration (where we assume that the author and the later victim are distinct individuals) – and which do so by invoking the fact that the author’s “successor” is supposedly a non-person. These arguments, I have suggested, are unconvincing.

First, I have argued that Buchanan and Brock fail to show that it is an implication of the non-identity claim that demented Margo is a non-person, and that this is so mainly because they fail to show why we should accept their assumption that we ought to adopt a low threshold of psychological continuity. The bigger problem with their argument, however, is that even if we find it reasonable to conclude that the demented patient is a non-person, it is very doubtful whether this fact would have the necessary moral implications needed to secure the moral authority of the advance directive.

Whereas Buchanan and Brock’s appeal to the author’s supposed “quasi-property right” might sound like a rather solid defence of the moral authority of the advance directive, a closer look at their argument reveals that they
seem to believe that this right would somehow have to be limited by the demented patient’s interests.

The last possibility I have considered in this chapter has thus been the possibility of claiming that in cases such as Margo’s at least, we can act on an advance directive suggesting non-treatment without actually violating any of her interests, and that this is so because we ought to question the often made assumption that non-person demented Margo would be harmed by a non-treatment decision. This argument, I have argued, fails too. Even though the supposed fact that demented Margo is a non-person quite possibly makes death less harmful to her than it would be to a person, it does not suggest that it is not in her best interest to carry on living.
6. Consequentialism and advance directives

Introduction

In the last chapter I argued that it is reasonable to assume that a happily demented patient – like demented Margo – would be harmed by a non-treatment decision. Thus, Kuhse’s argument in favour of granting advance directives moral authority in cases such as Margo’s seemingly fails.

I also suggested, however, that the harm death would cause a patient like demented Margo is rather limited, mainly so because she is incapable of enjoying some of the things that persons usually enjoy – the opportunity of finishing her long-term projects, developing rewarding friendships etc. If I am right about this, it seems another type of argument presents itself. It could simply be argued that the value we derive from an established practice where advance directives are respected is very likely to outweigh the relatively little harm done to demented patients by letting them die. In other words we could justify granting advance directives moral authority on consequentialist grounds, pointing to the likeliness that respecting the advance directive, even in cases such as Margo’s, would have better consequences than any alternative policy. In this chapter I will thus consider the possibility of providing such a consequentialist defence.

I will argue that even if an established practice where advance directives are respected is likely to have some good consequences (such as the comfort and assurance most people derive from knowing that their wishes will be respected), there is a problem with the suggested argument. The problem is that it seemingly rests on the assumption that an established practice where we abide by advance directives without exception has better over-all consequences than an established practice where these kinds of cases are exempt. This assumption, I will suggest, can be questioned by pointing to some likely negative consequences of abiding by advance directives in cases such as Margo’s.
Classic utilitarianism

As the name suggests, consequentialist normative theories all share the basic assumption that the rightness or wrongness of an act depend solely on its consequences, or – as we shall later see – on the consequences of something that relates to that act, such as for example the general rule requiring acts of the same kind. Apart from this basic assumption, however, there is great diversity between different consequentialist theories.

The paradigm consequentialist theory is classic utilitarianism, i.e. hedonistic act utilitarianism. According to this theory – held, for example, by Jeremy Bentham and Henry Sidgwick – an act is morally right if and only if it maximises utility, i.e. if and only if the total amount of good minus the total amount of bad is greater than this net amount for any incompatible act available to the agent on that occasion. In addition to this, hedonism asserts that the standard of value is pleasure and the absence of pain. There are, however, other possible theories of value that are compatible with the utilitarian principle. Before moving on to the main argument of this chapter I will consider some of the differences between two such theories with regard to the advance directives.

An alternative theory of value

As we have seen, classical utilitarianism understands utility or well-being (only) in terms of pleasure and the absence of pain. Pleasure is thus, on this view, what we ultimately ought to promote. On this view, it seems, we have at least no direct reasons to abide by advance directives in a case such as Margo’s, simply because she is clearly deriving some pleasure from her existence. Many utilitarians, however, have rejected the hedonistic theory of value in favour of preferentialism.

