

My Body To Do With As I See Fit?

Didde Boisen Andersen

My Body To Do With As I See Fit?

PhD Dissertation

Politica

© Forlaget Politica and the author 2021

ISBN: 978-87-7335-286-1

Cover: Svend Siune
Print: Fællestrykkeriet, Aarhus University
Layout: Annette Bruun Andersen

Submitted September 5, 2021
The public defense takes place December 16, 2021
Published December 2021

Forlaget Politica
c/o Department of Political Science
Aarhus BSS, Aarhus University
Bartholins Allé 7
DK-8000 Aarhus C
Denmark

Table of Contents

Acknowledgements	7
Chapter 1: Introduction.....	11
Chapter 2: Methodology	15
Reflective Equilibrium.....	15
Consistency	18
The Use of Cases	18
Fact-free Principles and Principles for Regulation	22
Chapter 3: Personal Autonomy.....	23
The Conditions of Autonomy.....	23
Living Autonomously.....	26
Autonomy-Based Self-Regarding Duties.....	30
Chapter 4: Anti-Paternalism, Paternalism and the Non-Harm Condition.....	33
Paternalism and Anti-Paternalism.....	34
The Non-Harm Condition	36
Chapter 5: What's Up with Personal Identity?	39
The Shifting Identity Argument	41
The Personal Non-Identity Thesis.....	42
Advance Directives and the Personal Identity Problem	45
The Trilemma.....	46
Chapter 6: Conclusion.....	51
English Summary	55
Dansk resumé.....	59
Bibliography	63

Acknowledgements

As a young bachelor student of Public Health Science, I did not know that I was missing something until the day I was introduced to Political Theory. However, from that day on, I knew that I needed Political Theory to be part of my academic life. I therefore spent all the ECTS points that I could on getting to know Political Theory better and my happiness peaked upon being accepted to be a PhD student of Political Theory. Life as a PhD student, however, has been a great mixture of ups and downs and in spite of my love for Political Theory, I would not have been able to go through this journey without the help of the people around me. I am therefore very happy to be given the opportunity to express my gratitude to those people who have helped and supported me greatly.

Recently, I stumbled upon the ABCD reviewer-award (Above and Beyond Call of Duty Award) given to those reviewers ‘who were deemed worthy of special recognition for the helpfulness, civility, extensiveness and insight of their reviews.’¹ I would like to suggest that we make such a prize for supervisors and in such case; I would like to nominate my supervisor, Søren Flinch Midgaard. Søren, thank you for being the best supervisor one could have ever hoped for, for spending countless hours reading and discussing my papers and ideas with me and for helping me become a better researcher. I would also like to thank you for always challenging my liberal attitudes and for making me come up with arguments that even a paternalist would have a hard time refusing—I am positive that my papers are better because of it. Last, but not least thank for you for seeing the light in me when I first came to you with a dream of becoming a researcher, for your support when I was struggling and for your friendship. I would also like to thank my second supervisor, Kasper Lippert-Rasmussen, without whom my dissertation would definitely not have been of the same quality. I am so grateful to have been able to learn from you Kasper and to have been given the opportunity to discuss and receive comments from such a brilliant philosopher as you are. I was very proud when someone told me that my most recent paper sounded like something Kasper could have come up with. I would also like to thank you for always inviting me to conferences and workshops, for letting me exploit your extended research network, for your kindness and for your guidance.

I would like to express my gratitude to the members of the Political Theory section: Tore Vincents Olsen, Fabio Wolkenstein, Göran Duus-Otterström (a special thanks for encouraging and helping me to pursue my idea about the

¹ <https://omt.aom.org/awards/abcdaward>

non-identity problem), Juha Tuovinen, Renaud-Phillippe Garner, Marion Kathe Godman and Kristian Kriegbaum Jensen. I want to thank you for always being ready to discuss and comment on my papers and ideas and for having made it fun and inspiring to be part of the political theory section. To the PhD group in the political theory section Andreas Bengtson, Anne-Sofie Greisen Højlund, Lauritz Aastrup Munch, Anna Christine Hjuler Dorf, Jake Lehrle-Fry and Jonas Franzen it has been pleasure having such kind and talented people around me to share ups and downs with. A special thanks to my office mates Jens Jørund Tyssedal, Simone Sommer Degn and Vilde Lunnan Djuve for all the good and fun times at the office and for making it a joy to come to work. Andreas Brøgger Albertsen and Lasse Nielsen, I owe you both a special thank you for all of your help, guidance and support throughout my PhD. I am happy to have had you to look up to and consider both of you to be my good friends. Last, but not least I want to thank Viki Møller Lyngby Pedersen for being my lunch-box partner, my officemate, my paternalism-discussion-partner, but most of all for being my dear friend. Thank you for always being there for me and for supporting and encouraging me when I need it. It has been a privilege and so much fun. I hope that we will keep on being close friends in the future.

It has been a pleasure and an inspiration to be part of the Department of Political Science at Aarhus University. I am grateful for everybody I have met and I thank you all for being kind and helpful colleagues. A special thanks to the PhD leader, Christoffer Green-Pedersen, for always taken the time to answer any questions that I had and for guiding me when I had a hard time balancing life and work. Thank you to the members of the technical and administrative staff Maj Thimm Carlsen, Birgit Kanstrup, Annette Bruun Andersen, Helle M.H. Bundgaard, Lene Hjøllund, Malene Poulsen, Ruth Ramm, Susanne Vang, Anja Dalsgaard, Line Kjær Vesterbæk, Ida Warburg, Njall Beuschel and Henrik Normann Jørgensen for always answering my questions and helping me with a smile on your faces. Last, but not least, being part of such a great PhD group, as the group of Political Science is, has made my time as a PhD student much more fun and comfortable. Thank you for all the talks by the coffee machines, for all the fun times and for the empathetic environment in the group. I want to thank all of the members of the PhD group, but to name a few I especially want to thank Marie Kaldahl, Trine Høj Fjendbo, Amalie Trangbæk, Sarah Yde Junge, Thomas Artmann Kristensen and Mathilde Cecchini.

Unfortunately, Covid-19 made it impossible for me to go on my planned stay abroad to the Oxford Uehiro Centre for Practical Ethics. However, I would like to thank Professor Thomas Douglas for being so kind in wanting to host

me, for commenting on my work and for leaving the door open for me to hopefully visit some day in the future. Luckily, my friends in Aalborg at the Centre for Philosophy and Public Policy together with The Values, Welfare and Health research group and the Centre for Philosophy of Health at University of Southern Denmark came to my rescue and were kind enough to let me visit their research environments instead. For that, I am very grateful to all who welcomed me and commented on my work. Especially, I would like to thank Professor Jørn Sønderholm and Associate Professor Lasse Nielsen for hosting me.

I also want to take this opportunity to thank the members of my assessment committee; Professor Thomas Schramme, Assistant Professor Jessica Begon and Associate Professor Andreas Brøgger Albertsen. Thank you for taking the time to read and comment so extensively on my dissertation. I appreciate all of your comments and your challenge for me to think further about whether limits to self-regarding actions can ever be justified.

Outside the yellow walls, I am so lucky to have great people in my life to take my mind of work and for supporting me all the way through writing this dissertation. Karoline Steen Jensen, Trine Dalsgaard Gammelby, Camilla Rahr Tatari, Marie Lass, Isabella Stendrup Rasmussen, Tanja Nørrelund Jensen, Anne K. Devantier, Laura Lassen Jørgensen and Mette Wiuff Korsholm thank you for always being there for me, for listening to my endless talks about my newest idea and for all the fun times. I am grateful to have such great friends in my life. I want to thank my parents (all four of them): Linda Friis Boisen, Hans Henrik Dahl Andersen, Vibeke Strand Andersen and Brian Nielsen. Thank you for raising me to believe that I can do anything and for always being there for me when I need you. I also want to thank you for being amazing grandparents to my children and for stepping in whenever we need an extra hand. Thank you to my siblings Kristian Boisen Andersen and Emma Boisen Andersen for making sure that I know how to make my case—I am sure that my arguments today are better because of all our discussions throughout our childhood. A special thanks to you, Emma for always being there for me and for being the greatest aunt one could ever imagine to my children. I also owe thanks to my brother in law, Mikkel Witt Poulsen for all the fun times in our Food Club and for being a great uncle to my children.

Andrea and Martha, the joy of being your mother is incomparable to anything else in the world. You have taught me so much about life and about what is important and I love you more than I can ever express; though I am not sure you made it easier to write this dissertation! However, you definitely motivated me to work more efficiently so that I could leave the office at a reasonable hour to be with you.

The biggest thanks of them all I owe to my fiancé, Thomas Mahler Christensen. Thomas, I hope that you know how much I appreciate you and what a great father I think you are to our children. I want to thank you for challenging me to think about the real world and not just let me get lost in my spacy thought experiments. Thank you so much for always believing in me and for taking over at home when I sometimes disappear into my own little bubble of work. You are my best friend and I love you.

Didde Boisen Andersen
Aarhus, November 2021

Chapter 1: Introduction

Should I be allowed to do whatever I want with or to my own body as long as the contemplated acts affect only myself and my body? Should I, for example, be allowed to mutilate or change my own body in any way I see fit? Maybe I have always dreamt of looking like a tiger and therefore want to have tiger stripes tattooed all over my body,² or maybe I want to have my breasts enlarged considerably, because (I think) this look would be more in line with who I really am. Sometimes we harm or risk harming our own body in ways that will have profound effects on our wellbeing and/or our autonomy. Yet if such acts are voluntary, perhaps even crucial to us being who we want to be, then what is the problem?

A core liberal principles—which still shapes the contemporary debate about whether there are any justifiable limits on self-harming choices—is John Stuart Mill’s well-known harm principle:

the only purpose for which power can be rightfully exercised over any member of a civilized community, against his will, is to prevent harm to others. His own good, either physical or moral, is not a sufficient warrant. He cannot rightfully be compelled to do or forbear because it will be better for him to do so, because it will make him happier, because, in the opinions of others, to do so would be wise, or even right (Mill 2002 [1859]: 8).

Committing oneself to the harm principle means that in self-regarding matters, we seem to have no good reasons for restricting people’s conduct even in cases where significant and irreversible harm is at stake. This is true, according to Mill and Joel Feinberg, when people are acting in a sufficiently voluntary way. If, by contrast, a person is acting involuntarily, such actions are, ‘from the moral point of view, no different from the actions of someone else to which [the person] have not had an opportunity to consent [to]’ (Feinberg 1971: 107). In this regard, Feinberg relies heavily on *the Volenti Doctrine*, which says that ‘To one who consents no harm is done’ (1971: 106). This means that when a person consents (given that the consent satisfies relevant criteria for being deemed valid), or when she, in one-person cases where the agent inflicts harm on herself, acts voluntarily, liberals have no good reason to restrict the behaviour in question (cf. Feinberg 1971: 120; Shafer-Landau 2005: 171).

² This case is from my paper on self-regarding duties, autonomous living, and body modifications (Andersen 2020: 192-193).

The ultimate value behind the harm principle is arguably personal sovereignty or autonomy. That is, it seems to be a fundamental liberal premise that when individual persons are capable of making autonomous choices and planning their own lives in accordance with their own conception of the good life, they also have the right to decide for themselves within all areas of life (when the interests of other people are not directly involved) (cf. Arneson 2005: 259; Feinberg 1986: 59; Carter 2018: 826; Grill 2012; Savulescu 2007: 33-34). Accordingly, liberals adhering to the harm principle need to respect people's voluntary choices, even when these choices clearly seem unwise (cf. Andersen 2021: 129-130; Schramme 2008: 9; Carter 2018: 826).

Yet our liberal commitments of this kind may come under pressure in certain cases. Consider, for example, the case of a person who volunteers to become a slave, the case of the German Cannibal from 2001, where a man volunteered to be eaten (cf. Dworkin 2012: 317), or google 'Lizard Man', and you will be met with a bunch of pictures of a man who looks like a cross between a human being and a lizard, as a result of a series of treatments he voluntarily subjected himself to (cf. Schramme 2008: 11). Even when we assume that these people were acting voluntarily and were competent enough to make these decisions, such test cases might still raise some doubts about a consistent liberal dismissal of any interference in self-regarding matters (cf. Dworkin 2012; Shafer-Landau 2005: 170; Feinberg 1990: 128-129).

This dissertation critically discusses some of the liberty-limiting principles that have been proposed in the literature of relevance to self-harming conduct (e.g. principles related to paternalism, a liberal strategy of reconciliation, and the claim that we have self-regarding duties). It focuses, as indicated, especially on cases related to bodily self-harm; for instance, extreme body modification, living organ donation, and decisions regarding medical treatment. Yet, I believe that the arguments I put forward apply more generally to cases of self-harming behaviour. However, the body seems to be a paradigmatic case of something that is self-regarding, and many authors have tried to defend at least a negative right entitling people to decide what happens to their own body (cf. Schofield 2018: 77; Fabre 2006: 1; Andersen 2020: 179; Wilkinson 2011: 16). In light of this, I ask this overall research question:

What are the moral limits (if any) to what one may do to one's own body?

I try to answer this question by canvassing the following three questions that are arguably integral to it: 1) Is it morally wrong to inflict harm on one's own body? 2) Can harm count as a relevant reason for restricting self-harming behaviour? 3) Can we justify restrictions on self-harming behaviour in a non-paternalistic way? I examine the questions in the following four papers. The

shortened titles (in parentheses) are used in the summary to make it more reader-friendly.

- Andersen, DB. (2020). Self-Regarding Duties, Autonomous Living, and Body Modification. *Ethical Perspectives* 27(2): 177-144 ('Self-regarding duties')
- Andersen, DB. (2021). May I give my heart away? On the permissibility of living vital organ donation. *Bioethics*, DOI: 10.1111/bioe.12935 ('Organ Donation')
- Andersen, DB. (2021). I Have Got a Personal Non-identity Problem: On What We Owe Our Future Selves, *Res Publica* 27: 129-144 ('Non-identity')
- Andersen, DB. On the Moral Authority of Advance Directives, unpublished manuscript ('Advance Directives')

The aim of this summary is to present the arguments I put forward in the papers in a unified form. The thesis I defend is that there are few (if any) substantial moral limits regarding what one may do to one's own body. I have organized the summary in the following manner. In chapter 2, I present the overall methodology of the dissertation. In chapter 3, I present and motivate the account of personal autonomy that my argument employs. Especially, I show how posing substantive requirements to what it means to live an autonomous life seems to be in clear tension with the core ideas of autonomy. This means, I argue, that the only moral limits to what one may do to one's own body are the procedural requirements to the process behind one's decisions about what kind of life one wants to live.

