Anti-paternalism and Public Health Policy

Kalle Grill

Stockholm 2009
ABSTRACT


This thesis is an attempt to constructively interpret and critically evaluate the liberal doctrine that we may not limit a person’s liberty for her own good, and to discuss its implications and alternatives in some concrete areas of public health policy. The thesis starts theoretical and goes ever more practical. The first paper is devoted to positive interpretation of anti-paternalism with special focus on the reason component – personal good. A novel generic definition of paternalism is proposed, intended to capture, in a generous fashion, the object of traditional liberal resistance to paternalism – the invocation of personal good reasons for limiting of or interfering with a person’s liberty. In the second paper, the normative aspect of this resistance is given a somewhat technical interpretation in terms of invalidation of reasons – the blocking of reasons from influencing the moral status of actions according to their strength. It is then argued that normative anti-paternalism so understood is unreasonable, on three grounds: 1) Since the doctrine only applies to sufficiently voluntary action, voluntariness determines validity of reasons, which is unwarranted and leads to wrong answers to moral questions. 2) Since voluntariness comes in degrees, a threshold must be set where personal good reasons are invalidated, leading to peculiar jumps in the justifiability of actions. 3) Anti-paternalism imposes an untenable and unhelpful distinction between the value of respecting choices that are sufficiently voluntary and choices that are not. The third paper adds to this critique the fourth argument that none of the action types typically proposed to specify the action component of paternalism is such that performing an action of that type out of benevolence is essentially morally problematic. The fourth paper ignores the critique in the second and third papers and proposes, in an anti-paternalistic spirit, a series of rules for the justification of option-restricting policies aimed at groups where some members consent to the policy and some do not. Such policies present the liberal with a dilemma where the value of not restricting people’s options without their consent conflicts with the value of allowing people to shape their lives according to their own wishes. The fifth paper applies the understanding of anti-paternalism developed in the earlier papers to product safety regulation, as an example of a public health policy area. The sixth paper explores in more detail a specific public health policy, namely that of mandatory alcohol interlocks in all cars, proposed by the former Swedish government and supported by the Swedish National Road Administration. The policy is evaluated for cost-effectiveness, for possible diffusion of individual responsibility, and for paternalistic treatment of drivers. The seventh paper argues for a liberal policy in the area of dissemination of information about uncertain threats to public health. The argument against paternalism is based on common sense consequentialist considerations, avoiding any appeal to the normative anti-paternalism rejected earlier in the thesis.

Keywords: Alcohol Interlocks; Altruism; Anti-paternalism; Epistemic paternalism; Group consent; Harm principle; Interference; Invalidation of reasons; Liberalism; Limiting liberty; Private sphere; Product safety regulation; Public health policy; Reason-actions; Self-regarding; Social responsibility; Uncertain information; Withholding of information.
LIST OF PAPERS

This doctoral thesis consists of an introduction and seven papers:

2. Grill, K. ‘Anti-paternalism and Invalidation of Reasons’. Submitted manuscript.

Front cover picture: Odysseus and the Sirens. Detail from an Attic red-figured stamnos (pottery), ca. 480-470 BC. From Vulci.

Author: Kalle Grill, Department of Philosophy and History of Technology, Royal Institute of Technology, Stockholm, Sweden. Email: kgrill@kth.se.
PREFACE
Almost six years ago, I set out to understand what paternalism was all about and why some people were so fiercely against it. I considered myself a liberal of sorts, and still do, but it seemed to me that the value of liberty was invoked too readily and with too much confidence, in this area as in many others. After all, there seemed to be some agreement that paternalism involved on the one hand a limiting of liberty, but on the other hand a promotion of good. Preserving liberty and promoting good seemed to me both important. Why then should paternalism always be (prima facie) morally wrong?

This early impression has stayed with me. I still consider myself a liberal and I am still uncertain what I mean by that. However, I am ever more convinced that for all the greatness of liberty, it is not the moral trump card it is too often made out to be. Individual health and well-being are very important, and can sometimes be secured at the expense of liberty. As tyrannical as it may seem to restrain a person for her own good, as cruel can it be to stand by and let someone perish from her own mistakes or confusion. It may yet be that some liberal self-realization or Millian individuality is the highest form of life, though a life of community and shared joys and sorrows is at least a close second. However, it cannot be that mere everyday restriction, coercion, or intrusion should never be suffered for preserved health or survival.

I sometimes wish with the liberal tradition that there were rational true selves with unshakable preferences inside all of us, who could be asked to direct our lives when we need steadying. However, for better and worse we are just the fallible human beings that we are, prone to bias and misjudgement and heavily influenced by our surroundings. Such are the selves that should be at liberty. Equally fallible, of course, are governments and other authorities. Nowhere can we turn for enlightened direction, neither to ourselves nor to some external director. We must simply make do with what little ability we have, helping ourselves and each other enjoy the most liberty and the most well-being that we can attain, making difficult choices in the process.

ACKNOWLEDGEMENTS
Writing this thesis has mostly been lonesome work, as academic work often is. That said I am very grateful for the many opportunities I have had to discuss my ideas with friends and colleagues at the Division of Philosophy, at quite a few international conferences, and during my visit to the University of California in San Diego (UCSD). I am no less grateful for the love and support of my close friends and family over the years, especially my mother Lisa Grill for her constant interest in and encouragement of my sometimes esoteric work, my friend and former girlfriend Hanna Ögren for her support and her refreshing ‘naive philosophy’, and my present girlfriend Camilla Claesson for her embrace of my various philosophy projects, as well as for her analytic mind and open heart.

Two of my close friends are also my best critics – my fellow graduate students Lars Lindblom and Niklas Möller. When it comes to the intricacies of liberal political philosophy as well as my particular takes on it, Lars has been my most dependable and
thorough critic. Niklas has consistently provided excellent comments on argument and presentation more generally (he has also vigorously tried to slow down my thesis work by involving me in courses, readings and discussions on meta-ethics and philosophy of mind and language). My main supervisor Sven Ove Hansson has provided useful critique on terminology and argument. Without Sven Ove this thesis would not have been written, as he authored the original research plan and secured the funds. My assistant supervisor Martin Peterson has provided reliable and speedily feed-back on presentation and sometimes argument, ever urging me to get clearer. This has been good advice more often than I realized at the time. Other colleagues at the Division of Philosophy have provided useful critique on parts of the thesis, especially Sara Belfrage and Dan Munter.

During my six months at UCSD, Richard Arneson was my host and very generous discussion partner. Our common probings into the subtleties of anti-paternalism were crucial in giving me a sense that I was on the right track, or at least not on an unreasonable one. The UCSD philosophy department in general was very welcoming and I am thankful.

Financial support from the Swedish Council for Working Life and Social Research (FAS) and, for the UCSD visit, from the Swedish Foundation for International Cooperation in Research and Higher Education (STINT), is very gratefully acknowledged.