According to a preference-utilitarian what is ultimately good is desire satisfaction or the fulfilment of preferences, and what is bad is the frustration of desires or preferences. Combined with this theory of value, then, utilitarianism tells us to act so as to maximise desire satisfaction, and this regardless of whether or not this “cashes out” in feelings of satisfaction or enjoyment.

Given this line of thought, there seems to be a very natural way for the preference utilitarian to support the established practice of granting advance directives moral authority: doing so is, if not necessarily to maximise, at least to promote that which is of final value – namely the fulfilment (rather than the frustration) of a person’s desires. So, unless we have good reason to believe that this good is outweighed by some negative consequences of this act, we ought to abide by the advance directive. (We must not forget, however, that demented Margo’s preferences, if she has any, also has to be in-

---

cluded in the calculation, as well as any other persons’ preference-fulfilment/frustration that might be the result of our actions).

One common criticism levelled against preference utilitarianism is that the fulfilment of some preferences just doesn’t seem to be that good. Is, for example, the fulfilment of former preferences conducive of value? Or what about preferences that are misinformed (such as wanting to take a sip out of the glass in front of me believing it is red wine when in fact it is acid), or just trivial (wanting to wake up on my right side every other Saturday morning)? It seems very doubtful that the fulfilment of such preferences really constitutes anything of final value.

In response to this line of criticism the preference utilitarian might suggest that only certain kinds of preferences ought to be taken into account. Some preference utilitarians believe that former preferences ought to be exempt from the utilitarian calculation. Others have argued that only informed preferences ought to count in the utilitarian calculation. Yet another common suggestion of how to demarcate between preferences that should be taken into account and those that should not, is that only preferences a person has about what he/she does or gets (so called personal or self-regarding preferences) should be taken into account. Preferences that a person has about what other people do or get, on the other hand, should not.

If either of these kinds of demarcation could and should be maintained, the metaphysical claim (i.e. the claim that pre-demented Margo and demented Margo are numerically distinct) seems to undermine the suggested preference-utilitarian defence of granting advance directives moral authority (at least in the cases under consideration). If the preferences formulated in the advance directive are preferences that are actually about someone else (as would be the case if the metaphysical claim was true) the preference utilitarian claiming that only self-regarding preferences adds to the value of someone’s life, or so it seems, would see no (direct) reason at all for conforming to those preferences. Likewise, if the preferences formulated by pre-demented Margo were formed under the impression that she would be (identical to) the future patient, it could be argued that her preferences were uninformed and therefore no value would be realised by the fulfilment of those preferences.

In other words, what theory of prudential value one combines with the principle that value should be maximised, have different implications regarding the reasons for respecting advance directives. The kind of preferentialist that could claim that we have direct reasons for abiding by advance directives in the cases under consideration, are thus only those who believe, first,

---

127 See e.g. Hare (1981).
129 See e.g. Dworkin (1977).
130 See also Furberg (2011).
that former preferences ought to count, and who see no reason to limit their
theory of value to self-regarding or informed preferences.

We must not forget, however, that another feature of classic utilitarianism
is its assertion that the moral rightness of an act depends on its consequences
for all sentient beings. Therefore, when we decide on whether or not to abide
by an advance directive we ought not restrict ourselves to Margo or the con-
sequences our acts have for her, but rather the total consequences of these
acts. Thus, even if demented Margo would have been better off had she been
kept alive, other persons might benefit from us abiding by the advance di-
rective, and if the total benefit to others outweighs the loss made by dement-
ed Margo, this is what we ought to do.