Following this, I present my argument for the claim that we are not acting morally wrong even when we modify or mutilate our own body in an irreversible and severe manner. Specifically, I examine whether we have autonomy-based self-regarding duties regarding what we ought to do (or not to do) to our own bodies and grant that if such self-regarding duties do exist, they initially seem to suggest moral limits to what one may do to one's own body. However, as I argue in 'Self-regarding Duties', such duties only impose *prima facie* restrictive moral limits on what one may do to one's own body. As it turns out, such duties should be seen against the backdrop of (and sometimes in opposition to) a more important concern of living autonomously—a concern which on my view is key.

In chapter 4, I reject the suggested liberty-limiting intuition that risk of harm is always a good and relevant reason for restricting self-regarding behaviour. I challenge what I call *the non-harm condition*, which says that living vital organ donation is never ethically permissible when significant harm is done to the donor, the harm being dying earlier than one would otherwise

have died (Organ Donation: 4). Accordingly, the non-harm condition argues that the harm in question is a good and relevant reason in favour of limiting who should be considered eligible for living vital organ donation. If we were to accept this kind of restriction, we need to abandon the harm principle and accept, as I argue in ‘Organ Donation’, an objectionable form of paternalism. This rejection gives indirect support to the claim that there are no moral limits to what one may do to one’s own body.

In chapter 5, I present what is known as the *liberal strategy of reconciliation*, which is a strategy that seeks to justify restrictions on self-harming behaviour that we find unsettling in a *non*-paternalistic way. One way to do so is to show that the putatively self-harming behaviour in question is in fact a sort of *other*-regarding harm. In that way, we can justify restrictions without compromising the harm principle (Non-identity: 130).³ The reconciliation strategy addressed in chapter 5 derives from rethinking our notion of personal identity. That is, the argument put forward in favour of the reconciliation strategy is that instead of seeing ourselves as being one and the same person throughout our biological life, we need to acknowledge that sometimes we do in fact shift identity in such a way that our future selves should be considered a numerically different person (Non-identity: 130). Accepting this notion of personal identity means that when we act in ways that will impose severe and irreversible consequences upon our future selves, we are in fact harming another person. In chapter 6, I show how this line of thought at first seems to justify restrictions on putatively self-harming behaviour and limit our right to medical self-determination. However, despite the ingenuity and attractiveness of these arguments, I show how they are not necessarily successful in imposing limits on what one may do to one’s own body, even when we abandon our standard notion of personal identity.

In chapter 6, I pull the strings together from my answers to the three questions and conclude on how we should respond to the overall research question.

³ On the liberal strategy of reconciliation cf. Viki M. L. Pedersen (2019); Peter de Marneffe (2006); Feinberg (1986).

Chapter 2: Methodology

This dissertation is written within the field of analytical political theory, which is characterized, in its nature, as a normative discipline, since it seeks to show how the world *should* be and what we *ought* to do (or not to do) rather than how the world *is* or what we actually *do* (Holtug 2011: 277). The research question thus requires normative discussions of potentially liberty-limiting principles pertaining to what one may do to one's own body. In particular, I employ the method of reflective equilibrium in order to test, identify and formulate particular moral principles related to the question of what the moral limits (if any) are to what one may do to one's own body.

However, analytical political theory also involves conceptualizing, evaluating, and testing particular theories, principles or claims, in order to show whether such should be considered valid in light of more or less universal methodological requirements (cf. Holtug 2011, List & Valentini 2016: 526, 539). In this dissertation, I focus especially on the methodological requirement of consistency and show how this requirement forces us to revise or maybe completely abandon one or more normative claims. Below, I present the methodological framework of the dissertation.

Reflective Equilibrium

Reflective equilibrium is the standard method in contemporary analytical political theory. Reflective equilibrium can be seen as a method of theory (or principle) testing, in which we test theories or a set of moral principles by considering the implications of those principles in light of different cases. It allows us to show and/or defend whether to endorse a particular set of moral principles (List & Valentini 2016: 541-542).

More precisely, when working within the framework of reflective equilibrium, one should aim to construct a coherent system (a reflective equilibrium) in which there is acceptable coherence between (a) our considered moral judgements (some say intuitions)⁴ about particular cases, (b) a set of moral

⁴ I will be using the terms 'considered moral judgements' and 'intuitions' interchangeably. Note, however, that intuitions should not simply be seen as "gut feelings" or first impressions, but, as John Rawls himself puts it, as a 'judgements ... given under conditions favourable for deliberation and judgment in general' (1999: 42).

principles and (c) relevant background theories (Daniels 1979: 258-259; Daniels 2020; Slavny et al. 2020).⁵ The considered moral judgments in (a) relevant for constructing a reflective equilibrium are those we feel relatively confident about when confronted with a particular case, since they have been made under favourable conditions (Daniels 1979: 258). As John Rawls (1999: 42), puts it:

we can discard those judgments made with hesitation, or in which we have little confidence. Similarly, those given when we are upset or frightened, or when we stand to gain one way or the other can be left aside. All these judgments are likely to be erroneous or to be influenced by an excessive attention to our own interests.

As the considered moral judgments in (a) are about one particular case, a moral principle in (b) is, by contrast, a statement that tells ‘agents what (they ought, or ought not) to do’ (Cohen 2003: 211) potentially regarding *more than one* particular case (List & Valentini 2016: 535). Moreover, I follow Norman Daniels’ rather broad interpretation of what a background theory in (c) is (exemplified with regards to Rawl’s theory of justice): ‘a theory of the person, a theory of procedural justice, general social theory, and a theory of the role of morality in society (including the ideal of a well-ordered society)’ (1979: 260).

The idea is that we should rectify incoherence between (a), (b) and (c), since only then will we be able to justify the normative claims (Daniels 2020; Slavny et al. 2020: 1). That is, the moral principles or normative claims obtain their justification by being part of a coherent system of (a), (b) and (c) (Daniels 2020). One of the relevant criteria for the degree of coherence is consistency, since inconsistency decreases the degree of coherence (Lippert-Rasmussen 2005: 281; Holtug 2011: 291). Furthermore, the more explanatory relations there seem to be between (a), (b) and (c), the more the degree of coherence increases. That is, some elements in a coherent system need to ‘provide support or provide a best explanation for others’ (Daniels 2020; see also Lippert-Rasmussen 2005: 281).

The method involves a deliberative process where one seeks to align our considered judgements about particular cases and our moral principles, going back and forth, in light of the relevant background theories (Lippert-Rasmussen 2011: 298). We may need to revise our considered judgments and/or our moral principles to find an acceptable coherent alignment of these (Hansen 2016: 23-24). Often the way to go about this is to consider variations of a particular set of cases, ‘test’ the moral principles against these cases to see how we judge the implications of the principle and then revise and/or specify the

⁵ Yet as Rawls puts it reaching an actual reflective equilibrium ‘is a point at infinity we can never reach’ (1995: 142)

moral principle to accommodate our considered judgements. We may also revise our considered judgements if they do not fit the moral principles we become inclined to accept (Daniels 2020).

But what role should background theories play? Some have said that reaching a stage where our considered judgments and the moral principles fit together in a coherent system, we have ‘only’ reached a *narrow* reflective equilibrium (cf. Daniels 2020). However, such equilibrium does not take background theories into account, and therefore we need to do so to achieve a *wider* reflective equilibrium. As noted, I understand background theories very broadly as facts and theories about human nature and different metaphysical assumptions (cf. Daniels 1979: 258, 260).

I believe there are several ways of taking relevant background theories into account and by doing so widening the reflective equilibrium one seeks to achieve. One way to take background theories into account is to very explicitly investigate what happens to our considered judgments and our moral principles when we change or accept different background theories. In ‘Non-identity’, I explore what happens to our commonly accepted moral principles when we accept a reductionist notion of personal identity. Take for example Mill’s harm principle, which deems it impermissible to forcibly prevent acts of merely self-regarding harm. Only actions that harm other people can rightfully be prevented (Non-identity: 129; Mill [1859] 2000). The implications of the harm principle seem to depend on a specific theory of personal identity, in that the principle seems to assume that there is a clear distinction between self-regarding and other-regarding harm. However, as I write in ‘Non-identity’, accepting a reductionist view on personal identity implies that ‘the moral difference between us and other people becomes less important. This indicates a sceptical view towards the moral principles relying on an assumption of a deep morally relevant difference between the person him- or herself and other people’ (Non-identity: 142). Since the harm principle is a principle of such kind, appealing to a reductionist view on personal identity seem to indicate that the harm principle, as it is, cannot be useful in terms of reaching a reflective equilibrium.⁶

⁶ In the paper, I suggest that we might revise the harm principle such that it should be seen as a scalar principle, i.e., ‘the less psychologically connected one is to one’s past self, the more state interventions are justified since the less connected this ‘other’ future person will be’ (Non-identity: 142)

Consistency

In my thesis, I assess, in part, as mentioned above, the consistency of particular claims and principles. Consistency is a common method or criterion of theory testing that many disciplines besides political theory aspire to achieve. However, it seems particularly important when we seek to identify universal moral principles, i.e., moral principles applicable to more than one particular case or person (cf. Holtug 2011: 282; List & Valentini 2016: 539). By judging whether a particular set of claims is logically consistent, we make sure that endorsing this set of claims does not involve self-contradictions. Everything and nothing seem to follow from a set of inconsistent claims, since a set of claims ‘containing both “*p*” and “*not p*”, is of little use, whether for explanatory, predictive, evaluative, or prescriptive purposes’ (List & Valentini 2016: 539; see also Holtug 2011: 282; Weber 2000 [1922]: 163-164). In ‘Advance Directives’, I use this *internal* requirement, as some call it (cf. List & Valentini 2016: 539), explicitly to show that with regards to the moral authority of advance directives, we are in fact facing a trilemma where the demand for consistency forces us to reject one of the three normative claims (Holtug 2011: 283). Importantly, however, showing that a particular set of normative claims are inconsistent does not in itself give us any clear answer concerning which claim to reject (Holtug 2011: 283). In ‘Advance Directives’, I remain neutral on the question which claim to reject, since my aim is, more modestly, simply to point out an inconsistency.

The Use of Cases

Whether working within the framework of reflective equilibrium or testing theories, principles or claims by appealing to demands of consistency, a central component in analytical political theory is the use of cases. Cases may be hypothetical or real-life, and they can have multiple functions in the process of theory testing (e.g. when we seek reflective equilibrium). Cases are supposed to help us focus, disentangle and assess moral reason in order to either test, support, formulate or systematize moral principles (cf. Slavny et al. 2020: 5).

Hypothetical cases allow us to design a case that isolates the factor we want to examine and are therefore often preferable to real-life cases. When using cases to elicit moral judgements, we must ensure that the judgments are not affected by distracting features in the case. It is therefore an ambition when formulating cases that they are as ‘clean’ as possible (Tadros 2011: 7). Real-life cases are often ‘messy’, meaning that they typically involve complexity and nuances and may cloud or mud our judgements about a particular case

(Tadros 2011: 7). In this way, it may be hard to know which factors actually determine our judgments about particular cases. This line of thought is similar to experiments or case-control studies in empirical science, where we also aspire to make cases as clean as possible in order to test a specific factor. For example, if we want to test the effect of an intervention, we need to make sure that the effect is not affected by other factors (cf. Holtug 2011: 286-287).

However, using hypothetical cases instead of real-life cases has raised concerns about whether our moral judgments about such cases can be trusted. If we use more or less unrealistic hypothetical cases, our judgments risk being affected by our inability to imagine such a case, since our capacity to form ‘correct moral intuitions about cases does not extend to far-fetched cases’ (Elster 2011: 242) or so the objection goes (cf. Tadros 2011: 8; Brownlee & Stemplowska 2017: 23). The worry is that when cases are too far from reality, we find it hard to fill in the blanks of, for instance, how particular features affect the social structure, the psychological effect of the situation or some other feature of the world (Elster 2011: 251). A colleague once asserted that he simply could not judge a case of voluntary vital organ donation, since his judgement was too affected by his inability to image that any person in their right mind would ever want to do such a thing.

Admittedly, some of my cases may seem a bit far-fetched and unrealistic. For instance, in ‘Self-regarding Duties’, I ask readers to imagine a person who breaks both his legs and lets them grow back together in a way that will make him walk more like a clown, or a person who wants to have tiger stripes tattooed all over his body. Some might argue that our judgments about such cases could be affected by the fact that it is hard to imagine that any person would do such a thing voluntarily, or that we need more information about the person’s psychological make-up or background. However, both cases are actually inspired by real-life cases, and if we should not trust people’s intuitions about such cases, we might not trust people’s intuitions about real-life cases either. Take, for example, the case mentioned in the introduction of a man who volunteers to be eaten by a cannibal. I guess reality sometimes exceeds imagination, and the concern about trusting our intuition regarding such cases might therefore not be about whether the case is real or hypothetical, but more about how different the case is from our current experiences (cf. Tadros 2011: 8).

However, it is important to keep balance in mind when using cases to test the validity of moral principles. Of course, we should take the concern mentioned above seriously and for that reason not make cases more far-fetched than they need to be. However, we should make sure that this does not come at too great a cost with respect to how clean the case needs to be (Tados 2011: 8). Sometimes we need to make strange and unrealistic cases in order to isolate the factor we want to examine (Tadros 2011: 8). With these concerns in

mind, I have tried to keep the balance of realism versus cleanliness when formulating the cases I use in my papers and be very clear about the settings of the cases.

As stated above, cases can have multiple functions, and sometimes cases are deployed simply to illustrate or specify a principle or an argument. In ‘Self-regarding Duties’, I establish that a concern for living autonomously is more important than refraining from acting in ways that will diminish or destroy one’s capacities for autonomy or limit one’s array of options. I use the cases of Tiger Tim and Leopard Leonard to illustrate the implications of my argument but also to specify more precisely what the argument entails. The cases show that the argument needs to be specified in a way that leaves room for some minor constraints on what one ought to do or not do to one’s own body in order to avoid counter-intuitive implications.