Stockholm Kalle Grill
February 2009
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>iii</td>
</tr>
<tr>
<td>Preface</td>
<td>v</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>v</td>
</tr>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1. Overview</td>
<td>1</td>
</tr>
<tr>
<td>2. A Brief History of Anti-Paternalism</td>
<td>2</td>
</tr>
<tr>
<td>3. Definitions</td>
<td>5</td>
</tr>
<tr>
<td>3.1 Paternalism 1972-2008</td>
<td>5</td>
</tr>
<tr>
<td>3.2 Discussion of Definitions and My Own View</td>
<td>10</td>
</tr>
<tr>
<td>3.3 Descriptive and Normative Definitions</td>
<td>13</td>
</tr>
<tr>
<td>4. Reasons and the Good</td>
<td>15</td>
</tr>
<tr>
<td>4.1 Knowing Best</td>
<td>15</td>
</tr>
<tr>
<td>4.2 Acting on Judgements</td>
<td>17</td>
</tr>
<tr>
<td>4.3 Health Promotion and Moralism</td>
<td>19</td>
</tr>
<tr>
<td>5. Values, Reasons and Principles</td>
<td>20</td>
</tr>
<tr>
<td>5.1 Reasons without Influence</td>
<td>21</td>
</tr>
<tr>
<td>5.2 Anti-Paternalism as an Influence-Regulating Principle</td>
<td>22</td>
</tr>
<tr>
<td>5.3 Against Anti-Paternalism</td>
<td>22</td>
</tr>
<tr>
<td>5.4 Constant Overriding</td>
<td>25</td>
</tr>
<tr>
<td>5.5 Liberalism without Anti-Paternalism</td>
<td>26</td>
</tr>
<tr>
<td>5.6 More General Theories</td>
<td>27</td>
</tr>
<tr>
<td>6. Liberty or Health?</td>
<td>29</td>
</tr>
<tr>
<td>6.1 Health</td>
<td>29</td>
</tr>
<tr>
<td>6.2 Liberty</td>
<td>30</td>
</tr>
<tr>
<td>6.3 Libertarian Paternalism</td>
<td>32</td>
</tr>
<tr>
<td>7. Overview of Papers</td>
<td>34</td>
</tr>
<tr>
<td>References</td>
<td>36</td>
</tr>
<tr>
<td>Paper 1 - The Normative Core of Paternalism</td>
<td>39</td>
</tr>
<tr>
<td>Paper 2 - Anti-Paternalism and Invalidation of Reasons</td>
<td>53</td>
</tr>
<tr>
<td>Paper 3 - Paternalistic Interference</td>
<td>77</td>
</tr>
<tr>
<td>Paper 4 - Liberalism, Altruism and Group Consent</td>
<td>105</td>
</tr>
<tr>
<td>Paper 5 - Anti-Paternalism and Public Health Policy - The Case of</td>
<td>125</td>
</tr>
<tr>
<td>Product Safety Regulation</td>
<td></td>
</tr>
<tr>
<td>Paper 6 - Responsibility, Paternalism and Alcohol Interlocks</td>
<td>135</td>
</tr>
<tr>
<td>Paper 7 - Epistemic Paternalism in Public Health</td>
<td>153</td>
</tr>
</tbody>
</table>
Introduction

[It is possible, and at times justifiable, to coerce men in the name of some goal (let us say, justice or public health) which they would, if they were more enlightened, themselves pursue, but do not, because they are blind or ignorant or corrupt.

–Isaiah Berlin

1.

1. OVERVIEW

This is a thesis on political morality and practical ethics. To paraphrase Joseph Raz (1986, p. 4), it is a thesis on ethics, which concentrates on certain moral issues because of their political implications. The moral issues are the nature of paternalism and anti-paternalism, and the reasonableness of the latter. The political implications are quite general but are in the thesis explicitly exemplified in the area of public health policy, with special attention to the tendency of such policy to limit the liberty of groups of people. Product safety regulation, mandatory alcohol interlocks, and the dissemination of information about uncertain threats to public health are investigated in some detail. The most novel contribution is perhaps the attention paid to reasons and their role in anti-paternalist doctrines, both as formulated in theory and as applied in practice.

Throughout, the focus is more on normative issues than on conceptual analysis. Paternalism is (in the first paper) defined generically, to allow more detailed conceptions to capture variations on the opposition to benevolent limiting of liberty that is part of the liberal tradition. This opposition is then analyzed and criticized (in paper two and three). Setting aside commonsense consequentialist or empirical arguments, the purely normative principle that benevolence can never justify limiting liberty is found unreasonable on several grounds. Nonetheless, an attempt is made (in paper four) to contribute to the anti-paternalist cause by suggesting some rules by which the liberal may possibly justify limiting the liberty of groups of individuals without invoking their good, but only their (partial) consent. Both the more constructive interpretation of anti-paternalism and the more negative rejection of this doctrine in the more theoretical first part of the thesis inform the discussion of concrete policy areas in the more practical later part (papers five, six and seven).

In this introduction, I will first briefly describe the history of anti-paternalism as I see it, with a focus on John Stuart Mill and Joel Feinberg as the main proponents of this doctrine. After so setting the stage, I will move on to consider seven definitions of paternalism and discuss their strengths and weaknesses. This will further demonstrate the context and background against which this thesis has developed and will naturally lead me to explain my own view of how paternalism should be understood. In the fourth chapter I discuss what it means to act for the good of someone else and argue that the inherent morally problematic aspect of paternalism is not overriding a person’s own

conception of her good, but rather limiting her liberty. In connection, the distinction between paternalism and moralism is defended. Not until the fifth chapter do I give a summary of the arguments against anti-paternalism that are developed in the first three papers. I place these arguments in the wider context of practical reasoning and in particular in the context of what I call influence-regulating principles – principles that block reasons from influencing the moral status of actions according to their strength. In this chapter I also explain how liberalism can do without paternalism and how my argument against anti-paternalism relates to more general theories of (political) morality. In chapter six I try to say something more positive about the value of health and of liberty and consider so called libertarian paternalism as one way to justify public health policy. The very short seventh chapter contains a brief survey of the six articles that follow upon this introduction.

2. A BRIEF HISTORY OF ANTI-PATERNALISM

Much of contemporary debate on paternalism takes as its starting point John Stuart Mill’s *On Liberty* (1991 [1859]). Mill does not himself use the term paternalism, but famously formulates this liberty principle:

> That principle is, that the sole end for which mankind are warranted, individually or collectively, in interfering with the liberty of action of any of their number, is self-protection. That 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 opinion of others, to do so would be wise, or even right. (p. 14)

This moral principle has a positive and a negative part. The positive part is the harm principle, saying that preventing harm to others is a valid purpose for limiting liberty (or interfering with liberty – I make no distinction). The negative part is that no other purpose is valid. This part can be further divided into resistance to particular purposes – to promote physical good, to promote moral good, to promote doing the right thing (being moral). In his debate with Patrick Devlin over the relationship between law and morality and the proposal to decriminalize private homosexuality, H.L.H. Hart (1963) made a point of distinguishing legal moralism – ‘to enforce positive morality’ – from paternalism – ‘to protect individuals against themselves’ (p. 31).