An (act-) utilitarian defence of advance directives

One commonly cited problem with the act-utilitarian principle is that it
seems to overlook justice and rights. To illustrate this problem, imagine a
case where there are five patients at a hospital who will die failing an organ
transplant. One of them needs a heart, another needs a liver, and the third
one needs a kidney and so on. In the same hospital another patient comes in
for a routine test and it is discovered that this patient’s tissue is compatible
with all five of the above mentioned transplant patients. Now, let us assume
that the act of killing the patient who is in for a routine control to save the
five transplant patients would produce more good than the act of letting the
possible donor live and letting the five transplant patients die. If this were the
case the suggested normative principle would clearly suggest that the mora-
ly right thing to do is to kill the healthy patient and transplant his o
r

131 See e.g. Harman (1977).
132 See Urmson (1953) and Brandt (1959).
it is reasonable to assume that the act of not killing the innocent patient in order to save the five others is one that conforms to the rule the following of which brings about the best over all consequences (for example, “don’t kill innocent people”), then we ought to act in this way even in cases where the act of conforming to this rule is not the particular act that will maximise utility.

Although this manoeuvre does seem to bring the utilitarian theory closer to some of our intuitions, critics have sometimes accused rule-utilitarianism of being too rigid. If we assume, for example, that “don’t kill innocent people” is the best rule (i.e. the rule the following and/or acceptance of which produces the largest sum of total welfare), then it seems, on a rule-utilitarian view, that we are allowed no exceptions – not even if we would have every reason to believe that one particular instance of abiding by such a directive would have disastrous consequences.

The implication that an act with disastrous consequences would be the right act is perhaps an unfortunate implication for any normative theory and in particular, one might think, for a normative theory that purports to be consequentialist. So, what seemed to be one of the advantages of rule-utilitarianism also seems to illustrate one of the disadvantages of this view.133 Further, if the rule-utilitarian were to respond to this challenge by suggesting more fine-grained rules that would, in a sense, allow for “exceptions” (e.g. “don’t kill innocent people except when doing so is the only way to prevent disaster” or “don’t kill innocent people except when doing so is the only way to save the lives of five people”), it seems rule-utilitarianism would be vulnerable to the challenge that it collapses into act-utilitarianism: for every situation where compliance with some rule would not produce the greatest expected good, the rule-utilitarian seems driven to favour compliance with the rule which has the case at hand as an “exception”. And if so, rule-utilitarianism will end up requiring the very same acts that act-consequentialism requires.

Now, there are possible ways for the rule utilitarian to respond to this challenge. One such possible – and, I believe, satisfactory – reply would be to point to the fact that excessively complicated rules would not be the ones to maximise the expected good. This is so because rule-utilitarianism does not rank systems of rules (only) in terms of the good that results from the compliance of them, but also in terms of the good that come from their general acceptance. Were we constantly to make exceptions to rules (or constantly change them), this would probably produce less well-being all things considered since the comfort of having a somewhat predictable set of moral rules would be lost. For example, the fact that people in general speak the truth and not only do so in cases where they believe doing so maximises

133 Naturally, there have been arguments in support of rule-utilitarianism in light of this particular objection. Brandt (1992) suggests that the rule-utilitarian simply endorses a rule requiring one to prevent disaster, even if doing so requires breaking other rules.
utility, increases the overall well-being in society because it enables people to trust one another. Or at least so the rule-utilitarian would argue. As we shall see, however, appeals to the good consequences that would come from a general acceptance of rules, are also open to the act-utilitarian.

One act-utilitarian argument in favour of granting advance directives moral authority that appeals to the very same kind of indirect consequences as do rule-utilitarianism has been forwarded by Torbjörn Tännsjö. On consequentialist grounds, Tännsjö explicitly suggests that:

An express living will should be accepted unless there are very special reasons to go against it. In particular no attempts should be made by doctors, nurses or courts to ascertain whether it is really still “valid”, when needed for a decision.

Tännsjö’s main argument in favour of this conclusion with regards to advance directives is what he dubs the argument from juridical security. According to this argument it should not be recommended that people in general – and judges, doctors, nurses and social workers in particular – in each situation act in accordance with what they believe maximises utility, even though these are the right acts. According to Tännsjö, recommending this “method of decision-making” would “result in many arbitrary and strange decisions”. Rather, he continues, the actions of doctors and nurses etc. ought to be regulated through laws or established practices, and this is so because we want to be able to predict their behaviour. In addition to this, Tännsjö continues, the system of laws and established practices that ought to regulate our actions ought to be “such that, when we adopt it and abide by it, to the degree that we actually do, it produces the best possible consequences”.