The cases of Tiger Tim and Leopard Leonard are also good examples of how cases may be used to consider whether a specific factor has any moral relevance. The two cases are designed to be identical in every aspect (harm, irreversibility, the effect on the agent’s array of options) except the factor I am interested in. The only factor that varies across the two cases is how important the specific body modification is for the fulfilment of the agent’s autonomous life plans.⁷ This means that we are comparing two scenarios that only deviate on the factor we want to investigate, which allows us to observe singular effects (cf. Holtug 2011: 286-287; Slavny et al. 2020: 9; McDermott 2008: 17-18).

Working within the framework of reflective equilibrium, I mostly use cases to ‘test’ moral principles by considering the implications of the principle in a range of cases (cf. Tadors 2011: 4). Such cases are sometimes called *corner cases*, which ‘test how the principle fares in the most extreme settings and assess its robustness’ (Slavny et al. 2020: 8), or *counterexamples* when the case challenges a principle by revealing counter-intuitive implications (cf. Slavny et al. 2020: 9). In ‘Self-regarding Duties’, I use the case of a transsexual man as a counterexample to the argument that we should never harm our bodies in an irreversible manner. If I am right that the argument has counter-intuitive implications (e.g., it seems counterintuitive that the man would be acting morally impermissibly if he were to get an operation that would prevent him from bearing children even when getting the operation is what he most desires), it gives us a *pro tanto* reason for rejecting the argument.

⁷ ‘To look like a tiger is what Tim mostly desires and it is very important in order for him to be living autonomously, whereas it is not an essential part of Leonard’s autonomous life plan look like a leopard for the rest of his life’ (Self-regarding Duties: 192-193).

The case of Mr. Lived-a-good-life is an example of a corner case (Organ Donation: 6-7):

After reflecting about his life, Mr. Lived-a-good-life decides that he has had a very good life. He is not ill, dying or even very old (let us say 65 years old). However, he has reached a stage where, if he died tomorrow, he would die a happy man. However, it is important to him that other people will also have the chance of a happy life, just as he himself has had, and he knows that the best way for him to help others is to become an organ donor. However, he knows that if he waits until he is dead, most of his organs will no longer be suitable for transplantation, and therefore, he voluntarily decides to become a living vital organ donor now.

I use the case to test our commitments to what I call the non-harm condition, that is, that vital organ donation is only ethically permissible when and only when no significant harm is done to the donor. I use the case of Mr. Lived-a-good-life to show that the non-harm condition seems objectionably paternalistic, since it is unclear how we may deny Mr. Lived-a-good-life the opportunity to become a living donor and still proclaim to be giving considerable weight to the value of autonomy (cf. Organ Donation: 6-7).

By testing the validity of moral principles in this way, it becomes clearer whether the principles should be endorsed, since we seem to have greater reason to endorse a principle if its implications are intuitive (Tadros 2011: 4, 6; Nielsen 2013: 48). This approach follows a hypothetical deductive logic, which tests a hypothesis by scrutinizing the implications derived from the hypothesis (cf. Nielsen 2013: 46). We can use the case about the transsexual man to illustrate the hypothetical deductive logic:

- I. Hypothesis: We should never harm our bodies in an irreversible manner.
- II. Implication derived from the hypothesis: the transsexual man is acting morally wrong by undergoing operations that will change his body in an irreversible manner, even when this is voluntary and an important part of his autonomous life plan.
- III. It seems counter-intuitive that this man is acting morally wrong, which weakens the hypothesis.

The hypothesis is established based on an existing theory or moral principle, and since the implication is derived from the hypothesis, the conclusion must either accept the implication and thereby strengthen the hypothesis or not accept the implication and thereby weaken the hypothesis (Nielsen 2013: 46). In this case, the examination gives us reason to revise the hypothesis since we cannot accept the implication.

Fact-free Principles and Principles for Regulation

Most of the arguments and moral principles I put forward in my dissertation should be seen as putative so-called ultimate or fact-free principles. Identifying such principles is valuable in and of itself but also necessary if we are to ever formulate principles for regulation. Ultimate principles do not say anything about how we should regulate society or practices, but they establish a foundation in light of which principles for regulation may be created⁸ (Cohen 2003: 241-242). In my discussion of self-regarding duties and autonomous living in ‘Self-regarding Duties’, I leave out some facts about the world and other normative concerns which might be necessary to take into account if we were to formulate a principle for regulation. If we for example find that we have a self-regarding duty to refrain from doing something to our own body, one important question to answer with a view to formulating any principles for regulation would be whether such duties should be enforceable by others. In ‘Organ Donation’, I take other values and various practicalities into consideration when formulating a principle on when vital organ donation is ethically permissible. This means that such a principle could at some point be used as a rule for regulation, ‘that is, as a certain type of social instrument, to be legislated and implemented’ (Cohen 2003: 241).

⁸ I follow G. A. Cohen’s perception of a principle for regulation being “a device for having certain effects,” which we adopt or not, in the light of an evaluation of its likely effects, and, therefore, in the light of an understanding of the facts’ (2003: 241).

Chapter 3: Personal Autonomy

The concept of personal autonomy⁹ (from now on autonomy) is generally understood as something that refers to ‘self-government’, ‘self-authorship’ or ‘self-determination’ etc. (cf. Christman 2020; Feinberg 1986: 27; Raz 1986). Feinberg (1986: 59) describes it as there being a sphere or a realm that exclusively belongs to the person and ‘where the interests of others are not directly involved’. Within this sphere, each person ‘must be the one to decide—for better or worse—what is to be done’ (1986: 59). But what does it mean to be autonomous? In this chapter, I specify the conditions of autonomy and how I situate myself within the debate about them. To understand the concept of autonomy, I find that we need to clearly distinguish between different aspects of autonomy, that is, what it means to be an autonomous person, to make autonomous choices and to live autonomously. These distinctions, which I present below, demonstrate the theoretical backdrop against which my dissertation has been developed.

This chapter is structured as follows: First, I briefly present the conditions of autonomy based on two of the most influential accounts of autonomy in the literature. Second, I distinguish the conditions of autonomy from what it means to be living autonomously. Especially, I show how posing substantive requirements to what it means to be living autonomously seems clearly to be in tension with core ideas of self-authorship and authenticity. Third, I explain how I apply the distinction between the conditions of autonomy from autonomous living in ‘Self-regarding Duties’ to show that autonomy-based self-regarding duties, *if they exist*, only impose prima facie moral limits to what one may do to one’s own body.

The Conditions of Autonomy

Despite the fact that Gerald Dworkin (1988) and Joseph Raz (1986) formulated their theories of autonomy more than two decades ago, they still represent the two most influential contributions to the literature. Accordingly, I rely on their accounts of autonomy to present the conditions of autonomy. In order

⁹ I focus on personal autonomy rather than moral autonomy, since personal autonomy focuses more generally on what it means to author oneself in every aspect of life, not just regarding one’s moral agency. Moral autonomy, by contrast, specifically concerns the capacity to impose moral laws on oneself or whether one’s moral principles can be considered one’s own (cf. Eidelson 2015: 139; Dworkin 1988).

to grasp the complexity of the concept of autonomy, it seems relevant to distinguish between the *capacity* for being an autonomous person and the *external* conditions necessary for making autonomous choices (Eidelson 2015: 140-141; Feinberg 1986: 28). In the next section, I add one more aspect of autonomy, namely the aspects of autonomous living.

According to Dworkin (1988: 20), to be autonomous one needs to have a certain capacity and exercise it. Dworkin conceives autonomy as:

a second-order capacity of persons to reflect critically upon their first-order preferences, desires, wishes, and so forth and the capacity to accept or attempt to change these in light of higher-order preferences and values. By exercising such a capacity, persons define their nature, give meaning and coherence to their lives, and take responsibility for the kind of person they are.

This conception of autonomy is often known as the structural or hierarchical conception, since it is important that the desires or preferences we act on reflect our higher-order preferences and values. According to Dworkin, critically reflecting upon one's first-order preferences and changing these in light of one's second-order preferences is crucial for autonomy, since it makes sure that we actually want the preference in question. If we do not want the preference in question, we should change or at least not act upon it. Take smoking. If my desire to smoke is to be considered autonomous — in accordance with who I want to be as a person — I must also want to have a desire to smoke, otherwise I should alter my (first-order) desire to smoke or at least not make it effective in my actions (Dworkin 1988: 16-18; Self-regarding Duties: 186; Andersen & Midtgaard 2017: 281).

According to Dworkin, autonomy should essentially be distinguished from concepts such as independence, freedom and liberty. However, such concepts are necessary for the reflective process to be considered one's own (Dworkin 1988: 16, 18). It is crucial that the process behind one's decisions about one's second-order values and preferences is independent. That is, one's reflections about what kind of person one wants to become must not be influenced 'by other persons or circumstances in such a fashion that we do not view those evaluations as being the person's own' (1988: 18). This is what Dworkin calls the condition of procedural independence.

To summarize: One is to be considered autonomous when and only when:

- 1) one possesses the capacity to critically reflect upon one's first-order preferences and change these in light of one's second-order preferences and values, and
- 2) the process of critical reflection is independent (cf. Self-regarding Duties: 186).

Raz's (1986) notion of autonomy also emphasizes the importance of having requisite capacities for autonomy. On his account, it is a condition for autonomy that one has the mental capacity to form sufficiently complex intentions, comprehends the means required to realize one's goals and also uses this capacity when choosing one's plan of life (1986: 372-373). This condition seems similar to the one Dworkin emphasizes.¹⁰ According to Raz's (1986: 372) notion of autonomy, autonomy also includes two external conditions: i) an adequate range of options and ii) independence. As I see it, these two external conditions are not related to the question whether a person is autonomous.¹¹ Take for example the case of a person who is imprisoned for most of her life. If having some specific capacity or mental abilities suffices for being autonomous, being imprisoned does not necessarily eliminate or reduce autonomy (cf. Eidelson 2015: 140). Yet, being imprisoned often means not having an adequate range of options and one's choices are often not independent, which renders it impossible for one's choice to be considered autonomous.

Having an adequate range of options does not only mean having enough options,¹² but the options available must satisfy several distinct criteria (Raz 1986: 374). One's range of options must include options with long-term and severe consequences but also options with short-term or minor consequences (1986: 374). For a choice to be autonomous, the options available must vary in content, since being presented with hundreds of identical options is not really a choice (1986: 375-376). Lastly, for a choice to be considered autonomous it must be a choice between different good options. This is important, according to Raz, since a choice between good or bad (e.g., a choice of becoming an electrician or a murderer) is not a choice, and therefore we need to have several morally acceptable options available (1986: 378-379).

The last condition for autonomy suggested by Raz is independence. In the same way as Dworkin emphasizes procedural independence, Raz suggests that coercion or manipulation diminishes or destroys the conditions necessary for

¹⁰ I do not mean to say that there is no relevant difference between the two conceptions of the capacity for autonomy. However, for my purposes, the differences are immaterial.

¹¹ Raz (1986: 372) does not distinguish between different aspects of autonomy. On the contrary, he actually writes that he sees capacities for autonomy in a very wide sense, including the two external conditions. More precisely, on his account of autonomy, possessing the relevant mental abilities is only one condition of autonomy, and the presence of specific external conditions is necessary for exercising such abilities. See also Benjamin Eidelson (2015: 140 note 32) on this point.

¹² Having an adequate range of options is not to be confused with maximization of the number of options (Blake 2002: 269). Having too many options may affect autonomy negatively, since it reduces our ability to assess options.

making autonomous choices, because subjecting ‘the will of one person to that of another ... violates his independence and is inconsistent with his autonomy’ (1986: 378).

Throughout the dissertation, I rely on the different conditions necessary for being an autonomous person and making autonomous choices, which is presented in this section. In ‘Self-regarding Duties’, however, I emphasize that capacities for autonomy, having adequate options to choose from and independence are only three aspects of autonomy and only necessary regarding what it means to be an autonomous person and making autonomous choices.¹³ I argue that we need to distinguish these from another aspect of autonomy, namely autonomous living (Self-regarding Duties: 185).

Living Autonomously

As I state in ‘Self-regarding Duties’ (p. 185), ‘I believe that one is living autonomously when, and only when, one is living in accordance with one’s autonomous life plan’.¹⁴ Living autonomously is, of course, related to having capacities, adequacy of options and independence but signifies something very distinct. Contrary to the conditions necessary for making autonomous choices or being an autonomous person, to be living autonomously, on my account, having an adequate array of options, living independently or even possessing specific mental abilities are not necessary conditions. It is, of course, important if one is to be living autonomously that one lives according to one’s *own* life plan. For a life plan to be considered one’s own, it needs to be the result of a reflective and independent process on one’s higher-order preferences and values, that is, the process needs to satisfy the conditions of autonomy presented above.

In my view, it is important that the concept of autonomous living we adopt is value neutral, since what seems to matter with regard to living an autonomous life is not the content of one’s life plan but rather that one is living according to one’s *own* perception of the good life. I therefore favour the structural conception of autonomy (represented by Dworkin), since it subscribes to

¹³ In the paper, I do not present the distinction as a distinction between conditions for being an autonomous person, making autonomous choices and living autonomously, but as the conditions for autonomy (mental capacities, adequate array of options and independence) being aspects of autonomy, which need to be seen against the backdrop of another more important aspect of autonomy, namely autonomous living.

¹⁴ Note that a life plan, on my view, need not be ‘long-termed and detailed – a life plan could, for example, be to live spontaneously and still be an autonomous plan’ (Self-regarding Duties: 185).

a content-neutral idea of what constitutes an autonomous life plan (Self-regarding Duties: 186; see also Christman 2020). This means that as long as the autonomous life plan can be considered one's *own* in the sense that it has been chosen through a critical reflection process on a second-order level and satisfies the condition of procedural independence, there are no restrictions on the life one is living in order for it to be considered autonomous. Accordingly, one can choose to live a life with very few or no options (e.g., to become a slave¹⁵) and still be living an autonomous life.

Behind the notion of autonomous living as content-neutral lies an assumption that what gives meaning to one's life is pursuing one's *own* preferences and projects. In this regard, I applaud authenticity, since what seems to matter is that we live according to our own perception of the good life, if we are to be considered authors of our own lives (Christman 2020; Self-regarding Duties: 186). As I write in 'Self-regarding Duties', 'my position rests on the idea that there are certain life-shaping decisions, plans, projects, and so on that are essential if one is to live an autonomous life as opposed to [and more important than] just retaining the capacities and adequate options' (Self-regarding Duties: 190).