Hart’s characterization of paternalism is suggestive but imprecise. Neither Mill nor Hart were especially concerned with conceptual analysis. Mill was above all concerned to specify the principle he advocated, Hart the principle he rejected. Hart

---

2 Following Feinberg (1986, p. ix). Sometimes ‘the harm principle’ refers to what I (following e.g. Arneson 1989, p. 409) just called ‘the liberty principle’. 
rejected moralism, but accepted paternalism. His common sense arguments for this acceptance are instructive:

> Choices may be made or consent given without adequate reflection or appreciation of the consequences; or in pursuit of merely transitory desires; or in various predicaments when the judgment is likely to be clouded; or under psychological compulsion; or under pressure by others of a kind too subtle to be susceptible of proof in a court of law. (p. 33)

As a result, people make choices that harm them, and when they do we should sometimes stop them.

Our understanding of the many ways in which we are products of our environments have only deepened since Hart’s times. It is therefore not surprising that Mill’s most distinguished follower in the 20th century made voluntariness a central concept in his resistance to paternalism. Joel Feinberg (1971; 1986) restricted his liberalism to the subject of criminal law and defined *legal paternalism* so:

> 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.’ (1986, p. 4)

Feinberg rejected this principle in favour of what he would have preferred to name *soft anti-paternalism* but, yielding to convention, called *soft paternalism* – the principle that the state may limit a person’s liberty for her own good when and only when her conduct is not voluntary enough, or if intervention is needed to establish how voluntary it is (1986, p. 12). 3 *Hard paternalism*, in contrast, allows limiting a person’s liberty even when her conduct is fully voluntary. Feinberg spends a large part of his article and later book discussing how voluntary is voluntary enough. The criteria include basic competence (not an infant or insane or comatose); absence of manipulation, coercion and duress; informedness; and absence of distorting circumstances (fatigue, agitation, passion, drugs, pain, neurosis, time pressure) (p. 115).

When we fall under the threshold of sufficiently voluntary, we are not in tune with what Feinberg sometimes calls our ‘true self’, and so restraining our non-voluntary actions is not really limiting our liberty. While the harm principle protects a person from others, soft paternalism protects him from his ‘non-voluntary choices’, which, being the genuine choices of no one at all, are no less foreign to him.’ (1986, p. 12) When we stay over the threshold, on the other hand, there are no good reasons to intervene since no wrong is done: What harm we voluntarily bring on ourselves is not wrongful, in line with the Roman Law maxim of *volenti non fit injuria* – ‘To one who freely consents to a thing no wrong is done, no matter how harmful to him the consequences may be.’ (1971, p. 107)

---

3 Feinberg thereby rejects not only legal paternalism as specified, but also the stronger version where ‘sometimes’ is substituted for ‘always’, given that the actor acts voluntarily.
Soon after Feinberg’s first rejection of hard legal paternalism in 1971, the debate on paternalism turned more conceptual, as will be explored in the following section. On the more normative side, most of the attention was on soft paternalism (e.g. Hodson 1977; Van de Veer 1986) and on further restricting anti-paternalism by making exceptions for the promotion of liberty or autonomy (e.g. Dworkin 1972; Kleinig 1983; Sneddon 2001; Husak 2003; De Marneffe 2006). Gerald Dworkin (1972) argues that liberty-preserving hard paternalism should be accepted and that Mill paved the way for such acceptance with his rejection of voluntary slavery. Others followed Dworkin in proposing other accounts of liberal values that similarly justify hard paternalism.

The most noteworthy recent development on the normative side of the debate is arguably the embrace of paternalism by Richard Thaler and Cass Sunstein (2003a; 2003b) under the provocative label *libertarian paternalism*. Libertarian paternalism is paternalism in the sense that ‘it attempts to influence the choices of affected parties in a way that will make choosers better off.’ (2003b, p. 1162) It is libertarian in the sense that ‘people should be free to opt out of specified arrangements if they choose to do so.’ (p. 1161) Thaler & Sunstein justify benevolent influence on choice by appeal to a sort of expanded soft paternalism. To briefly return to Mill, Hart found his picture of the human being naive:

[A] middle-aged man whose desires are relatively fixed, not liable to be artificially stimulated by external influences; who knows what he wants and what gives him satisfaction or happiness; and who pursues these things when he can.

We might in turn find Hart’s picture of the middle-aged man naive (and sexist), but the point is that Mill assumed that, with a few exceptions he did not spend much ink discussing, people act on their set preferences. Feinberg recognised that we often do not so act, but apparently assumed that we all have a middle-aged man inside of us, our true self, with equally set preferences. The point that Thaler & Sunstein has driven home is that because of the great impact of default rules, starting points and framing effects, most of the time we simply have no preferences independently of the choice situations we face. Therefore, adjusting those choice situations to promote our good does not limit our liberty.

Much has been written on paternalism and anti-paternalism that is not captured in this brief history of course. I will turn to some more conceptual points in the following section. Parallel to the more principled arguments surveyed, there has always been straightforward consequentialist arguments for and against paternalism, mainly of two kinds: 1) People, and especially people in the government, cannot be trusted to limit other people’s liberty for their own good, because people are selfish, power corrupts, etc. 2) Everything we do and are affects others one way or other and so there is always a non-paternalistic rationale for limiting liberty. Both of these arguments are valid, but they do not address what I take to be the central normative dimension of paternalism – whether or not the promotion (not just as an attempt or an excuse but actually) of a person’s good can contribute to justifications for limiting her liberty.
3. Definitions

As noted, I argue in the first paper for defining paternalism on normative grounds. More specifically, detailed conceptions of paternalism should be based on substantial normative views on what reasons are morally problematic when invoked for what actions, or on attempts to describe such substantial normative views. In this chapter, I will defend this approach in general terms. I will also explain more in detail what I think this normative approach entails concerning the defining and understanding of paternalism. First, however, I will survey seven definitions from the literature on paternalism over the last decades and discuss their differences, their strengths and weaknesses. This is not a comprehensive account (more thorough overviews, though less up to date, can be found in Van de Veer 1986 and Nikku 1997). I merely aim to contrast the approach to anti-paternalism defended and employed in this thesis to other influential approaches.