According to Tännsjö’s view, then, we can motivate a rule suggesting that we at all times abide by advance directives (even if this sometimes means our acting wrongly – i.e. in a way that does not maximise utility – when doing what these rules prescribe), at least if it is reasonable to believe that such a rule is part of the system of rules, such that when we adopt and abide by it, it maximises utility.

134 Tännsjö (1999).
135 Ibid. p. 47.
136 Ibid. p. 3.
137 Ibid.
Some problems with the suggested argument

So, then, ought we adopt – and abide by – a rule that suggests that we grant advance directives moral authority? Well, because the answer to this question – according to the utilitarian – hinges on whether or not such a rule maximises utility (or whether or not this rule is part of the system of rules that does so) we first of all seem faced with rather serious epistemic problems. How can we possibly know when a rule, or even worse perhaps, when a system of rules, maximises utility? To answer this question we first need to calculate the total consequences of the rules we adopt, but we also need to calculate the total consequences of every alternative rule. To say the least, it seems we need rather extensive empirical studies to argue in favour of one rule (or system of rules) over another.

But perhaps this is too much to ask. After all, it doesn’t seem unreasonable to assume that the use of advance directives has some positive consequences. For example, it might be suggested that the use of advance directives enables us to take precaution that our present will be respected even in future situations when we are no longer be able to make contemporaneous decisions, and this possibility will surely provide some comfort and assurance. Yet another possible reason to believe that an established practice of abiding by advance directives would have good consequences is to argue that such a practice is an essential part of the more general practice (within the healthcare system or in society at large) of respect for autonomy. Respect for patient autonomy seems to be a general principle very few of us would like to see exchanged for some paternalistic alternative. This line of thought is present also in Tännsjö’s argument. He suggests that if such a principle is respected, “then, on the whole, we will have a better system … than we would have if the ideal was not respected.”

Now, it does seem reasonable to assume that a policy of respecting advance directives would have some good consequences. However, a consequentialist defence of granting advance directives moral authority in cases such as Margo’s seemingly assumes that an established practice where we abide by advance directives without exception has better over-all consequences than an established practice where these kinds of cases are exempt. This assumption, I will argue, can be questioned, and below I will offer some arguments intended to support this claim.

Before doing so, however, it should be noted that it seems virtually impossible to say with any certainty that one of the suggested established practices is conducive of more total well-being than the other. It seems difficult enough to say which out of two rules that is conducive of more well-being when we compare, say, a rule of not abiding by advance directives with a rule that says that we ought not to – but when it comes to determining the

---

consequences of two more fine-grained rules (such as “abide by advance directives” and “abide by advance directives except in cases such as X”) these already rather serious epistemic problems seems to be magnified.

Now, the main reason for questioning the assumption that a rule prescribing that we abide by advance directives also in cases such as Margo’s, is that including theses cases seem to have some likely negative consequences. Knowing that people like happily demented Margo who is clearly deriving some pleasure out of her existence are let to die, I believe, is likely to make people at least slightly uneasy. After all, Margo is a human being, entrusted to our care and as such, we seem to have some obligations towards her. And, further, even though the question of whether or not pre-demented Margo and demented Margo are numerically distinct beings doesn’t seem to matter directly to the hedonistic utilitarian, whether or not people think of demented Margo as a different person from the person who formulated the advance directive, does seem to matter. In other words: if people think that abiding by the advance directive in a case such as Margo’s is somehow illegitimate (possibly because they think that lack of psychological connectedness results in a different person) – this also seems to speak in favour of making these kinds of cases exempt.

Apart from this, one might perhaps also expect that care takers and healthcare professionals would find it troubling to abide by advance directives in cases such as Margo’s. People working within the healthcare system, one might assume, usually has as their main objective to save life and restore well-being, and would therefore, perhaps, feel uneasy about having to let a patient like demented Margo die when her life could easily have been saved.