Several authors criticize the structural notion of autonomy, since it seems too inclusive with regards to the content of what should in fact be considered an autonomous life. According to the critics, living a life with very restricted options seems inconsistent with autonomy (Oshana 1998; Christman 2020). Marina A. L. Oshana stresses that a conception of autonomous living that only requires second-order reflection and procedural independence is insufficient in the sense that there seem to be external circumstances which simply make it impossible to live autonomously even when such external circumstances reflect an individual's second-order preferences (Oshana 1998: 85).

What seems to be a paradigmatic case of a life that can never be considered autonomous even when it satisfies the conditions of second-order reflection and procedural independence is a very constrained life without control over one's external circumstances and with very limited freedom to decide for oneself, e.g., the life of an enslaved person (cf. Oshana 1998: 86-88). According to Oshana, there are objective external conditions to autonomy independent of the condition for second-order reflection and procedural independence (1998: 86). On her account, autonomy is incompatible with losing one's freedom, and

¹⁵ Throughout my time as a PhD student, I have been asked whether I find slavery morally acceptable more times than I can remember, and my answer has been and still is that with regards to autonomous living, I do not see how living a life of a slave cannot be considered an autonomous life, yet there are of course other concerns such as domination that make the practice of slavery morally problematic.

slavery therefore becomes a paradigmatic case of a life that is definitely *not* autonomous. Oshana (1998: 86) grants that it may be possible to hold a conception of the good life that does not include an interest in remaining free to choose for oneself, i.e. choosing a life of a slave autonomously (by satisfying the conditions necessary for autonomous choosing). However, since the slave's 'external environment renders him incapable of functioning in a self-governing way' (1998: 97), the slave can no longer be seen as governing their own life (i.e. living autonomously). As Oshana (1998: 87) has it, '[b]eing a slave means that *how* he shall live is no longer up to him.'

With Dworkin I find it important to distinguish between autonomy and freedom (1988: 14-15, 18), and this is why we should operate on different levels of preferences (first- and second-order). According to Dworkin, we simply fail to capture something important if we only focus on whether a person is enslaved or not. We also need to distinguish between two people who are enslaved, and one is living according to his autonomous life plan and the other is not (and on a second-order level resents living this way) (Dworkin 1988: 19).¹⁶ Usually, people might see freedom as an integral part of their autonomous life plan and appreciate on a second-order level to be able to act freely. Yet, as Dworkin (1988: 129) puts it:

There is nothing in the idea of autonomy that precludes a person from saying, 'I want to be the kind of person who acts at the command of others. I define myself as a slave and endorse those attitudes and preferences. My autonomy consists in being a slave.' If this is coherent, and I think it is, one cannot argue against such slavery on grounds of autonomy. The argument will have to appeal to some idea of what is a fitting life for a person and, this, be a direct attempt to impose a conception of what is 'good' on another person.

Maintaining objective external conditions involves only accepting those following a life plan that is acceptable from a certain understanding of the good life to be considered autonomous living (Christman 2020). In my view, it seems too restrictive and inconsistent with the ideal of living a life authored by oneself if one's array of options is limited to a specific set of 'acceptable' options. Furthermore, such objective external conditions are, according to Dworkin, also inconsistent with other important values such as commitment, loyalty and objectivity. On the structural account of autonomy, what is valuable about living autonomously is that one is living as the person one wants to be. If that involves committing to a cause or giving up one's freedom, why should that be less autonomous than living a life with complete freedom (1988: 26-28; Self-regarding Duties: 189)? Imagine if I wanted to become a

¹⁶ See also Amartya Sen's distinction between starving and fasting (1992: 52).

nun.¹⁷ My preference for becoming a nun satisfies the conditions of procedural independence and second-order reflections—this is who I want to be. I would be submitting myself to the rules of the convent and would therefore not be able to choose freely what I want to wear, where I want to live, what career to pursue etc. Some might even say that how I will live is no longer up to me. Yet, it seems plausible to conceive of my life as autonomous.

One concern about a content-neutral idea about autonomous living is that it is vulnerable to people changing their minds or revising their second-order preferences. Living as a slave entails, of course, that one cannot choose another life; it is an irreversible choice, and the slave cannot change his mind (well, he can, but he cannot make it effective in how he lives). If the slave were to change his mind about his autonomous life plan, he would therefore not be living autonomously anymore. I agree that the slave would then not be living an autonomous life, since he would not be living according to his *new* autonomous life plan. Yet in ‘Self-regarding Duties’, I show that this objection can be met.

I do this in part by considering cases in which a person chooses a certain life and thus incurs some irreversible consequences. First of all, I show that it cannot be irreversibility alone that deems a life *nonautonomous* even when such actions make it almost impossible for us to change our life plan in the future. Take having children. It seems counterintuitive if we were to say that all parents live nonautonomous lives (Self-regarding Duties: 188). Secondly, consider the case of a transsexual man (Self-regarding Duties: 188) who wants to change his body in an irreversible manner:

This man feels that he was born in the wrong body. The body he possesses is the physical body of a woman; yet, he identifies as a man. If we can rule out that this man’s wish to get the body of a man is involuntary for some reason, I guess many would argue that this man should be allowed to get the body of a man if it is what he autonomously desires. However, this includes operations that will change his body in an irreversible manner (e.g. he might give up his ability to bear children), prohibiting him from changing his mind about his life plan as a man.

If the transsexual man or a parent were to change their minds about their life plan, they would not be able to live accordingly—i.e., on my account, not live autonomously. However, in my view, ‘there are actions that are necessary if one is to live autonomously in spite of possible regrets in the future’ (Self-regarding Duties: 189), and it seems as too great a loss to a conception of autonomous living if we refrain from acting out of fear of possible future regrets. I grant that when making choices with irreversible consequences, we risk

¹⁷ This case is due to Clare Chambers (2004).

changing our minds in the future and thereby no longer be living autonomously. However, as I see it, that is a risk worth taking, because the alternative would mean not acting on our current autonomous life plan whenever it involves irreversible choices, and that is not living autonomously either. Furthermore, it seems as an integral part of the structural conception of autonomy that one takes one's future preferences into consideration, that is, in one's reflective and critical process one would need to consider possible future regrets (Self-regarding Duties: 187).

Autonomy-Based Self-Regarding Duties

In 'Self-regarding Duties', I use the distinction put forward here between different aspects of autonomy to show that even if we have autonomy-based self-regarding duties pertaining to what we should or should not do to our own body, these duties are only *prima facie* duties and therefore do not involve the sort of liberty-limiting restrictions that one may initially be inclined to think. I set aside the question whether self-regarding duties exist but simply assume that they do and show that *if they do exist*, they seem at first to propose a clear moral limit to what one may do to one's own body.

The paper focuses on autonomy-based self-regarding duties relating to what one should or should not do to one's own body. I argue that if we have certain self-regarding duties, we at least have the following three duties of this kind:

- 1) a duty to refrain from modifying our own body in ways that will diminish or destroy our capacities for autonomy
- 2) a duty to avoid limiting our array of options in such ways that it becomes inadequate
- 3) a duty to sometimes actively manipulate or modify our own body to enhance our capacities for autonomy and/or our array of options (Self-regarding Duties: 184-185).

Offhand, such self-regarding duties seem to limit how we ought to treat our body, as they indicate that we should not engage in activities that harm or risk harming our mental intelligence or ability to critically reflect upon our higher-order preferences (taking drugs, boxing etc.) (Self-regarding Duties: 182). Furthermore, we should not modify our body in ways that will limit our bodily function such as our ability to move or change our appearance in ways that will diminish our array of options in life (including career choices and partner options) in such a way that it will become inadequate (Self-regarding Duties: 182-183). Such duties seem immediately very restrictive in terms of what we should or should not do to our own body.

In the paper, however, I argue that such duties are only *prima facie* duties since they should be seen against the backdrop of (and sometimes in opposition to) the concern of autonomous living, and that a concern for living autonomously is more important. Consider the case of the Jehovah's Witness from 'Self-regarding Duties' (p. 191):

[the Jehovah's Witness] has been in an accident and needs a blood transfusion to survive. Because of his religion, however, which he identifies himself with on a second-order level, he believes that it is impermissible for him to accept a blood transfusion.

In this case, the Jehovah's Witness is acting in a *prima facie* morally wrongful way since he is violating his self-regarding duties to preserve his mental capacities and his array of options. However, taking the transfusion would be acting contrary to his autonomous life plan—contrary to who he wants to be.

I use this case (and others) to show that living according to one's autonomous life plan outweighs concerns for not violating autonomy-based self-regarding duties. There might of course be extreme cases where the scale would tip in the way of not violating one's self-regarding duties. As I state in 'Self-regarding Duties' (p. 191), if the Jehovah's Witness could gain 100 extra autonomous life years and there was a real chance that he would actually come to appreciate having gotten the transfusion, concerns for preserving his capacity for autonomy could in such extreme cases outweigh a concern for autonomous living. However, even here there is at least a presumption in favour of living autonomously *vis-à-vis* preserving one's capacities for autonomy. The way I see it, the fulfilment of such self-regarding duties may 'be seen as something from the outside that makes a person live a life that he or she does not endorse' (Self-regarding Duties: 191). I argue that living our lives according to our autonomous life plan is more important than fulfilling our self-regarding duties, when such prevent us from becoming the kind of person we want to be. Actually, I suggest that 'capacities for autonomy, as well as having an adequate array of options, are appropriately conceived as instruments toward securing autonomous living' (Self-regarding Duties: 191). Therefore, even if the autonomy-based self-regarding duties in question exist, they only impose *prima facie* restrictive moral limits to what one may do to one's body, since they are outweighed by a concern for living autonomously.

My argument in 'Self-regarding Duties' suggests that self-regarding duties are quite permissive. That is, people are not acting morally wrong even when they inflict harm on their body or modify it in any way that will have severe consequences for their capacity for autonomy and/or their array of options. However, it is important to note that my arguments are not permissive when it comes to body modifications that are not part of one's autonomous life plan

or based on trivial desires. To illustrate this point, I use the case of Leopard Leonard, who decides to get tattooed like a leopard in order to be able to participate and socialize with his co-workers at the annual theme party at his work place (Self-regarding Duties: 193). However, the leopard look will only cater to one of his desires (socializing with his co-workers), which is not even a very important desire for him. At the same time, it ‘comes at the expense of all other aspects of his autonomous life plan’ (career, friends, wife). The case of Leopard Leonard is supposed to illustrate a situation where a person violates his self-regarding duty to avoid limiting his array of options in such ways that it will become inadequate and jeopardize his chances of ever living autonomously. His choice to be tattooed as a leopard is morally problematic since his behaviour ‘is based on what amounts to shallow reasons relative to what constitutes autonomous living for him’ (Self-regarding Duties: 194), as opposed to people who change their body to be who they actually want to be.

Chapter 4: Anti-Paternalism, Paternalism and the Non-Harm Condition

In ‘Organ Donation’, I reject the non-harm condition, which says that vital organ donation is only ethically permissible when and only when no significant harm is done to the donor (harm being dying earlier than one would otherwise have died) (Organ Donation: 4). Standing firm on this condition seems objectionably paternalistic, since it prevents every donation involving any harm (or risk of harm) to the donor even when the donor consents to donating. Acting paternalistically means, roughly, interfering with a person’s autonomy or liberty to promote the good, interest or wellbeing of the person being paternalized (Dworkin 2020; Grill 2012). Upholding the non-harm condition seems to be an act of this kind, since it prevents people from becoming living vital organ donors if they are not imminently dying, that is, if they risk being harmed (by dying earlier than they would otherwise have died) by donating.

The charge against paternalism (represented by the non-harm condition) in ‘Organ Donation’ is that it fails to give considerable weight to autonomy. In this chapter, I specify how one can fail to give sufficient weight to autonomy in at least two ways: i) by balancing concerns for the harm of the donor against concerns for respecting the donor’s autonomous wish to donate, and in that balance giving too little weight to the concerns of autonomy, ii) by letting concerns for the harm of the donor count in any way as a weight on the scale (cf. Feinberg 1986: 26).

This chapter is structured as follows. First, I briefly present the paternalistic principle and its implications. Second, I lay out the opposing position, viz. anti-paternalism, which says that concerns for the wellbeing of the person in question should never be considered a good and relevant reason for interference. Lastly, I show how even if we accept the paternalistic principle, the non-harm condition still seems objectionable. Relating to the research question, this means that the non-harm condition does not impose any limits on what one may do to one’s own body.

Paternalism and Anti-Paternalism

So-called *hard* paternalism is the form of paternalism of relevance to my dissertation. Contrary to *soft* paternalism (which focuses on involuntary actions),¹⁸ hard paternalism focuses on voluntary actions (voluntary is here perceived as tantamount to acting autonomously¹⁹). That is, hard paternalism aims to promote the good, interest or wellbeing of the person even when the person is acting voluntarily (cf. Feinberg 1986: 12; Dworkin 2020; Begon 2016: 357). Paternalists believe that, sometimes, reasons to respect people's choices or their autonomy are outweighed by the reasons we have to promote the good of a person (cf. Feinberg 1986: 25-26; Pedersen & Midtgaard 2018: 773; Shafer-Landau 2005: 186-188). This leads to the paternalistic principle, formulated by Feinberg (1986: 4):

It is always a good and relevant (though not necessarily decisive) reason in support of a criminal prohibition that it will prevent harm (physical, psychological, or economic) to the actor himself.²⁰

In this way, paternalists accept that preventing harm is in fact a 'good and relevant' reason for not permitting living vital organ donation even when this means disrespecting the autonomy of the donor. However, the benefits of preventing harm to the donor must be weighed against the bad involved in disrespecting autonomy (cf. Grill 2015: 47; Pedersen & Midtgaard 2018: 773). This means that sometimes one would be justified in acting against the donor's autonomous wish, if the good promoted (in terms of wellbeing or harm prevented) substantially exceeds the bad involved in violating autonomy. In such

¹⁸ Most agree that soft paternalism is not paternalistic at all. Despite the name and the fact that it justifies interfering in people's self-regarding actions out of concern for the good of the person, it seems rather to be a moderate version of anti-paternalism (as opposed to hard anti-paternalism), since it only finds it to be a good and relevant reason to prevent people from harming themselves when and only when people are acting involuntarily (cf. Begon 2016: 356; Feinberg 1986: 12). For a recent critical discussion of soft paternalism, see e.g. Jason Hanna (2018).