3.1 Paternalism 1972-2008

In his argument for liberty-preserving paternalism, Gerald Dworkin (1972) adopted this brief characterization:

By paternalism I shall understand roughly the interference with a person’s liberty of action justified by reasons referring exclusively to the welfare, good, happiness, needs, interests or values of the person being coerced. (p. 65)

The characterization comes very close to the anti-paternalist part of Mill’s liberty principle. Paternalism is basically benevolent interference. Later in the article, it becomes clear that Dworkin finds paternalism morally problematic only when it is not consented to (p. 77).

Dworkin is mainly concerned with the normative question of whether benevolent limiting of liberty is ever justified. He supposedly neither intends a full conceptual definition nor a precise description of a normative position, but merely to set the stage for his argument (as indicated by the ‘roughly’). The debate on paternalism took a more conceptual turn with Bernard Gert and Charles Culver (1976) objecting to defining paternalism in terms of interference with liberty of action. Gert & Culver point out that paternalism need not involve an attempt to control behaviour, but can be a case of ‘killing, causing pain, disabling, depriving of pleasure, deception, or breaking a promise.’ (p. 51) These things cannot, they claim, be understood as interferences with liberty of action. Dworkin (1983) later concedes this point but defends his earlier definition as appropriate given his interest in ‘the proper limits of state coercion’ (p. 105). Gert and Culver explicitly do not restrict themselves to legal paternalism. They offer this general definition:

A is acting paternalistically toward S if and only if A’s behaviour (correctly) indicates that A believes that

1. his action is for S’s good
(2) he is qualified to act on S’s behalf
(3) his action involves violating a moral rule (or doing that which will require
him to do so) with regard to S
(4) he is justified in acting on S’s good independently of S’s past, present, or
immediately forthcoming (free, informed) consent
(5) S believes (perhaps falsely) that he (S) generally knows what is for his
own good. (pp. 49-50)

What is most noteworthy about this definition is that it is put entirely in the head of the
paternalist. No interference with liberty is required, but neither is an actual violation of a
moral rule. As noted by Donald Van de Veer (1986, p. 37), I may, according to this
definition, believe that I am exorcising your demon through some obscure but harmless
ritual and be acting paternalistically, even though there is no benefit to you nor any other
effect on you. The focus is very much on the inner life and character of the agent, and
not at all on the action and its consequences. This is particularly inapt for discussing
paternalism in politics, where effects are, arguably, more important than motives.4

Apart from the heavy focus on beliefs, Gert & Culver’s definition is traditionally
Millian in its focus on benevolence (condition 1) and lack of consent (condition 4). The
interference condition (3) is formulated in terms of violating a moral rule, which, if it
were not for the belief qualification, would ensure that paternalism raises moral issues.
Condition 5 contains the traditional exclusion of incompetents (infants, animals), a very
weak voluntariness condition. The purpose of condition 2 seems to be to exclude, for
some reason, cases where the agent is obviously less competent than the person she acts
towards, such as when the agent is ‘a small child’ (pp. 50-51).

Van de Veer (1986) has another issue with Dworkin’s rough characterization.
The term ‘justified’ implies, he argues, that Dworkin’s ‘definition’ is normative in that
paternalism is supposed to be justified, and that Dworkin fails to consider the motives
of the agent. An action that is justified by personal good reasons may be “‘paternalistic in
result’”, Van de Veer admits, but the action as such is not paternalistic (p. 27). Van de
Veer sometimes (p. 26) takes Gert & Culver, too, to assume that paternalism is justified
(by condition 4) and sometimes not (because of the belief qualification) (p. 37). In
contrast to their (overly) rich definition, Van de Veer’s (1986) own is minimalist in its
core content, though excessive in its detail:

A’s doing or omitting some act X to, or toward, S is paternalistic behavior if and
only if

1. A deliberately does (or omits) X

and 2. A believes that his (her) doing (or omitting) X is contrary to S’s operative
preference, intention or disposition at the time A does (or omits) X [or
when X affects S—or would have affected S if X had been done (or
omitted)]

4 In a later article, Gert and Culver (1979) repeat their definition but treats condition 3 as if it was not
phrased in terms of beliefs, but in terms of actual violation (pp. 199-200 and esp. note 4).
and 3. A does (or omits) X with the primary or sole aim of promoting a benefit for S [a benefit which, A believes, would not accrue to S in the absence of A’s doing (or omitting) X] or preventing a harm to S [a harm which, A believes, would accrue to S in the absence of A’s doing (or omitting) X].

(p. 22)

Interestingly, Van de Veer’s definition is as subjectivist as Gert & Culver’s. On this definition, too, it may be paternalism if I believe that I am exorcising your demon through some obscure but harmless ritual.

Van de Veer takes it to be an advantage that on his definition, paternalism need not raise any moral issues. Very reasonably, however, Van de Veer is mainly interested in specifically those cases of paternalism that do raise moral issues. As he notes, ‘many paternalistic practices happen to involve acts which are presumptively wrong’ (p. 21). We might then infer a Van de Veerian definition of morally interesting paternalism by adding this interference condition:

and 4. A’s doing (or omitting) X presumptively wrongs S.\(^5\)

This condition would also offer a connection with actual effects, independently of beliefs (it could of course also come in a belief variation).

Van de Veer’s condition 1 only requires that the doing is an intentional action, something that is taken for granted in other definitions. Condition 2 is foremost a consent condition, though without the added condition 4 it also serves the role of a weak interference condition. While in Gert & Culver’s definition the agent believes she acts ‘independently of’ consent, in Van de Veer’s she believes she acts ‘contrary to’ consent. The first formulation includes acting where there is no (belief that there is a) preference either way. This makes sense, as there is an important distinction to be made between expressed consent and lack of expressed consent. Important because an interference that is explicitly consented to is arguably not morally problematic, it does not limit liberty in a morally relevant sense.

Van de Veer’s condition 3 is the benevolence condition, which in his rendering explicitly includes both harm-prevention and good-promotion. It is, like Gert & Culver’s, subjective to the agent. No actual benefit is required. While Gert & Culver requires only that the agent believes the action will be beneficial, Van de Veer settles for the stronger requirement that the action must be primary or solely intended to benefit the person acted towards. However, this difference is softened by Gert & Culver explaining that ‘insofar as A’s behaviour toward S is paternalistic, it is only S’s good, not the good of some third party, which is involved.’ (p. 50)

Like Van de Veer, David Archard (1990) tries to improve on Dworkin’s and Gert & Culver’s accounts (without position himself in relation to Van de Veer). Like Van de Veer’s, his definition is explicitly designed to be non-normative or morally neutral.

\(^5\) Van de Veer would not appreciate this addendum. In fact, he laments his own earlier ‘fixation on the morally interesting subset of cases of paternalism, interesting because controversial’ (p. 35).
However, in contrast to Van de Veer, Archard means to incorporate the morally interesting aspect of paternalism in his definition:

P behaves paternalistically towards Q iff:

1. P aims to bring it about that with respect to some state(s) of affairs which concerns Q’s good Q’s choice or opportunity to choose is denied or diminished;
2. P’s belief that this behaviour promotes Q’s good is the main reason for P’s behaviour;
3. P discounts Q’s belief that P’s behaviour does not promote Q’s good.