There are, however, also several ways to argue against the above suggestions. For example, a utilitarian defender of granting advance directives moral authority in cases such as Margo’s might very well doubt that people in general believe cases such as Margo’s to be problematic or that the advance directive is somehow “illegitimate”. Perhaps because they do not – psychological alterations notwithstanding – regard demented Margo as a different person from the author. Also, it might be argued, people in general are more worried about paternalism within the healthcare system, than they are about the possibility of letting a happy human being die.

Yet another argument available to the utilitarian who believes we ought not make exceptions to the policy of abiding by advance directives, would perhaps be to point to the fact that cases such as Margo’s are likely to be rare. So rare, she might continue, that the negative consequences are not very likely to be very serious. However, if it is true that cases such as Margo are very rare, this could work as an argument also in favour of a policy where we do exempt these kinds of cases. This is so, one might suggest, because if they are rare it is not very likely that making exceptions for them would have a negative affect at all on the positive effect that might be expected by a general policy suggesting that we abide by advance directives. For example, one
might argue, the general level of comfort and assurance people derive from knowing that their wishes will be respected would be preserved even in light of these few exceptions.

A problem with all arguments above, however, is that they rest on pure speculation, for example when it comes to questions such as whether or not people in general believe that it is legitimate to abide by advance directives in cases such as Margo’s. To make a steady case for either view it thus seems we need empirical data that relate to these kinds of questions.

Conclusion

In this chapter my main objective has been to consider the possibility of providing a consequentialist defence of granting advance directives moral authority, even if we assume that a case such as Margo’s presents us with an inter-personal conflict of interests. As we have seen, only preference utilitarians seemed able to provide such a defence by appeals to direct consequences. Whether or not such a defence would really hold in a case where we assume that the author of the advance directive and the later patient are numerically distinct beings, however, depends also on how a preference utilitarian chooses to demarcate between preferences that ought to count in the utilitarian calculations and preferences that ought not.

Even if a hedonistic utilitarian is unlikely to defend the moral authority of advance directives by appeals to direct consequences, we have seen that a defence can be provided by appeals to indirect consequences. The problem with this view, I suggested, is that it rests on the assumption that an established practice where we abide by advance directives without exception has better over-all consequences than an established practice where these kinds of cases are exempt. This assumption, I argued, can be questioned by pointing to some likely negative consequences of abiding by advance directives in cases such as Margo’s.
Summary

The arguments in this book have taken as their point of departure one of the most commonly discussed medical-ethical arguments against granting advance directives moral force: the Objection from Personal Identity. The adherers of this objection basically asserts that when there is lacking psychological continuity between the person who formulated the advance directive and the later patient to whom it supposedly applies, this seriously threatens the directive’s moral authority. And, further, that this is so because lacking sufficient psychological continuity implies that the author of the advance directive is numerically distinct from the later patient.

Although this argument has some initial appeal, most philosophers in the advance directives debate maintain that the Objection from Personal Identity fails, but suggest different reasons as to why. In this book I have investigated some of the most influential of these arguments. Here follows a brief summary of what has been said so far.

The first argument I considered was the claim that the Objection from Personal Identity rests on the faulty assumption that we persist by virtue of our psychology (the Psychological View). Rather, critics have recently suggested, we are animals, and as such we persist by virtue of our biology. If this is true, it seems the adherer of the Objection from Personal Identity is mistaken as no degree of psychological discontinuity could threaten the identity relation between the author of the advance directive and the later patient. So, then, are recent animalist arguments — intended to show that we ought to abandon the Psychological View in favour of Animalism — convincing? I argued that they are not, and that a case can be made for the Psychological View. It thus seems there is no reason to dismiss the Objection from Personal Identity on the basis that it presupposes a faulty view on identity.