¹⁹ Cf. Feinberg's standard of voluntariness (1986: 117-121).

²⁰ As Viki M.L. Pedersen & Søren F. Midtgaard (2018: 773) suggest, this formulation of the paternalistic principle seems too narrow. I agree, since it confines paternalism to cover only means of prohibition. However, there are many more ways to interfere with a person's autonomy than imposing prohibitions (prohibitions might not even interfere with a person's autonomy). To reflect interferences of both autonomy and liberty, Pedersen & Midtgaard (2018: 773) propose adding 'any other mean apart from rational persuasion' to the formulation, since that seems to cover a broader concern for autonomy and liberty.

situations, many paternalists would find interference justified (Dworkin 1988: 8, 127; Pedersen & Midtgaard, 2018: 775).

It is important to note that this kind of weighing could also result in *non-interference* when the good promoted does not substantially exceed the bad involved in violating autonomy. According to Daniel Groll (2012: 693), however, treating the will of the person merely as a weight on the scale disrespects the person's autonomy even when we actually end up not interfering or limiting the person's autonomy (i.e. permitting or not preventing some kind of putatively harmful and self-regarding act).²¹ Take the case of a patient who is told by his doctor that he needs surgery in order to survive. The patient is competent to make a decision and decides that he does not want the surgery (Groll, 2012: 692, 707). Suppose now that the doctor answers, 'I have decided that you should not have the surgery, because you do not want it' (2012: 707). How should the patient react? Most would agree that it is not the doctor's decision to make, but solely up to the patient to balance the relevant considerations for making such decision. What seems wrong in this situation is that the doctor considers the will of the patient 'as part of a larger set of considerations about what is good' for the patient (Groll 2012: 707), when the will of the patient should have been the only decisive consideration.²²

This line of thought indicates that to give sufficient weight to autonomy, we need to treat the will of the person as *structurally* decisive—as authoritative, not because it outweighs other considerations such as the good, interests or wellbeing of the person, but because it silences other considerations (Groll 2012: 701).²³ This is the anti-paternalistic position, since it opposes paternalism by holding 'that paternalistic reasons [should] never have *any* weight on the scales at all' (Feinberg 1986: 25-26; see also Grill 2015: 47; Feinberg 1984: 15). That is, it denies the paternalistic principle as this is set out above. According to the anti-paternalistic position, paternalistic reasons 'are morally illegitimate or invalid reasons by their very natures, since they conflict head on with defensible conceptions of personal autonomy' (Feinberg 1986: 26).²⁴

It is clear that autonomy is what motivates anti-paternalism, the idea that people should be allowed to pursue or not prevented from pursuing their *own*

²¹ Groll (2012: 707, 719) finds it disrespectful even when we treat the will of the person as *substantially decisive*—that is, as a weight on the scale that always outweighs all other reasons.

²² Groll (2012) refers to this as treating one person's will as structurally decisive.

²³ See also David Enoch (2016), who presents it as treating the will of the person as an exclusionary reason. See also Jessica Begon (2016: 269) on this.

²⁴ See also Kalle Grill (2015: 52), who describes this interpretation of the anti-paternalistic position as a filter approach.

preferences and projects. We should therefore not interfere with people's voluntary choices regardless of the consequences (cf. Begon 2016 357, 360; Flanagan 2017). In this way, autonomy seems to become the good that paternalism fails to respect by letting other considerations count as 'good and relevant' reasons for interference (Oshana 1998: 82).

The Non-Harm Condition

Irrespective of whether you embrace paternalism or anti-paternalism, respecting people's autonomous wishes seems at least to constitute a 'good and relevant' reason against interference when it comes to self-regarding choices. Even paternalists are willing to accept that sometimes the reasons in favour of respecting people's autonomous choices outweigh the benefits of preventing harm (Organ Donation: 6-7). In 'Organ Donation', I argue that accepting the non-harm condition would mean invoking a rather strong interpretation of the paternalistic principle—one I guess even most paternalists would be inclined to find objectionable (Organ Donation: 6-7).

The non-harm condition is suggested by proponents of what is called *organ donation euthanasia* (ODE). ODE is a proposal keen on expanding the pool of eligible donors for vital organ donation. In a nutshell, they suggest permitting vital organ donation from consenting patients who are imminently dying. Specifically, they suggest that 'when a person wants to have his or her life-sustaining treatment (LST) withdrawn, which most likely will cause the death of the person, it should actually be acceptable to procure the person's vital organs before death [with the patient's consents]' (Organ Donation: 2). According to ODE, vital organ donation is ethically permissible when and only when

1. the donor (or a surrogate) has voluntarily consented (the autonomy condition),
2. more viable organs will be available than if we wait until the donor is declared dead (the benefit condition), and
3. no significant harm is done to the donor (the non-harm condition) (Organ Donation: 4)

In the paper, I show that on a plausible understanding of the autonomy and the benefit condition, there is no compelling reason backing the requirement that the patient is imminently dying. A more far-reaching suggestion for regulation is warranted, that is, people who are not imminently dying should also be considered eligible as living vital organ donors. However, the non-harm condition blocks this suggestion, since, according to this condition, there must be no significant harm involved when retrieving the organs. According to the

non-harm condition, vital organ donation is only permissible in ‘those unfortunate situations in which individuals are very ill and imminently dying—either “naturally” because of their illness or by their own decision to withdraw LST [life-sustaining treatment]’ (Organ Donation: 3). In such cases, there seems to be no real harm involved in retrieving vital organs even when it leads to the death of the donor.

Yet upholding the non-harm condition seems objectionably paternalistic, I argue in the paper, since this would imply not giving any (or at least very little) weight to the reasons for respecting autonomy. Upholding the non-harm condition means, essentially, if the donor risks being harmed (even only moderately), they should not be permitted to donate (Organ Donation: 6).

Imagine the case from the paper of *Mr. 5-years-left*:

Mr. 5-years-left is told that because of illness, he has a maximum of 5 years left to live. There is nothing more the doctors can do to prolong his life, but they can make sure that he will not be suffering if he undergoes treatment every week for the rest of his life. He will be able to pursue some of his usual activities, but for every day that passes, his chances of becoming an organ donor decreases because his illness is wearing his organs down. This means that when he dies, his organs will not be suitable for transplantation. Mr. 5-years-left has a strong desire to help other people in need of an organ transplantation and, therefore, voluntarily chooses to become a living vital organ donor (Organ Donation: 6).

If we retain the non-harm condition, Mr. 5-years-left should not be permitted to donate his vital organs, even though there are strong autonomy-based reasons in favour of donation. Mr. 5-years-left would of course die five years earlier than he would otherwise have died, but not being allowed to donate would subject him to live according to a conception of the good life that he does not endorse. This way of limiting Mr. 5-years-left’s autonomy only prevents him from sacrificing five years of his life (and of course the opportunities that might have occurred in that period), which does not really seem to be a heavy weight to put on the scale in favour of not permitting the donation. However, either way this seems to be enough for the proponents of ODE to outweigh the reasons in favour of respecting autonomy.

Following the anti-paternalistic position set out above, giving this much weight to the (few) benefits of preventing harm is obviously and strongly objectionable, since such reasons should not even count as a weight on the scale at all. However, even in a revised (more permissible) form, the non-harm condition would always, from the perspective of an anti-paternalistic position, be objectionable because it implies that we are willing to consider harm as a ‘good and relevant’ reason for interference. So from the perspective of anti-paternalism, it seems clear that we should jettison the non-harm condition. Perhaps more interesting and surprising, though, *even from a plausible paternalistic*

perspective, it is unclear how we may stand firm on the non-harm condition and still proclaim to be giving considerable weight to the reasons in favour of respecting autonomy (Organ Donation: 6-7). Preventing harm may be a 'good and relevant' reason, but it should still be balanced against the reasons for respecting autonomy, and I guess 'most people (even paternalists) are willing to accept that people may make some important decisions regarding their own lives and take responsibility for instances of non-trivial harms that may be involved in such' (Organ Donation: 7). This means, as I end up claiming in the paper, that unless we are willing to accept an objectionable form of paternalism, the non-harm condition should at the very least be revised to permit some amount of harm, meaning that people who are not imminently dying should also be considered as eligible living vital organ donors (when the autonomy and benefit condition are satisfied) (Organ Donation: 7).

Chapter 5: What's Up with Personal Identity?

As it became clear in the previous chapter, the contemporary struggle between paternalists and anti-paternalists concerns whether reasons for promoting people's own good should ever be considered a 'good and relevant' reason for interference. One of the objections standardly raised against paternalism is that it disrespects people when we interfere in people's self-regarding matters. Such interferences inevitably seem to express that the people interfered with 'are incapable of making serious decisions regarding their own lives, which is why the state [or others] needs to circumscribe people's liberty in various ways' (Organ Donation: 6; see also Anderson 1999: 301-302).

This line of thought follows the foundational liberal premise that 'individual persons have an equal status as agents capable of making choices and planning their own lives, that such planning might well involve the intrapersonal distribution of resources between different stages of one's life, and that respecting people as agents with this capacity to plan involves according them freedom' (Carter 2018: 826). This goes hand in hand with Mill's harm principle that when it comes to people's voluntary self-regarding choices, others are not justified in interfering regardless of the consequences.

However, when anti-paternalists talk about the importance of respecting people's autonomous choices or their right to self-determination in relation to choices involving long-term and irreversible consequences, the underlying assumption seems to be that the person making the decision is also the one facing the consequences. As Richard Arneson describes the idea of making autonomous choices: 'making a voluntary choice a person takes on responsibility for all the foreseeable consequences to himself' (1980: 475). However, part of the reason we think autonomy or self-determination is important or valuable is that we rely on the above-mentioned assumption that the person making the self-regarding decision is also the one being harmed by it (harm being facing severe and irreversible consequences). But what if this is not the truth (cf. White 2020: 32)? What if we actually harm one or several other people when we make putatively self-regarding choices?

When we make choices with irreversible and severe consequences, it will have an impact on our future self—we are therefore making such choices on behalf of ourselves *and* our future selves or purely on behalf of our future selves if the consequences only appear in the future (cf. White 2020: 34). On

what has been named *the Simple View* of personal identity,²⁵ one where we take the person to be one and the same over time, it seems perfectly aligned with the foundational liberal premises that the person's future self, who is the still the same person, is the one facing the consequences (cf. Carter 2018: 826, 838; Parfit 1973: 140, 147; Kogan 1976: 835-836).

Recently, however, holding such a view of personal identity has been criticized, since we should not always see ourselves, it is argued, as being one and the same person throughout our biological life but instead acknowledge that our future selves might sometimes be numerically different from our past and current selves (Non-identity: 130). On this view of personal identity, presented as *the Complex View*, 'identity can, in its nature, be reduced to facts about psychological connectedness and continuity' (Non-identity: 134). According to the Complex View, if the psychological characteristics of one's past self and one's future self are sufficiently different, they would be considered as numerically different. In such cases, we would say that the past self has ceased to exist, and a new numerically different person (the future self) has come into existence.²⁶ This means that when we make choices with irreversible and severe consequences, we are not only making them on behalf of ourselves but also on behalf of one or several other numerically different people—viz. our future selves.

In this chapter, I present two arguments for why giving up the Simple View of personal identity may lead us to a justification of some moral limits on what one may do to one's own body. The first argument, *the shifting identity argument*, shows how putatively self-regarding actions can be subsumed under the harm principle and thereby justify restrictions on people's choices when they involve irreversible and severe consequences. The second argument or more precisely challenge, which is often referred to as *the personal identity problem*, challenges the moral authority of advance directives by implying that such directive should have nothing to say when its author has ceased to exist. By relying on the arguments from 'Non-identity' and 'Advance Directives', I show how neither the shifting identity argument nor the personal identity problem necessarily succeeds in justifying limits on what one may do to one's own body.

²⁵ Cf. Derek Parfit on this (1984: 325; 1973: 137-140).

²⁶ On this, see also John Kleinig 1983: 45-46; Terry S. Kogan 1976: 826; Parfit 1973: 137-141; 1984: 206; Allen Buchanan 1988: 294-295.

The Shifting Identity Argument

The shifting identity argument can be characterized as a ‘liberal strategy of reconciliation,’ since it aims at justifying restrictions on putatively self-regarding harm in a *non*-paternalistic way.²⁷ By appealing to the Complex View of personal identity, proponents of this strategy argue that because the future self should sometimes be seen as numerically different from the past self, then what at first might seem as an act of self-regarding harm in fact constitutes harm to other people (one’s numerically different future selves) (cf. Kogan 1976: 844; Reagan 1983: 125; Knight 2015: 129). This means that restrictions are justified by appealing to the harm principle²⁸ (cf. Reagan 1983: 122; Kleinig 1983: 45).

Take the famous case of the cyclist Bert,²⁹ who rides without a helmet. If Bert crashes and is injured, then according to Carl Knight, ‘it seems very likely that a significant psychological change will occur ... [in such a way that] post-accident Bert may be considered sufficiently different from pre-accident Bert that we do not assign the costs of latter’s choices to the former’ (2015: 129). Donald H. Reagan presents a similar version of the Bert case. He argues that ‘it is [of course] not certain that having the accident will produce any such change in the cyclist. But it seems likely to ... If the cyclist changes in this way, she is a [numerically] different person, who deserves protection against the foolish behavior of her earlier self’ (1983: 125). This means that when Bert chooses to ride without a helmet, Bert is in fact harming his future self who is a numerically different person. Accordingly, restricting Bert’s act of driving without a helmet would be justified by appealing to the harm done to his numerically different future self and in this way be perfectly in line with the harm principle.

The shifting identity argument thus proposes a moral limit to what one may do to oneself, since people are no longer morally permitted to act in ways where they risk harming their future selves (which includes severe and irreversible harm), because they are in fact not only harming themselves (Non-identity: 132-133). The shifting identity argument can be summarized like this:

Premise 1: If putatively self-harming acts at T1 resulting in putative self-harm at T2 are such that the agent at T1 is a numerically different person from the person who is harmed as a result of the acts at T2, then the harm principle

²⁷ Cf. de Marneffe 2006; Feinberg 1986; Pedersen 2019.

²⁸ Assuming of course that the harm principle expresses a sufficient condition of when such restrictions are justified.