Part of the background to Archard’s interference condition (1) is that Dworkin (1983), going conceptual, argues that Gert & Culver’s definition is too narrow, since paternalism need not involve the violation of a moral rule. It is paternalism, Dworkin argues, for a husband to hide his sleeping pills from his suicidal wife (p. 106). Archard’s definition is meant to accommodate this case. Though the husband is not obligated to share his pills, his wife’s opportunity to choose is diminished because he would have done so “[i]n the normal course of events” (p. 37). On the other hand, increasing a person’s choices cannot, Archard argues, be paternalistic. The definition is morally neutral because we have duties to promote other’s well-being, but also, supposedly, duties to respect their choice, and the definition does not say which is stronger (pp. 41-42).

Archard’s benefit condition (2) is all but identical to Van de Veer’s (assuming a sole reason is a main reason). Condition 3 can be understood as a rich consent condition. Archard finds it essential to paternalism that it involves ’the usurpation of one person’s choice of their own good by another person.’ (p. 36) Therefore, it is not paternalism if P’s belief about Q’s belief is incorrect, as there is then no actual usurpation (p. 39). Further, Archard prefers an ‘independently of’ understanding of consent, and accepts that the consent may be given previously but not that it be merely anticipated (p. 40).

Seana Shiffrin (2000), like Archard, aims to provide ‘a conception of paternalism that fits and makes sense of our conviction that paternalism matters’ (p. 212):

[P]aternalism by A toward B may be characterized as behavior (whether through action or through omission)

(a) aimed to have (or to avoid) an effect on B or her sphere of legitimate agency
(b) that involves the substitution of A’s judgment or agency for B’s
(c) directed at B’s own interests or matters that legitimately lie within B’s control
(d) undertaken on the grounds that compared to B’s judgment or agency with respect to those interests or other matters, A regards her judgment or agency to be (or as likely to be), in some respect, superior to B’s. (p. 218)
What is most noteworthy with this definition (characterization) is that there is no benevolence condition. Shiffrin argues that it may be paternalism to interfere with a person for the sake of a third party, or without concern for anyone’s welfare, as long as the interference condition is satisfied (pp. 215-7). Condition d is supposed to specify a motive; Shiffrin seems to hold that the sheer belief that one can do something better (in some unspecified sense) than another can move one to act, giving the example of articulating a point at a talk. Indeed, this motive is the central normative component of paternalism on Shiffrin’s account, since it ‘delivers a special sort of insult to competent, autonomous agents.’ (p. 220). The three remaining conditions all concern the specification of interference (conditions a and c are all but identical). However, with its focus on legitimate control, this specification may be taken to incorporate a consent condition. Shiffrin earlier in the article explicitly prefers a ‘independently of’ to a ‘contrary to’ understanding of consent (and supposedly of legitimate control) (p. 214).

In contrast to Archard, Shiffrin thinks that increasing a person’s range of options (her choices) can be paternalistic, if the person prefers not to have those extra choices. This is so because in such cases, where A increases B’s options against her will, ‘A forcibly substitutes her judgment about the right way for B to exercise and develop her agency.’ (p. 214) In other cases, the substitution is not so much of judgment as of agency – A and B may agree on what should be done but B may think that she can make it happen more effectively, even though it falls under A’s ‘sphere of legitimate agency’. It is this focus on substitution of judgment or agency that leads Shiffrin to reject the benevolence condition. She argues that ‘we should have the same sort of normative reaction’ regardless of who, if anyone, benefits from the interference.

Peter De Marneffe (2006) objects to Shiffrin’s anti-paternalism and argues that substitution of judgment is common and not necessarily insulting. The motive for imposing speeding limits may be the good of drivers or the good of third parties – in either case the government substitutes its judgment for that of the driver (pp. 77-79). We are imperfectly rational and so may make mistakes both in our considerations of the interest of others and of our own interests (p. 80). Paternalism would be morally problematic if paternalistic reasons were silenced somehow, or if there were moral rights against paternalism, but De Marneffe finds no good reason to accept any of these ideas (at least not on a Scanlonian account of rights) (pp. 83-86).

De Marneffe observes that policies may be paternalistic in their effects without a paternalistic motive and is therefore drawn to an account in terms of justification. However, he finds no such account that allows that unjustified policies can be paternalistic. Therefore, and perhaps to accommodate linguistic intuitions, he proposes this hybrid definition:

[A] government policy is paternalistic toward A if and only if (a) it limits A’s choices by deterring A from choosing to perform an action or by making it more difficult for A to perform it; (b) A prefers A’s own situation when A’s choices are not limited in this way; (c) the government has this policy only because those in the relevant political process believe or once believed that this policy will benefit A in
some way; and (d) this policy cannot be fully justified without counting its benefits to A in its favor. (pp. 73-74)

The definition is rather traditional, with a the interference condition in terms of choice, reminiscent of Archard’s; b a consent condition (where if A has no explicit preference we consider what A would prefer if she thought about it [p. 73, note 15]); c a motivational benevolence condition; and d a justificationary benevolence condition. What is novel is the hybrid between a motivational and a justificationary definition.

In a definition of paternalism for The Stanford Encyclopedia of Philosophy, first formulated in 2002, Dworkin (2008) drops the ‘of action’ bit from his early characterization of interference with liberty and adds interference with autonomy as an alternative, proposing this very Millian definition:

I suggest the following conditions as an analysis of X acts paternalistically towards Y by doing (omitting) Z:

1. Z (or its omission) interferes with the liberty or autonomy of Y.
2. X does so without the consent of Y.
3. X does so just because Z will improve the welfare of Y (where this includes preventing his welfare from diminishing), or in some way promote the interests, values, or good of Y.

There is very clearly an interference condition (1), a consent condition (2) and a benevolence condition (3). In an earlier article, Dworkin (1983) equates violation of autonomy with substitution of judgment (p. 107), moving him close to Shiffrin. The consent condition is of the ‘independently of’ kind. As pointed out by De Marneffe, Dworkin’s benevolence condition is ambiguous between motivational and justificationary interpretations.

Dworkin’s recent definition completes the survey. I have returned repeatedly to the three conditions that specify the form of paternalism that Mill rejected in his liberty principle. In the following section, I will summarize the conceptual debate in relation to the anti-paternalism of Mill and Feinberg.