In the next chapter, however, I suggested that there is another possible argument to make against the Objection from Personal Identity. This is to suggest that (strict) identity is not what matters to our moral and practical concerns, and that therefore the proponent of the Objection from Personal Iden-
tity is mistaken in believing that identity is a necessary condition for the moral authority of advance directives.

This argument, I suggested, is sound: although a four-dimensionalist view of persons could possibly escape the implication of the argument from fission, the more traditional three-dimensionalist view cannot. Therefore, it could reasonably be argued that lacking psychological continuity does not threaten the moral authority because it implies that the author of the advance directive is numerically distinct from the later patient (even though, as I argued in chapter two, this could quite possibly be an implication of lacking psychological continuity). Still, it seems overwhelmingly natural to conclude that psychological continuity is what matters to our moral and practical concerns, and therefore it is still reasonable to assume that the relation that matters to the moral authority of an advance directive is threatened by severe and permanent neurological damage, for example that due to advanced dementia. Or so I argued.

In chapter four I considered an argument that took as its point of departure the fact that the non-identity claim seemingly implies that the author of the advance directive has gone out of existence. If this is the case, the argument continues, nothing that we do can retroactively affect the value of this person’s already completed “life”. This latter claim, I argued, is mistaken. At least some of our interests, I suggested, are such that their thwarting can harm us or make us worse off even after we have gone out of existence. Therefore, the non-identity claim (i.e. the claim that the author of the advance directive and the later patient are numerically distinct individuals) does not give us reason to simply disregard the interests of the former person.

At the end of chapter four, however, I pointed to the fact that even if we accept the claim that we can be harmed by posthumous events (and thus reject the immunity thesis) this does very little to support the moral authority of advance directives. It merely allows for the possibility of a genuine conflict between the “surviving” interests of the former person and the patient’s current ones. Whose interest that ought to be given priority, it seems, is till an open question.

Therefore, the task in the two following chapters was to critically examine two different suggestions of how we might normatively argue to the affect that we ought to grant the advance directive moral authority, even if the case at hand ought to be viewed as an inter-personal conflict.

The first suggestion was an argument offered by Buchanan and Brock. Although Buchanan and Brock agreed both with the assumption that the correct theory of personal identity will include the claim that psychological continuity is (at least) a necessary condition, as well as with the suggestion that we can be made worse off by events that occur after our death, they
argue that the inter-personal conflict ought to be solved in favour of the former person – i.e. in favour of the advance directive. When the demented patient is a non-person, they suggested, we can provide a justification for granting the advance directive moral authority by appeals to “a right of disposal” or a “quasi-property right”.

I argued, however, that Buchanan and Brock’s defence of the moral authority of advance directives is not as solid as they make it out to be. The fact that the demented patient is a non-person does not, according to their view, imply that we have no moral reason to consider her interests, and it is far from obvious from Buchanan and Brock’s argument that the interests of the current non-person always ought to be overridden in favour of the interests of the former person. I also argued against the claim that the supposed fact that the demented patient is a non-person implies that she cannot be harmed by a non-treatment decision. She is harmed, I suggested, because such a decision deprives her of a valuable future.

In chapter six I considered the possibility of a consequentialist defence for granting advance directives moral authority even in cases such as Margo’s. I argued that even if an established practice where advance directives are respected is likely to have some good consequences (such as the comfort and assurance most people derive from knowing that their wishes will be respected), there is a problem with the suggested argument. The problem is that it seemingly rests on the assumption that an established practice where we abide by advance directives without exception has better over-all consequences than an established practice where these kinds of cases are exempt. This assumption, I suggested, can be questioned by pointing to some likely negative consequences of abiding by advance directives in cases such as Margo’s.

Final words

The general aim of this book has been to critically examine some recent arguments intended to undermine the Objection from Personal Identity, and to answer the question of whether or not this objection really is an argument that we have good reason to dismiss. Having considered these arguments I think we ought to conclude, at least, that the Objection from Personal Identity cannot be as easily dismissed as the majority of philosophers in the advance directives debate seem to believe. The further question of interest is, however, what the arguments in this book, taken together, imply for the moral authority of advance directives.