²⁹ This case is originally due to Marc Fleurbaey (1995).

implies that interfering with the relevant putatively self-harming acts at T1 is permissible.

Premise 2: In a number of cases involving putatively self-harming acts at T1 resulting in putative self-harm at T2, these acts are such that the agent at T2 is harmed, and that agent is a numerically different person from the agent at T1.

Conclusion: In a number of cases involving putatively self-harming acts at T1 resulting in putative self-harm at T2, the harm principle implies that interfering with the relevant putatively self-harming acts at T1 is permissible (Non-identity: 135).

In this way, the shifting identity argument initially seems to justify limits on putatively self-harming behaviour in a non-paternalistic way by subsuming such behaviour under the harm principle. In the next section, however, I show how the shifting identity argument faces a profound challenge.

The Personal Non-Identity Thesis

The argument I put forward in ‘Non-identity’ shows that the implication of the shifting identity argument (that is, in relation to justifying restrictions on putatively self-harming behaviour) are very limited. I argue that even if we accept the Complex View of personal identity, meaning that the current self existing at T1 is in fact numerically different from the future self existing at T2, the current self is not harming anyone when acting putatively self-harming (Non-identity: 135). I do so by appealing to the non-identity problem, which is usually about what we owe future spatially different people or generations, and argue that it makes sense to talk about the non-identity problem in a personal context to see what we owe our numerically different future selves.

In a nutshell, the standard view of the non-identity problem is described as the situation when a person acts in some way at T1, which will have some consequences for the future person existing at T2. The person is not harming anyone presumably since the person at T2 is not worse off than he or she would otherwise have been, because he or she only exists because of the act done at T1. If the person at T1 had done something different or not done anything, the person at T2 would have been a numerically different person, since the act done at T1 determines the identity of the future person facing the consequences at T2 (cf. Boonin 2014: 2-4). Take the famous case of Wilma (due to Boonin 2014: 2-5):³⁰

³⁰ This is similar to what is often referred to as the case of the 14-year-old girl (cf. Parfit 1984: 358).

Wilma wants to become pregnant. Her doctor tells her that if she conceives now, she will have a blind child. Blindness would inevitably affect the child's quality of life negatively, however, it would not be so that the life of the child would not be worth living (people live fine lives despite being blind and despite the fact that their lives would probably have been better had they not been blind). However, Wilma can take a pill for the next two months and wait to conceive, and then she will have another child that would not be blind. Wilma finds it a bit inconvenient to take a pill for the next two months and decides to conceive now. She gives birth to the blind child (Non-identity: 136, note excluded).

In this case, Wilma is not harming anyone, because she is not making anyone worse off than they would otherwise have been (assuming that an act only harms another person if it makes this other person worse off than they would otherwise have been).

I believe that when it comes to cases of identity shift, we are facing a similar situation, what I call a personal non-identity situation. However, for this analogy to work, it is important that the identity of the future self existing at T2 depends on the putatively self-harming act done by the agent at T1. That is, we need to make sure that the future selves in question are not *actual* future selves, i.e. future selves who will exist no matter what we do,³¹ but merely possible future selves (Non-identity: 136-138). However, as I point out in 'Non-identity', it seems plausible to assume that in most cases of identity shift, the identity of the future person coming into existence would either be determined by the putatively self-harming act, or the act in question would be causing the identity shift to happen (cf. Non-identity: 138; 141). Of course, there might be cases where the identity of the future self is neither determined nor caused by the putatively self-harming act. However, most cases that are relevant for the shifting identity argument are cases of such severe and irreversible damages, for example driving (and crashing) without a helmet, that it seems plausible to assume that such cases would most often cause a shift in identity. As Knight (2015: 130) puts it:

It may be that the persons that result from the traumatic events typically described in abandonment cases rarely, if ever, stand fully in relation that matters to their earlier selves.

Now consider the version of the Bert case which I use in the paper. I call him Imprudent Bert and assume that when he crashes without wearing a helmet, sufficient psychological changes will occur in such a way that the future Bert, whom I call Prudent Bert, will be a numerically different person (Non-identity: 139). Had Imprudent Bert worn a helmet, Prudent Bert would never have

³¹ This would be the case if we inevitably shift identity over time, and the identity of the future self does not depend on the choices we make (cf. Non-identity: 136-137).

come into existence, since his existence depends on Imprudent Bert driving (and crashing) without a helmet. This warrants the construction of an analogy to the Wilma case and shows that because Prudent Bert only exists because of the choice Imprudent Bert made to drive without a helmet, Prudent Bert is not worse off (assuming that his life is still worth living) than he would otherwise have been, since he would otherwise never have existed. The analogy argument I put forward can be summarized like this:

Premise 1: Bert's act of driving without a helmet rather than wearing a helmet does not make prudent Bert worse off than he would otherwise have been.

Premise 2: If a's act harms b, then a's act makes b worse off than b would otherwise have been.

Conclusion 1: Bert's act of driving without a helmet rather than wearing a helmet does not harm prudent Bert.

Premise 3: Bert's act of driving without a helmet rather than wearing a helmet does not harm anyone other than prudent Bert.

Conclusion 2: Bert's act of driving without a helmet does not harm anyone (Non-identity: 140).

Since Bert is not harming anyone, the shifting identity argument fails to subsume this kind of behaviour under the harm principle. This means that even though the shifting identity argument at first seems appealing, since it tries to justify restrictions on putatively self-regarding behaviour involving severe and irreversible consequences in a non-paternalistic manner, the argument only holds in a very limited number of cases (i.e. cases where the identity of the future self is neither determined nor caused by the putatively self-harming act).

The argument I put forward in 'Non-identity' thereby shows that accepting the Complex View of personal identity inevitably affects how we should think about self-regarding harm, since when choosing to act in ways that will have severe and irreversible consequences, we are not just choosing on behalf of ourselves but also on behalf of our possible future self. However, as I show in this paper, this does not warrant any restrictions on such behaviour, since we are not harming anyone else.³²

³² There might of course be impersonal reasons in favour of restricting this kind of behaviour.

Advance Directives and the Personal Identity Problem

An advance directive is a legal document in which you state your decision regarding future treatments and medical care in case you become incompetent. In this way, you exercise your right to medical self-determination beforehand, since the advance directive will make sure that decisions made regarding you in case you become incompetent reflect your values and respect your wishes.³³ This means that by upholding the advance directive, others respect your right to medical self-determination (at least when the advance directive is deemed valid³⁴) (cf. Gligorov & Vitrano 2011: 149; White 2020: 39).

The moral authority of advance directives is based on an assumption of the Simple View of personal identity, i.e. that there is sufficient psychological connectedness between the author of the directive and the patient to whom the directive should apply (Witt 2020: 52; White 2020: 31). Accepting the Complex View of personal identity might therefore undercut the moral authority of advance directives if we have a reason to question whether we are dealing with numerically identical selves. This challenge to the moral authority of advance directives, often referred to as *the personal identity problem*, is what I address in ‘Advance Directives’.

The personal identity problem—qua its relation to the notion of personal identity pertaining to the Complex View of personal identity—questions whether numerical identity obtains in cases where advance directives usually come into play. More precisely, the problem is that when facing advanced cognitive deterioration (which renders one incompetent and brings the advance directive into play), one ceases to exist, since the future self emerging because of the deterioration will not be psychologically continuous with the former self (Advance Directives: 2; see also Buchanan 1988: 280-281; Witt 2019: 3-5; 2020: 53). As Karsten Witt puts it: ‘The process [of becoming severely demented] is literally fatal to us. We cannot become severely demented because we cease to exist before that happens’ (2019: 4). This means that the incompetent person emerging because of the deterioration will be numerically different from the person authoring the advance directive. Therefore, upholding the advance directive can no longer be seen as an act of respecting the author’s right to self-determination, since the author no longer exists.

³³ Or at least the values and wishes you had at the time you authored the advance directive.

³⁴ In terms of the author of the directive being competent enough to make decisions regarding oneself and acting voluntarily.

In this way, the personal identity problem imposes a moral limit on one's decision regarding what should happen to one's own body in the future (e.g. in terms of medical treatment and care). If we accepted the Simple View of personal identity, the standard view in medical ethics seems to be that we should respect people's self-regarding choices (when voluntary) even when they are stated beforehand in an advance directive (cf. White 2020: 31, 39; Dworkin 1993: 222-229). However, as was the case in the former section, when we accept the Complex View of personal identity, we are suddenly no longer only choosing for ourselves but also on behalf of our possible future selves, who might be numerically different persons, and in the case of advance directives, we are choosing solely on behalf of our future (perhaps numerically different) self (White 2020: 34).

The Trilemma

In 'Advance Directives', however, I argue that advance directives might actually keep their moral authority even when we accept the Complex View of personal identity. I appeal to an argument of surviving interests in which advance directives should be seen as a tool for protecting such interests. As Allen Buchanan has it, '[a] person who issues an advance directive may do so not only to exercise control over what happens to herself after she becomes incompetent, but also to protect certain interests she has in what happens to her body after she, the particular person who she is, no longer exists' (1988: 286). These are what Buchanan calls surviving interests because 'whether they are satisfied or thwarted depends upon events that occur or do not occur after the person no longer exists' (1988: 287). This is in line with Feinberg's argument that a person's interests continue to exist independently of his or her death: 'we can think of some of a person's interests as surviving his death, just as some of the debts and claims of his estate do, and that in virtue of the defeat of these interests, either by death itself or by subsequent events, we can think of the person who was, as harmed' (1984: 83). Furthermore, only the interests that can no longer be helped or harmed by posthumous events die with the person, and the interests of what happens to one's body after death are not of such kind (1984: 86).

I argue that such interests are equivalent to the surviving interests one might have in what happens to one's organs after one dies, and therefore if the surviving interests always outweigh the interest of any living numerically different person in cases of organ donation (as they standardly seem to do³⁵), the surviving interests of the past selves should also always outweigh the interests

³⁵ Cf. T. M. Wilkinson 2011: 42-46.

of any living numerically different future self (Advance Directives). Accordingly, I argue that we are facing a trilemma with the following prongs:

- I) Regarding the moral authority of advance directives, we should never prioritize the surviving interests of the former self when the future self is a numerically different person.
- II) Regarding the moral authority of people's interests in what happens to their organs when they die, we should always give priority to the surviving interests of the holder of the organs, even when a numerically different person needs an organ transplant to survive.
- III) Since the cases in question in I) and II) are equivalent, such cases should be treated in the same way (Advance Directives: 3, 11).

In 'Advance Directives', I remain neutral on how to solve the trilemma. Yet, I show that we should accept III, as there seems to be no morally relevant difference between the cases relevant in I and the cases relevant in II. I do so by appealing to the following two cases:

No life-supporting treatment. Margo signs an advance directive saying that if she becomes severely demented, all life-support efforts are to be withdrawn. The cognitive deterioration due to the dementia has now become so severe that she has ceased to exist. The patient that remains, however, still possesses enough mental capacities to be considered a person and is therefore a new, numerically different person. Let us call her Older-Margo. Despite her illness, Older-Margo is described as undeniably happy—it is as if her illness leaves her carefree and always cheerful. Older-Margo has now contracted pneumonia and needs antibiotics to survive. She has an interest in staying alive (Advance Directives: 5, note excluded).

No organ donation. Amy signs an advance directive saying that if she in any way ceases to exist, she does not want to donate her organs. The cognitive deterioration due to the dementia has now become so severe that she has ceased to exist. The patient that remains is no longer a person and rapidly becomes brain dead. Carol, who is a numerically different third person, needs new organs to survive, and we know that the organs of Amy are a perfect match for Carol. However, if we uphold the advance directive, the organs should not be retrieved. If Carol does not receive the organs from Amy, she will die (Advance Directives: 6).

The two cases present us with two sets of *surviving* interests and two sets of *living* interests. Comparing the two surviving interests—i.e. Margo's interest in deciding what happens to her body after she ceases to exist and Amy's interest in deciding what should happen to her organs after she dies—I argue that there does not seem to be any morally relevant difference. As I state in the paper, Margo might 'find it degrading that her living remains, even though

they are not her, should be dependent on others ... [and] she might be concerned about what this numerically different person with dementia would do with and to her body and how it would affect her legacy' (Advance Directives: 7-8). Similarly, Amy might have interests in what happens to her body when she no longer exists; she might find it 'important to be buried with her remains intact' (Advance Directives: 8). Furthermore, I argue that the interests of the living numerically different people in the two cases (Older-Margo and Carol) are not morally relevantly different since they both need the body (or an organ) to survive. Both seem to have strong welfare-based interests in the body (or parts of it) (Advance Directives: 8).

As I point out in the paper, some might argue that an obvious disanalogy between the two cases will be that in the case of *No life-supporting treatment*, the future numerically different person already occupies the body, and by occupying the body the future self in some way acquires right over the body, which Carol in *No organ donation* does not seem to have (Advance Directives: 8-10).³⁶ However, consider the case from 'Advance Directives', which I call *Forest*:

Let us imagine that I have an ownership right to decide what happens to my property after I die. I own a beautiful forest and it is very important to me that the natural circle of life is not interrupted. I do not want my forest ever to be exposed to chemicals, even if it means that the forest eventually will be eaten by bugs. I state in my will that you can use the forest as much as you like, but in the event of an attack from bugs, you should not use chemicals to protect the forest. This will most likely make the forest decay. If you moved into my forest and lived there for many years, you may get some minimal ownership rights over the forest. You may decide who is allowed to visit you in the forest and so on, but should you have anything to say when the forest is attacked by bugs? (Advance Directives: 10).

I argue that when we only take ownership rights into account, then you should have nothing to say, and if you are in any way justified to go against my stated will (you might need the forest to have a home), it is not because you have acquired ownership rights over the forest but because you have a claim of necessity (Advance Directives: 10-11). This analysis leads us to accept III, which means that we need to give up either I or II in order to hold consistent views. As mentioned in the paper, I remain neutral on whether we should reject I or II. However, if we are not willing to accept that advance directives should have moral authority when the future self is a numerically different person, we need

³⁶ In 'Advance Directives' I also show that occupation in itself does not seem to be morally relevant.

to reject the idea that people have a right to decide what should happen to their organs when they die.

In relation to the research question, the personal identity problem suggests that there might be a moral limit to what one may decide about what happens to one's own body when one ceases to exist, and I cannot deny that this might actually be the case when we accept the Complex View of personal identity. However, if the personal identity problem does constitute such moral limit, it implies that we are willing to reject our basic intuitions regarding people's right to decide what happens to their organs after they die, even when other people need them to survive (Advance Directives: 11-12).