3.2 DISCUSSION OF DEFINITIONS AND MY OWN VIEW

The survey of recent definitions shows a development in our understanding of the interference condition – from interference with liberty of action to presumptive wrongs and on to diminishing of choice or substitution of judgment or agency. However, Dworkin keeps or returns to the more traditional ‘interference with liberty’, adding ‘or autonomy’, perhaps with the thought that these two liberal notions may be specified to capture such things as diminishing of choice and substitution of judgment. Since, arguably, it is the interference condition that carries the moral weight of anti-paternalism, it is very important that it be adequately specified (in paper three I argue that it has not been).
Concerning consent, authors in general, very reasonably, agree that the important distinction is between acting with a person’s (expressed or possibly inferred) consent and without it (I have passed over a lot of more detailed issues on the nature of consent). Implicitly, Shiffrin incorporates consent into her account of interference, which is very reasonable on the assumption that acting towards a person with her consent is not really to limit her liberty. This assumption is supported by e.g. Husak (1981, p. 31): ‘It would seem that a necessary condition for describing an act as an interference with the freedom of the agent is that the agent did not consent to it.’ In fact, Mill could be read to make this assumption since he seems to equate ‘interfering with the liberty of action’ with ‘the exercise of power over someone against his will’, with an explicit consent condition only in the second formulation. In most of the papers in this thesis, I make the same assumption.

Merging interference and consent, we may talk of the action component of paternalism as that part which concerns the specification of the action type, in all its intricacies, including whether or not and in what sense the action has been consented to. This component can be further delimited in terms of other properties of the target of the action. Gert & Culver and Van de Veer restrict paternalism to targets over some degree of competence. Others explicitly do not, or leave the matter aside. When moving to normative considerations, however, most all find that voluntariness or some such property is essential to distinguishing morally problematic interference. As noted, one of Feinberg’s contributions is to take this condition very seriously. On Feinberg’s account, it is not interference, in the morally relevant sense, to restrain a person who acts insufficiently voluntarily.

The benevolence condition is rather stable over the definitions surveyed. With Shiffrin the notable exception, the joint assumption is that paternalism is acting for someone’s good or benefit, where no distinction is made between harm-prevention and benefit-provision, or between act and omission. This is consistent with Mill’s and Feinberg’s rejection of the act/omission distinction (Mill 1991 [1859], p. 15; Feinberg 1984, chapter 4). While Mill’s and Feinberg’s focus is on harm-prevention, this is only natural given that it is less morally problematic to limit a person’s liberty in order to prevent harm to her than to provide her with a benefit, if there is a distinction to be made.

On Shiffrin’s account, the essential aspect of paternalism is the unwelcome involvement with a person’s sphere of legitimate control in combination with the paternalist regarding herself as superior in some sense. Therefore, motive in a strict sense does not matter. De Marneffe criticises Shiffrin on the ground that substitution of judgment may have different motives, but this seems to be part of Shiffrin’s very point. Shiffrin explicitly admits that her account is dependent on an account of legitimate control that she does not provide while at the same time insisting that the ‘motive’ of disrespect for agency is not restricted to rights-violations (pp. 218-9). This implies that there is a sphere of authority where unwelcome involvement is prima facie wrong regardless of motive. This is an embryo to a defence of traditional anti-paternalism, but also more generally of anti-(unwelcome involvement). A shared rationale for both anti-
paternalism and for a broader liberalism of non-interference would to some extent muddle the distinction between benevolence and other reasons for interference. However, since the rationale for anti-paternalism is still very much under discussion, it seems reasonable to keep the possible distinction and confine paternalism to benevolent unwelcome involvement.

Pace Shiffrin then, there is widespread agreement that paternalism involves benevolence. Benevolence, however, can be either a motive or a justification, as noted by Van de Veer and brought out more clearly by De Marneffe. With the early Dworkin a possible exception, all definitions surveyed agree that paternalism requires a paternalistic motive. These motivational accounts have the consequence that if I, as a third party, observe you imposing a benefit on a person against her will, for some obscure reason (such as to make good on a bet), and I judge your action to be justified by this benefit, there is no paternalism involved. This is so even if I have the power to stop you and would have done so was it not for the benefit I want imposed on the person (as long as my omission does not count as an interference with the person). If I actively manipulate or force you into imposing this benefit, I may perhaps be said to be acting through you, though this is not at all obvious. Even such cases, therefore, may involve no paternalism. There is also the opposite or corresponding case where I, for some higher purpose, let you impose a benefit on a person, even though I could have stopped you, and even though you are blind to the higher purpose and act only for the sake of the benefit. On motivational accounts, the person is a victim of paternalism even if, were it not for the higher purpose, she would not have been imposed upon.

While there is widespread agreement on the importance of a benevolent motive, the surveyed accounts diverge on whether this motive should be the only reason or the main reason, or whether actions are paternalistic to the extent that actions have this motive (suggested by Gert & Culver and more explicitly proposed by Kleinig, 1983, p. 12), or whether the motive should be necessary, as specified by De Marneffe:

In my view, what is relevant to the paternalism of a policy is the truth of the counterfactual that it would not be the government’s policy had some government official not counted a paternalistic reason in its favor. (p. 74, note 16)

As long as motivational paternalism does not affect the justification of actions, these are all plausible accounts. Motivational paternalism may for example be relevant to the moral evaluation of a person’s character. If, however, paternalism is relevant for which actions are justified, as argued by Mill and Feinberg, then only De Marneffe’s account is adequate, since on all other accounts a personal good reason can tip the balance in favour of paternalistic action without making the action as such paternalistic. Such tipping would be rejected as unacceptable by anti-paternalists such as Mill and Feinberg. Both Van de Veer and Archard consider the possibility that benevolence may be a contributory reason for interference without being the main reason. They both conclude that such contributory reasons cannot make an interference paternalistic (Archard p. 38; Van de Veer pp. 27-28). Remember, however, Feinberg’s definition of legal paternalism:
A personal good reason is a good and relevant, _though not necessarily decisive_, reason in support of interference.

On my own view, the normative core of paternalism is the invocation, in favour of limiting a person’s liberty, of a reason that concern her good, regardless of whether this reason is the only reason, the main reason, a sufficient but secondary reason, a weaker contributory and possibly redundant reason, or, along de lines of De Marneffe, a necessary contributory reason in any set of sufficient reasons (this is the thesis of the first paper). If we are concerned with justification, the last form is the most salient. I agree with De Marneffe’s caption of justificationary anti-paternalism as the principle that if a policy (or an action) limits the liberty of a person A (without her consent), then ‘it is wrong for the government to adopt this policy unless it can be fully justified without counting any benefit to A in its favor.’ (p. 76) I take it that this is Mill’s and Feinberg’s position.⁶

All the definitions surveyed aim to define paternalistic policies, actions, or behaviour. I think this is a mistake. On what I call the _reason account_, what is paternalistic is always the combination of a reason and an action, or in other words a reason for an action. These two components can be specified in various ways to capture various normative positions. As noted, reasons can have different forms of impact and be on different levels of intentionality (motives, justifications). The reason account prepares the ground for principles that operate on reasons, as will be discussed below and in paper two. In contrast, the definitions surveyed are all committed to the _action account_, on which what is paternalistic is an action or policy, which is partly defined in terms of what reasons are (or could be) invoked to support it. The action account is perhaps more true to our linguistic habits, but it is unable to appropriately capture various forms of anti-paternalism.