Well, first of all it seems like appeals to surviving interests would provide us with at least a pro-tanto reason to abide by advance directives. However,
it also seems lack of psychological continuity (and possibly identity) between the author of the advance directive and the later incapacitated patient changes the ethical landscape in a way that makes it rather difficult, suggested normative arguments notwithstanding, to hold onto the belief that the former person’s advance directive ought to be granted moral authority. In fact, I believe that we should let go of this belief: a conflict between the surviving interests of the former person and the interests of the later patient ought to be resolved in favour of the patient. There is thus good reason to seriously question the rising tide of opinion in favour of advance directives.

Further issues

There are several issues touched upon in this book that deserves further attention, but that have been beyond the scope of this book. Let me mention some of these.

In this book my discussion has been limited to the moral authority of advance directives under particular circumstances, namely those when we have reason to question the assumption that the person who formulated the advance directive is the same person as the later patient. But what about cases where we have no reason to question this assumption? Are appeals to (precedent) autonomy, for example, sufficient to support the moral authority of advance directives under “ordinary” circumstances? If not, are there other possible justifications?

I argued in chapter four that the immunity thesis fails and that a case can be made for posthumous harm. That there are such harms, however, says little about the normative significance of “surviving interests” or what weight we ought to grant them in relation to other considerations.

In chapter five I considered Buchanan and Brock’s suggestion of what threshold level of psychological continuity we ought to employ when it comes to the moral authority of advance directives. Although I argued that they fail to show that we ought to employ a low (conservative) level, I did not argue in favour of any detailed alternative. However, if lacking sufficient psychological continuity really threatens the moral authority of an otherwise legitimate advance directive, it seems we need to formulate such a suggestion.

Finally, many of the arguments suggested in chapter six would benefit greatly from empirical studies. For example, one would like to know how people in general respond to cases such as Margo’s. Do they believe that the ad-
vance directive is authoritative, or do they think that it ought to be overriden in favour of the interests of the demented patient?


Stockholm Studies in Philosophy
Published by Stockholm University
Editors: Gunnar Svensson and Torbjörn Tännsjö


Subscriptions to the series and orders for single volumes should be addressed to any international bookseller or directly to the distributor: eddy.se ab, P.O. Box 1310, SE-621 24 Visby, Sweden.
Phone: +46 498 253900
Fax: +46 498 249789
E-mail: order@bokorder.se
http://acta.bokorder.se
ACTA UNIVERSITATIS STOCKHOLMIENSIS

Corpus Troporum Romanica Stockholmiensis
Stockholm Cinema Studies
Stockholm Economic Studies. Pamphlet Series
Stockholm Oriental Studies
Stockholm Slavic Studies
Stockholm Studies in Baltic Languages
Stockholm Studies in Classical Archaeology
Stockholm Studies in Comparative Religion
Stockholm Studies in Economic History
Stockholm Studies in Educational Psychology
Stockholm Studies in English
Stockholm Studies in Ethnology
Stockholm Studies in Film History
Stockholm Studies in History
Stockholm Studies in History of Art
Stockholm Studies in History of Ideas
Stockholm Studies in History of Literature
Stockholm Studies in Human Geography
Stockholm Studies in Linguistics
Stockholm Studies in Modern Philology. N.S.
Stockholm Studies in Musicology
Stockholm Studies in Philosophy
Stockholm Studies in Psychology
Stockholm Studies in Russian Literature
Stockholm Studies in Scandinavian Philology. N.S.
Stockholm Studies in Social Anthropology. N.S.
Stockholm Studies in Sociology. N.S.
Stockholm Studies in Statistics
Stockholm Theatre Studies
Stockholmer Germanistische Forschungen
Studia Fennica Stockholmiensia
Studia Graeca Stockholmiensis. Series Graeca
Studia Graeca Stockholmiensis. Series Neohellenica
Studia Juridica Stockholmiensis
Studia Latina Stockholmiensis
Studies in North-European Archaeology