Chapter 6: Conclusion

In response to the research question, this dissertation argues that there are few (if any) substantial moral limits to what one may do to one's own body. I defend this claim by showing that a concern for living autonomously (i.e. living according to one's autonomous life plan) outweighs a concern for not violating our autonomy-based self-regarding duties, and that there are no content conditions pertaining to what amounts to an autonomous life. Accordingly, I argue that as long as we satisfy the procedural conditions for a life plan to be considered one's own (our autonomous life plan being the result of an independent and critical reflection process), we are not acting morally wrong when inflicting harm on our body since we are living exactly the lives we autonomously want. However, if the acts in question do not reflect who we want to be, we are in fact acting morally wrong when we inflict certain harms on our body in ways where we risk diminishing or destroying our capacity for autonomy and/or our adequate array of options. That is, if our behaviour is based on trivial desires and jeopardizes our chances of living autonomously, then—since the harm in question is not part of who we want to be—the concern not to violate our autonomy-based self-regarding duties cannot be outweighed by the concern for living autonomously. This adds to the set of procedural conditions that the choices we make need to be the result of a deliberative process about which preference coheres best with who we want to be.

This suggests a very permissive view. Yet there are cases which challenge our liberal commitment to the harm principle or the claim that only harm to others is morally problematic. There might be situations in which the implications of respecting people's voluntary choices are hard to accept even when they live according to their autonomous life plan and only risk harming themselves. Hence it is relevant to consider whether certain liberty-limiting principles may be justifiable. This dissertation discusses some of the most prominent liberty-limiting principles that have been suggested in the literature seeking to justify restrictions on self-harming behaviour.

The non-harm condition, being one of these principles, implies that significant and non-trivial harm should count as relevant reasons for restricting self-harming behaviour. However, as I argue in 'Organ Donation', this condition can only be retained at the price of accepting an objectionable form of paternalism. Some might of course argue that in extreme situations involving, for example, a young person who faces some adversity and wants to end their

life, the harm in question might count as a relevant reason for considering restricting this kind of behavior (cf. Organ Donation: 6-7). However, I show that even if this is true, there is still a wide range of cases where we should be free to act in ways that may bring harm to us in various and non-trivial ways.

Yet there is another way for liberals to try to justify restrictions on self-harming behaviour. If it is possible to show that the putatively self-harming behaviour in question is in fact a sort of *other*-regarding harm, such behaviour can be subsumed under the harm principle, and restrictions would be justified. By appealing to a reductionist view of personal identity, the shifting identity argument seeks to show that when we act in ways where we risk harming ourselves in a severe and irreversible way, we are in fact harming our future self, who, appropriately conceived, is a numerically different person. Despite the ingenuity and attractiveness of this argument, I show that it is unsuccessful. By appealing to the debate about the non-identity problem, I argue that the past self is in fact *not* harming anyone when acting putatively self-harming, since the future self only exists because of the choice the past self made to act putatively self-harming. There might of course be cases where the past self would have come into existence independently of the act in question. If this is the case, the past self is in fact harming the numerically different future self, and the shifting identity argument would justify restrictions on putatively self-harming behaviour in such cases. However, as I argue in ‘Non-identity’, it seems plausible to assume that most (if not all) cases relevant for the shifting identity argument would involve such severe and irreversible damages that it would cause an identity shift and/or determine the identity of the future self. The rebuttal of the subsuming argument indirectly supports the positive claim that there are no moral limits to what one may do to one’s own body.

Finally, this dissertation contributes to answering the research question by considering a liberty-limiting principle which seeks to show that our right to medical self-determination when expressed in an advance directive is limited. By appealing to a reductionist view of personal identity, the so-called personal identity problem challenges the moral authority of advance directives by implying that when we face advanced cognitive deterioration, such as severe dementia, we actually cease to exist. This means that in cases where advance directives usually come into play, the person who authored the advance directive no longer exists, so why should we uphold the advance directive? However, as I show in this dissertation, under realistic assumptions, advance directives might still have moral authority because they should be seen as an expression of the former person’s surviving interests about what happens to their body after they no longer exist. In itself, the argument that advance directives are an expression of surviving interests does not give us a reason to uphold the advance directive. However, as I argue in this dissertation, if we

are not willing to give any weight to the surviving interests in cases of advance directives, we should not give any weight to the surviving interests people have in what happens to their organs when they die either. I remain neutral on whether this is the way to go, or whether we need to uphold advance directives because we are not willing to give up the idea about people having a right to decide what happens to their organs when they die. Accordingly, this dissertation does not give a clear answer to whether the personal identity problem actually succeeds in justifying a limit on the right to medical self-determination. In response to the research question, this means that there might be a moral limit to what one may do to one's own body, or more precisely a limit to what one may direct about what should happen to one's body in the future when one no longer exists. However, this limit hinges upon a willingness to accept that people do not have a right to direct what should happen to their organs (or their body) when they die.

Although I have examined key issues relating to the limits of what one may permissibly do to one's own body, my investigation is plausibly not exhaustive. I have chosen to focus on the limits connected to person-affecting considerations such as personal wellbeing and personal autonomy because they are usually taken to have more force than considerations for certain impersonal values (cf. Goold & Savulescu 2008: 56). Yet some might argue that we should not inflict harm on our own body or modify it in any way because of some impersonal value. By appealing to autonomy as intrinsically valuable, one argument might be that we fail to treat our autonomy as an end in itself when we choose a life as slave or modify our body in any way that leaves our array of options inadequate (cf. Bou-Habib 2006; Midtgaard 2015; *Organ Donation*: 7). We could also imagine an argument appealing to a duty to bring about valuable state of affairs (cf. Kymlicka 2002: 33; Hurley 2006: 332). Yet, while these factors are clearly valid, it is still far from clear that a concern for the intrinsic value of autonomy or our duty to bring about a valuable state of affairs is sufficiently strong to outweigh the person-affecting oriented concern for living autonomously.³⁷ Furthermore, it is open to serious doubt that others would necessarily be justified in restricting self-regarding behavior on the grounds of these impersonal values, even when we could say that one was acting morally wrong by inflicting harm on one's own body.

³⁷ I make a similar point in 'Organ Donation' (p. 7).

English Summary

This dissertation argues that as long as we live according to our autonomous life plan and satisfy relevant procedural conditions, there are few (if any) substantial moral limits to what one may do to one's own body, since we are living exactly the lives we autonomously want.

The first part of this dissertation shows that a concern for living autonomously outweighs any concern for not violating our autonomy-based self-regarding duties to not (risk) diminish(ing) or destroy(ing) our capacity for autonomy and/or our adequate array of options. Accordingly, if such self-regarding duties exist, they only impose *prima facie* moral limits to what one may do to one's own body. This means that we are not acting morally wrong when we inflict harm on our body as long as it is part of who we want to be (i.e. an integral part of our autonomous life plan).

This supports the thesis that there are few (if any) substantial moral limits to what one may do to one's own body, which suggests a very permissive view. Yet there are cases where people inflict very severe and irreversible harm on themselves in a way that challenges our liberal commitment to such a permissive view. The second part of this dissertation therefore explores whether significant and non-trivial harm should count as a relevant reason for restricting self-harming behaviour. By focusing on the case of living vital organ donation, this dissertation shows that even when people risk harming themselves in a significant and non-trivial manner (in this case sacrificing their life to donate their vital organs), the only way to justify restricting this kind of behaviour is to accept an objectionable form of paternalism.

The third part of the dissertation explores whether restrictions on self-harming behaviour can be justified in a non-paternalistic way by applying a liberal strategy of reconciliation. One such strategy is to show that we should actually be reductionist about personal identity in such a way that it is no longer certain whether our past, present and future selves are one and the same person. This means that when we make choices with severe and irreversible consequences, we are not just making them for ourselves but also on behalf of our future potentially numerically different self. The dissertation focuses on two arguments, the shifting identity argument and the personal identity problem, which both seek to justify restrictions on putatively self-regarding behaviour by subscribing to a reductionist view of personal identity.

The shifting identity argument seeks to justify restrictions on putatively self-harming behaviour by subsuming these under the harm principle. More precisely, the shifting identity argument shows how putatively self-harming behaviour is in fact *other*-regarding harm. That is, when we engage in self-

harm, we are harming our numerically different future self, and by appealing to the harm principle, we are justified in restricting behaviour involving other-regarding harm. However, by drawing on insights from the debate about the non-identity problem, the dissertation shows that the shifting identity argument cannot withstand scrutiny. I argue that when we accept that our future self might be a numerically different person, we are facing a personal non-identity situation in the cases in question, since the identity of this future self will inevitably depend on the choices we make now. This means that when I choose to drive without a helmet and crash, the person (who is a numerically different person) emerging because of the crash only exists because I chose to act the way I did. Had I not been driving (and crashed) without a helmet, the person would never have come into existence. This means that I do not harm my future self, even when this person is numerically different from me, since she would not be worse off than she would otherwise have been, assuming that her life is still worth living. Accordingly, we are not justified in restricting putatively self-harming behaviour.

The personal identity problem challenges the moral authority of advance directives. By appealing to a reductionist view on personal identity, the personal identity problem implies that when we face advanced cognitive deterioration, such as severe dementia, we cease to exist. This means that the cases where advance directives usually come into play, the person who authored the advance directive no longer exists, so why should we uphold the advance directive? The advance directive seems to have no moral authority over what should happen to the patient with whom we are now dealing, since this patient is a numerically different person. By appealing to an argument for the existence of surviving interests, the dissertation shows that advance directives need not lose their moral authority even when we accept that the author has ceased to exist. This means, I argue, that if it is important for one what happens to one's living remains, such interests might actually outweigh the numerically different future self's potential interests. To some, it may seem counter-intuitive that we would let the surviving interests of a former person outweigh the interests of a living numerically different person; however, this is exactly what we do in the case of organ donation. If people state that they do not want their organs to be donated when they die, it seems commonly accepted that we should follow their wishes, even when other living people might have an interest in receiving one or more of the organs. I argue that in order to be consistent, we either need to abandon the idea that people have a right to decide whether they want to donate their organs when they die, or simply accept that advance directives still have moral authority even when the author has ceased to exist.

This dissertation hereby shows that some of the most prominent liberty-limiting principles in the literature do not necessarily succeed in justifying restrictions on self-regarding behaviour, which indirectly supports the thesis that there are few (if any) substantial moral limits to what one may do to one's own body.

Dansk resumé

Denne afhandling viser, at så længe vi lever efter vores autonome livsplan og tilfredsstillende relevante procedurale kriterier, lever vi præcis det liv, vi autonomt ønsker at leve, og der er derfor kun få (om nogen) substantielle moralske grænser for, hvad man må gøre ved sin egen krop.

Den første del af afhandlingen viser, at hensynet til at leve autonomt vejer tungere end hensynet til at overholde vores autonomibaserede pligter over for os selv til ikke at risikere at begrænse eller ødelægge vores autonomikapaciteter og/eller vores tilstrækkelige række af muligheder. Dette betyder, at hvis sådanne pligter faktisk eksisterer, begrænser de kun prima facie, hvad vi må gøre ved vores egen krop. Vi handler derfor ikke moralsk forkert, når vi skader vores egen krop, så længe skaden er i overensstemmelse med det liv, vi gerne vil leve (dvs. en integreret del af vores autonome livsplan).

Dette støtter op om tesen, at der kun er få (om nogen) substantielle moralske grænser for, hvad man må gøre ved sin egen krop. Dette liberale synspunkt kan dog blive udfordret, når personer frivilligt påfører sig selv alvorlige og irreversible skader. Anden del af afhandlingen undersøger derfor, hvorvidt signifikant og ikke-trivielt skade udgør en god og relevant grund til at begrænse selvskadende adfærd. Med afsæt i spørgsmålet hvorvidt levende donation af vitale organer er etisk forsvarligt, viser afhandlingen, at den eneste måde, hvorpå vi kan retfærdiggøre begrænsninger på handlinger, hvor folk risikerer at skade sig selv på en signifikant og ikke-trivielt måde (i denne situation dø før tid), er at acceptere en problematisk form for paternalisme.

Den tredje del af afhandlingen undersøger, hvorvidt vi kan retfærdiggøre begrænsninger på handlinger, der kun vedrører en selv, ved at anvende en liberal forsoningsstrategi, hvor vi ikke appellerer til paternalistiske argumenter. En måde at gøre dette på er at vise, at vi skal acceptere et reduktionistisk syn på personlig identitet, hvor vi ikke længere nødvendigvis skal se vores fortidige, nutidige og fremtidige selv som en og samme person. Dette betyder, at når vi træffer beslutninger, træffer vi dem ikke kun på vegne af os selv, men også på vegne af vores fremtidige numerisk forskellige selv. I afhandlingen fokuserer jeg på to argumenter, der begge forsøger at retfærdiggøre grænser for, hvad vi må gøre ved vores egen krop, ved at appellere til denne forståelse af personlig identitet.

Det første argument, the shifting identity argument, forsøger at retfærdiggøre begrænsninger på den skade, man må påføre sig selv, ved at vise, at når vi risikerer at skade os selv alvorligt og irreversibelt, så skader vi faktisk en anden person, nemlig vores fremtidige numerisk forskellige selv. På denne

måde kan vi retfærdiggøre sådanne begrænsninger ved at appellere til skadesprincippet, som netop siger, at begrænsninger er retfærdiggjorte, når vi risikerer at skade andre. Med baggrund i debatten omkring the non-identity problem viser jeg dog, at vi bør afvise the shifting identity argument. Jeg argumenterer for, at i situationer hvor vi vil betragte vores fremtidige selv som en numerisk anden person, så vil vi stå over for et personligt non-identity problem. Dette skyldes, at vores valg uundgåeligt vil være styrende for, hvilken identitet vores fremtidige selv vil have. Når vi fx vælger at køre uden styrthjelm, og vi styrter, så vil den person, der vågner op efter styrtet, potentielt skulle betragtes som en numerisk anden person. Vedkommende eksisterer dog kun, fordi vi netop valgte at køre uden styrthjelm (og styrtede) til at begynde med. Havde vi ikke valgt at køre uden styrthjelm, ville denne person aldrig have eksisteret. Vi har altså ikke skadet vores fremtidige selv, selv når vi kører uden styrthjelm, da personen ikke er værre stillet, end personen ellers ville have været (da personen i så fald ikke ville have eksisteret). Dette betyder altså, at the shifting identity argument ikke kan retfærdiggøre begrænsninger på antageligt selvskadende adfærd.