### 3.3 Descriptive and Normative Definitions

The authors of the definitions surveyed above disagree on to what extent definitions should be normative. Some (Van de Veer) prefer definitions that do not raise moral issues at all, while others (Archard) aim to capture the negative side of paternalism, regardless of whether or not this aspect is balanced out by a positive side. Only Shiffrin explicitly aims to capture something that is prima facie wrong. I propose that this is exactly what definitions of paternalism should do – to capture something that is prima facie wrong, or is regarded to be so.

Dworkin (2008) states on defining paternalism that ‘[a]s a matter of methodology it is preferable to see if some concept can be defined in non-normative terms and only if that fails to capture the relevant phenomena to accept a normative definition.’ I disagree. Not only because, as Dworkin admits, “‘paternalism’ as used in ordinary contexts may be too amorphous for thinking about particular normative issues’, but because a concept that is mainly used to describe a normative position may as well be

---

⁶ On De Marneffe’s hybrid definition of paternalism a policy can only be paternalistic if benevolence is both justificationary and motivationally necessary. I take anti-paternalists to oppose beneficial limiting of liberty, regardless of motive.
normative. As far as I can tell, ‘paternalism’ has no distinct descriptive meaning in everyday conversation, nor in some specialized field or profession. The etymology of the word (Latin *pater* for father) supports characterization along the lines of treating someone as a (good) father would treat a child. In discussions of paternalism in moral and political philosophy, however, we are mainly interested in the benevolent limiting of liberty, not in other areas of parenthood. The main issue, furthermore, is the moral status of benevolent liberty-limiting, and especially the traditional liberal resistance to such liberty-limiting. Best then to define what exactly is resisted, instead of taking a detour around some non-normative definition.

Why prefer non-normative definitions? Donald Van De Veer (1986) argues: ‘If we wish to avoid begging the moral question (by simply assuming or supposing an act is wrong in labelling it “paternalistic”) we need to identify a morally neutral definition’ (pp. 16-17). This is not true. We can argue about whether or not that which is captured by the normative definition is indeed wrong, rather than just suppose that the labelling does the job. It seems that Dworkin and Van De Veer are interested in the concept of paternalism because of its alleged normative properties – its wrongness. They want their definitions to capture this moral controversy. However, for some reasons they do not want the moral controversy to be a defining criterion. I do not see why not. Van De Veer notes that there are many value-laden terms in our language, exemplifying with a long line of pejorative terms such as ‘bastard’ and ‘redneck’ (p. 16). However, there are also more theoretically interesting value-laden terms, such as ‘justice’ and ‘betrayal’. I do not know whether Dworkin would prefer to define those terms in non-normative terms or whether Van De Veer would find it question-begging not to do so. I propose, however, that such terms can only be fully understood by giving proper attention to their alleged normative properties.

What might make sense is to try to capture, somewhat more precisely and in less normative terms, what people who have taken an interest in paternalism have been interested in. A majority of the diverse treatments of paternalism over the past decades share a concern with actions that have all of the following four qualities (including those surveyed and also of importance Hodson 1977; Arneson 1980; Husak 1981; Dworkin 1983; Kleinig 1983; Arneson 1989; Archard 1994; Husak 2003; Arneson 2005, as well as many more specialized contributions in medical ethics and other areas of practical ethics):

1) It amounts to (or is believed or intended to amount to) involvement in some person’s life that can reasonably be interpreted as limiting or interfering with that person’s rights, liberty or self-determination.

2) It is not (or is not believed or intended to be) an immediate response to an informed, authentic and rational request by the person.

3) It is (or is believed or intended to be) to the benefit of the person.

4) Its moral status is not (or is not believed or intended to be) in any obvious way determined by other factors than the interests of the person.
This is the closest I will get to proposing my own non-normative characterization of paternalism. The first and second criteria ensure that the action is in some way problematic in relation to liberal values. The references to reasonable interpretation and obvious justification in criteria one and four, respectively, avoid dependence on normative elements while allowing for commonly accepted relevance boundaries. Criterion one excludes actions (such as asking for the time) that no one finds problematic. Criterion four excludes actions (such as stopping someone from blowing himself up in public; or spending a substantial part of the health budget on a program of extremely costly, forced plastic surgery) that most everyone agrees can be evaluated without considering the interests of the targeted person (‘no one’ and ‘everyone’ allows for interpretation). The element of subjectivity, as well as the disjunctive components and the inherent vagueness of many of the terms, make the criteria quite inclusive, while still providing some guidance as to in what way the problem (or family of problems) of paternalism is an ethical problem in its own right, distinguishable from more general questions of rightness.

4. REASONS AND THE GOOD

There is an action and a reason component to paternalism. Paper three is focused on scrutinizing accounts of the action component and paper two on the relationship between the two components. In this chapter I will discuss the reason component in somewhat greater detail than is done in any of the papers.

4.1 KNOWING BEST

It is very common to think that paternalism necessarily involves a presumption on the part of the paternalist to know better than the person targeted what is best for her. This is a mistake. For example, Richard Arneson states:

The essence of paternalism is overriding the individual’s own evaluation of where her own good lies (along with her decision as to the degree to which she will pursue that good by her choices rather than seek alternative goals). Restriction of people’s liberty intended to give effect to their current evaluations of where their own good lies and their own present will as to how far it should be pursued are not rightly deemed paternalistic, as many commentators have noted. (2005, p. 266)

I agree with the second sentence. Specifying the liberty value that anti-paternalists are interested in protecting is difficult (this is one of the main points of paper three). However, it seems that this value is not diminished if a (competent) person chooses to have her future options restricted. In other words, restricting a person’s options (her ‘liberty’) is immediately justified by her consent, and so consented-to option-restricting does not limit liberty in the relevant sense. We should notice, however, that it is certainly possible to help someone restrict her future options in accordance with her present will, without agreeing with her on what is her good (and to that extent ‘override’ her ‘own
evaluation of where her own good lies’). I may help Odysseus tie himself to the mast even though I think it harmless to jump ship when the Sirens sing, if tying him up promotes his good in some other way – such as by keeping him on a strict diet, something he himself in no way considers a good thing.