Det andet argument er nærmere beskrevet som et problem, kaldet the personal identity problem, der udfordrer behandlingstestamenters moralske autoritet. Ved at appellere til et reduktionistisk syn på personlig identitet peger the personal identity problem på, at når vi oplever alvorlig kognitiv funktionsnedsættelse, fx svær demens, så ophører vi faktisk med at eksistere. Dette betyder, at i de situationer, hvor vi normalt vil tage et behandlingstestamente i brug (fx når folk er svært demente), vil forfatteren af behandlingstestamentet ikke længere eksistere og står i stedet over for et valg vedrørende en numerisk anden person. Hermed har vi ikke længere nogen grund til at følge anvisningerne i behandlingstestamentet, og det kan derfor ikke være styrende i forhold til beslutninger om, hvilken behandling vi skal give (eller ikke give) patienten. Afhandlingen viser dog, at behandlingstestamenter ikke nødvendigvis mister deres moralske autoritet, selv når forfatteren ikke længere eksisterer. Med baggrund i et argument som viser, at behandlingstestamenter faktisk skal ses som et udtryk for folks overlevende interesser, argumenterer jeg for, at hvis det er vigtigt for en person, hvad der sker med vedkommendes krop, når han eller hun ikke længere eksisterer, bør sådanne interesser måske veje tungere end de interesser, det fremtidige selv måtte have. Nogle vil måske finde det kontraintuitivt, at folks overlevende interesser omkring, hvad der skal ske med deres krop, når de ikke længere eksisterer, kan veje tungere end de interesser, en levende person måtte have i den krop. Dette er dog præcis, hvad vi synes at være villige til at gøre, hvis folk ikke ønsker at donere deres organer, når de dør. Det synes at være en udbredt antagelse, at hvis en person ikke ønsker at donere, så skal vi respektere dette ønske, uanset om der er en levende

person, som kunne have interesse i at modtage den afdødes organer. Ved at appellere til kravet om konsistens peger jeg på, at vi er nødt til enten at opgive idéen om, at vi har ret til at bestemme, hvad der skal ske med vores organer, når vi dør, eller acceptere, at behandlingstestamenter stadig har moralsk autoritet, selv når forfatteren ikke længere eksisterer.

Samlet set viser afhandlingen, at nogle af de mest fremtrædende frihedsbegrænsende principper fra litteraturen ikke nødvendigvis kan begrunde moralske grænser for, hvad vi må gøre ved vores egen krop. Disse resultater støtter dermed indirekte op om tesen, at der kun er få (om nogen) substantielle grænser for, hvad man må gøre ved sin egen krop.

Bibliography

- Andersen, Didde B. (2020) Self-regarding Duties, Body Modifications and Autonomous Living. *Ethical Perspectives* 27(2): 177-199
- Andersen, Didde B. (2021) I Have Got a Personal Non-identity Problem: On What We Owe Our Future Selves, *Res Publica*, 27: 129-144
- Andersen, Didde B. (2021) May I give my heart away? On the permissibility of living vital organ donation. *Bioethics*, DOI: 10.1111/bioe.12935
- Andersen, Didde B. On the Moral Authority of Advance Directives. *Unpublished manuscript*
- Andersen, Didde B. & Midtgaard, Søren F. (2017) Stay Out of the Sunbed! Paternalistic Reasons for Restricting the Use of Sunbeds. *Public Health Ethics* 10(3): 276-288
- Anderson, Elizabeth S. (1999) What is the Point of Equality? *Ethics* 109(2): 287–337.
- Arneson, Richard (1980) Mill *versus* Paternalism. *Ethics* 90(4): 470-489
- Arneson, Richard (2005). Joel Feinberg and the Justification of Hard Paternalism. *Legal Theory* 11: 259-284
- Begon, Jessica (2016) Paternalism. *Analysis* 76(3): 355-373
- Blake, Michael (2002) Distributive Justice, State Coercion, and Autonomy. *Philosophy and Public Affairs* 30(3): 257-296
- Boonin, David (2014) *The Non-Identity Problem and the Ethics of Future People*. Oxford: Oxford University Press
- Bou-Habib, Paul (2006) Compulsory Insurance without Paternalism, *Utilitas* 18(3): 243-263
- Brownlee, Kimberly & Stemplowska, Zofia (2017) “Thought Experiments” in *Methods in Analytical Political Theory*, A. Blau (ed.), Cambridge: Cambridge University Press pp. 21-45
- Buchanan, Allen (1988) Advance Directives and the Personal Identity Problem. *Philosophy and Public Affairs* 17(4): 277-302
- Carter, Ian (2018) Equal Opportunity, Responsibility, and Personal Identity. *Ethical Theory and Moral Practice* 21: 825–839
- Christman, John (2020) Autonomy in Moral and Political Philosophy, *The Stanford Encyclopedia of Philosophy* (Fall 2020 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/fall2020/entries/autonomy-moral/>
- Chambers, Clare (2004) Are breast implants better than female genital mutilation? Autonomy, gender equality and nussbaum’s political liberalism. *Critical Review of International Social and Political Philosophy* 7(3): 1-33
- Cohen, G. A. (2003) Facts and Principles. *Philosophy and Public Affairs* 31(3): 211-245
- Daniels, Norman (1979) Wide Reflective Equilibrium and Theory Acceptance in Ethics, *The Journal of Philosophy* 76(5): 256-282
- Daniels, Norman (2020) Reflective Equilibrium, *The Stanford Encyclopedia of Philosophy* (Summer 2020 Edition), Edward N. Zalta (ed.), URL

- =<https://plato.stanford.edu/archives/sum2020/entries/reflective-equilibrium/>
- de Marneffe, Peter (2006) Avoiding Paternalism. *Philosophy & Public Affairs* 34(1): 68–94.
- Dworkin, Gerald (1988) *The Theory and Practice of Autonomy*, Cambridge University Press, Cambridge
- Dworkin, Gerald (2012) Harm and the Volenti Principle. *Social Philosophy and Policy* 29(1): 309–321.
- Dworkin, Gerald (2020) Paternalism. *The Stanford Encyclopedia of Philosophy* (Fall 2020 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/fall2020/entries/paternalism/>
- Dworkin, Ronald (1993) *Life's Dominion - An Argument about Abortion, Euthanasia, and Individual Freedom*. New York: Vintage Books
- Eidelson, Benjamin (2015) *Discrimination and Disrespect*. Oxford: Oxford University Press
- Elster, Jakob (2011) How Outlandish Can Imaginary Cases Be? *Journal of Applied Philosophy*, 28(3): 241-258
- Enoch, David (2016) What's Wrong with Paternalism: Autonomy, Belief and Action. *Proceedings of the Aristotelian Society* 116(1): 21-48
- Fabre, Cécile (2006) Whose Body is it Anyway? Justice and the Integrity of the Person. Oxford, NY: Oxford University Press.
- Feinberg, Joel (1971) Legal Paternalism. *Canadian Journal of Philosophy* 1(1): 105–124.
- Feinberg, Joel (1984) *Harm to others*. New York: Oxford University Press.
- Feinberg, Joel (1986) *Harm to Self*. New York: Oxford University Press.
- Feinberg, Joel (1990) *Harmless Wrongdoing*. New York: Oxford University Press.
- Flanigan, Jessica (2017) Seat Belt Mandates and Paternalism. *Journal of Moral Philosophy* 14(3): 291–314.
- Fleurbaey, Marc (1995) Equal Opportunity or Equal Social Outcome? *Economics and Philosophy* 11: 25–55.
- Gligorov, Nada & Vitrano, Christine (2011) The Impact of Personal Identity on Advance Directives. *Journal of Value Inquiry* 45(2): 147-158
- Goold, Imogen & Savulescu, Julian (2008) In favour of freezing eggs for non-medical reasons, *Bioethics* 23(1): 47-58
- Grill, Kalle (2012) “Paternalism” in *Encyclopedia of Applied Ethics*, 2d ed., Chadwick, R. (ed.), Elsevier.
- Grill, Kalle (2015) “Antipaternalism as a Filter on Reasons” in *New Perspectives on Paternalism and Health Care*, Schramme, T. (ed.), Basel: Springer, pp. 47–63.
- Groll, Daniel (2012) Paternalism, Respect, and the Will, *Ethics* 122(4): 692-720
- Hanna, Jason (2018) *In Our Best Interest: A Defense of Paternalism*. Oxford: Oxford University Press
- Hansen, Rasmus S. (2016) ”Metode I normative politisk teori” in *Metode i Normativ Politisk Teori*, Hansen, R.S. & Midtgarrd, S.F. (eds), Samfundslitteratur
- Holtug, Nils (2011) Metode i politisk filosofi. *Politica* 43(3): 277-295

- Hurley, Susan (2006) "The 'What' and the 'How' of Distributive Justice and Health" in *Egalitarianism: New Essays on the Nature and Value of Equality*, Holtug, N. & Lippert-Rasmussen, K. (eds.), Oxford: Oxford University Press, pp. 308-334
- Kleinig, John (1983) *Paternalism*. New Jersey: Rowman & Allanheld
- Kogan, Terry S. (1976) The Limits of State Intervention: Personal Identity and Ultra-Risky Actions. *Yale Law Journal* 85(6): 826–846.
- Kymlicka, Will (2002) *Contemporary Political Philosophy An Introduction*. Oxford, New York, Oxford University Press
- Knight, Carl (2015) Abandoning the Abandonment Objection: Luck Egalitarian Arguments for Public Insurance. *Res Publica* 21: 119–135.
- Lippert-Rasmussen, Kasper (2005) Erik Rasmussens værdirelativisme: indhold og gyldighed. *Politica* 37(3): 274-286
- Lippert-Rasmussen, Kasper (2011) Eksperimentel Politisk Teori. *Politica* 43(3): 296-314
- List, Christian & Valentini, Laura (2016) "The Methodology of Political Theory" in *The Oxford Handbook of Philosophical Methodology*, Cappelen H., Gendler TS & Hawthorne J. (eds), Oxford University Press, pp. 525-553
- McDermott, Daniel (2008) "Analytical political philosophy" in *Political Theory: Methods and Approaches*, Leopold, D. & Stears, M. (eds.), Oxford; New York: Oxford University Press, pp. 11-28
- Midtgaard, Søren F. (2015) Non-Renounceable Rights, Paternalism and Autonomy. *Utilitas* 27(3): 347-364.
- Mill, John S. (2000 [1859]) *On Liberty*. New York: Dover Publications, Inc.
- Nielsen, Lasse (2013) Om metoden i normative politologi. *Politik* 16(3): 45-54
- Oshana, Marian A.L. (1998) Personal Autonomy and Society. *Journal of Social Philosophy* 29(1): 81-102
- Parfit, Derek (1973) "Later Selves and Moral Principles" in *Philosophy and Personal Relations: An Anglo-French Study*, Montegiore A. (ed.), London: Routledge & Kegan Paul Ltd., pp. 137–169
- Parfit, Derek (1984) *Reasons and Persons*. Oxford: Clarendon Press.
- Pedersen, Viki M.L. (2019) On the Anti-Paternalist Project of Reconciliation. *Utilitas* 31(1): 20–37.
- Pedersen, Viki M.L. & Midtgaard, Søren F. (2018) Is Anti-paternalism Enough? *Political Studies* 66(3): 771-785
- Rawls, John (1999 [1971]) *A Theory of Justice*. Rev. ed., Cambridge, Massachusetts: The Belknap Press of Harvard University Press
- Rawls, John (1995) Political Liberalism: Reply to Habermas. *The Journal of Philosophy* 92(3): 132-180
- Raz, Joseph (1986) *The Morality of Freedom*. Oxford: Clarendon
- Reagan, Donald H. (1983) "Paternalism, Freedom, Identity, and Commitment" in *Paternalism*, ed. Sartorius, RE., Minneapolis: University of Minnesota Press, pp. 113–138.

- Savulescu, Julian (2007) "Autonomy, the Good life, and Controversial Choices" in *The Blackwell Guide to Medical Ethics*, Rhodes, R., Fancis, LP. & Silvers, A. (ed.), Malden, MA : Blackwell Pub., pp. 17-37
- Schofield, Paul (2018) Paternalism and Right. *The Journal of Political Philosophy* 26 (1): 65–83.
- Schramme, Thomas (2008) Should We Prevent Non-Therapeutic Mutilation and Extreme Body Modification? *Bioethics* 11: 8–15.
- Sen, Amartya (1992) *Inequality Reexamined*, New York; Oxford; Russel Sage Foundation; Clarendon
- Shafer-Landau, Russ (2005) Liberalism and Paternalism, *Legal Theory* 11: 169-191
- Slavny, Adam, Spikerman, Kai, Lawford-Smith, Holly & Axelsen, David V. (2020) Directed Reflective Equilibrium: Thought Experiments and How to Use Them, *Journal of Moral Philosophy* 18(1): 12-25, Pre-proof version.
- Tadros, Victor (2011) *The Ends of Harm: The Moral Foundations of Criminal Law*. Oxford: Oxford University Press
- Weber, Max (2000 [1922]) *Makt og byråkrati – Essays om politikk og klasse, samfunnsforskning og verdier*. Gyldendal Norsk Forlag ASA
- White, Lucie (2020) "Personal Identity and Self-Regarding Choice in Medical Ethics" in *Theories of the Self and Autonomy in Medical Ethics*, Kühler, M. & Mitrovic, V. L. Mitrovic (ed.) Cham: The International Library of Bioethics, vol 83. Springer, pp. 31-47
- Wilkinson, T. M. (2011) *Ethics and the Acquisition of Organs, Transplantation and Rights over our Bodies* Oxford: Oxford University Press
- Witt, Karsten (2019) In Defence of Advance Directives in Dementia. *Pacific Philosophical Quarterly* 10(1): 2-21. DOI: 10.1111/papq.12292
- Witt, Karsten (2020) "Personal Identity and the Moral Authority of Advance Directives" in *Theories of the Self and Autonomy in Medical Ethics*, Kühler, M. & Mitrovic, V. L. Mitrovic (eds.), Cham: The International Library of Bioethics, vol 83. Springer, pp.