My objection is to the first sentence. I agree with the bit in parenthesis – paternalism must be unwelcome in some sense, it must conflict with a person’s decision or will. It is normally important that a person can decide freely how to pursue what she thinks is her good and paternalism is most typically a value conflict involving liberty in this sense. However, overriding a person’s decision for her good need not involve disagreeing with her on what is her good, since she need not act on her own view of her good (which is in fact rather straightforwardly entailed by the ‘to the degree’ and ‘how far’ bits in the quote). Consider the nicotine addict who believes that smoking is bad for her and yet wants to smoke. She may even believe that in the long run smoking is worse for her than the effects of temporary coercion, so that her all things considered judgment of what is conducive to her good favours coercively preventing her from smoking. Still, as long as she does not consent to being coerced, coercing her in conflict with her present will and for her good would arguably involve paternalism (this case is discussed in paper two and named ‘The Smoker’). One might avoid this conclusion by claiming that whenever we can contribute to what a person considers her good by limiting her liberty she is irrational and therefore doing so involves no paternalism. This strong claim seems to depend on the unreasonable assumption that paternalism is only relevant in relation to people who always maximise their own good (or more precisely who promote their good so efficiently that no one else can ever help promote it by the slightest limiting of their liberty).

Furthermore, it is not problematic to disagree with a person on what is her good, as long as one does not limit her liberty. In fact, it is not even problematic to actively promote, for her good, what she does not consider her good, as long as doing so is in accordance with her present will. For example, it involves no paternalism if I help a (competent) person, in response to her request, to destroy what she but not I think is her own good, in order to promote what I but not she thinks is her good. Her motive might be to promote what she thinks is some greater cause. Assume, more concretely, that Anne thinks that her good is best promoted by her having as much money as possible. She asks Betty for help in giving away money to charity in order to promote the greater good of alleviating suffering. Betty does not believe in alleviating suffering but believes that it is better for Anne to give away some of her money (since psychological research shows that practicing altruism makes us happier) and helps her for this reason. Betty is helping to give effect to Anne’s decision, but disagrees with her on what is her good.

All this shows that what is morally problematic, and therefore what involves paternalism, is the overriding of decisions, not the overriding of evaluations of good. From the perspective of a presumptive good-promoter, there are three important factors in Arneson’s quote – whether we agree with the person on what is her good, whether we restrict her options, and whether we act in accordance with her present will. These three factors can be combined in eight ways and all combinations are feasible. I have agreed
with Arneson that it involves no paternalism to restrict a person’s options in accordance with her will. I have pointed out that this is true regardless of whether or not we agree with her on what is her good (Odysseus on a diet). I have argued against Arneson that it does involve paternalism to restrict a person’s options against her will even though we agree with her on what is her good (The Smoker). If we do not agree on what is her good, restricting her options against her will is of course a typical instance of paternalism. I have also argued that it does not involve paternalism to disagree with a person on what is her good as long as we do not restrict her options (Anne and Betty). This is true regardless of whether or not the person wants us to restrict her options, as long as we have no duty to do so. If we do have such a duty, not restricting her options, for her good, may involve paternalism. Two non-problematic possibilities remain – we may agree with a person on what is her good and not restrict her options. The person may want us to restrict her options, but if there is no disagreement on what is her good, choosing not to will not involve paternalism (we may just prefer to do something else).

Arneson is an influential interpreter of anti-paternalism but not (any longer) an advocate. In an earlier article (1989), he makes the adequate and interesting observation that the insistence on respecting people’s own views of their good can actually undermine anti-paternalism: ‘From a single-party welfarist consequentialist standpoint, the insistence that “autonomy trumps” is just another species of perfectionist imposition of values on the agent in defiance of the agent’s own considered evaluation.’ (p. 435) This is true. However, I prefer, more generously, to assume that the anti-paternalist tendency to insist that people be allowed to define their own good is a confusion.

4.2 Acting on Judgements

In order for anti-paternalism to be of any interest, the doctrine must operate on reasons that would be relevant and valid was it not for the doctrine. A claim that we have no reason to ever help others does not support anti-paternalism, but rather makes it redundant. To reject anti-paternalism is not to make a claim about what are in general good reasons for action, nor about what is in fact good for a person. It is to claim that reasons that concern a person’s good are as appropriate for actions that limit her liberty as they are for any other actions.

Good reasons are most obviously such that refer to the actual effects of an action. Paternalism is perhaps most appealing when limiting a person’s liberty does in fact promote her good. Indeed, it is a common position that only facts provide reasons (e.g. Parfit manuscript, section 1.1; Raz 1990. p. 18). However, since what is in fact good is a matter of philosophical and often empirical dispute, and in any case very hard to know, we may want to distinguish between acting on such beliefs as are well supported and such that are not (this point is also made in footnote 11 of the second paper). I therefore propose that we accept as reasons for action also judgements to the effect that something of value will be affected.

We may distinguish a number of different judges or ‘authorities’, who can make judgements on what is good for a person. Admittedly, a theory of the good can identify the good with what some authority judges to be the good. In such a case, the fact
of the matter and the judgment of the authority coincide (if the theory is correct). It may still be worthwhile to distinguish the two, since people may agree on the normative status of the judgment of the authority while disagreeing on what is the correct theory of the good. An authority may make judgements both with regard to what is the correct theory, and with regard to what the correct theory entails in a certain case (e.g. what makes a certain person happy in certain circumstances). The two most salient authorities in the context of paternalism are the person acting and the person(s) affected by the action, though possible authorities also include third parties, such as experts, people who know the affected person well, and possibly even general opinion. More complex kinds of judgment are also possible and important, especially the acting person’s judgment on the affected person’s judgment on what is for her good. We may also introduce idealized or qualified versions of these authorities such as the person acting when reasonable and informed, or the person affected when expressing her settled values.

From a liberal perspective, acting toward someone based on one’s own judgment on what is for her good may seem dubious, while acting toward someone based on her judgment may seem unproblematic. Things are, however, not that simple. As argued in the previous section, people may well resist an action even if this action promotes their good according to their own judgment. Also, acting on one’s judgment on what is for the good of others may be very reasonable and even trivially non-problematic. We may certainly act on our own judgment of what is good for others if this is within our rights, as for example when we make a decision concerning what to give a friend for her birthday based on what we think would make her most happy (rather than what we think that she thinks would make her happy). Often, we do not know (and cannot in any reasonably efficient way find out) what is truly for the good of others, nor what they judge to be for their good. Equally often we cannot help but to significantly affect others one way or other (by our action or inaction). It would then seem preferable to act on one’s own judgment, rather than to simply refrain from reasoning, or to disregard the good of others. We may of course try to guess what others would judge to be in their interest, but sometimes we have very little information on which to base such guesses.

I propose that anti-paternalism is equally reasonable or unreasonable for any kind of benevolence, regardless of whether the reasons are based on facts or on the judgment of various authorities. For those who disagree, the doctrine could be restricted to reasons provided by certain kinds of judgments. However, if the resistance to paternalism is based on the idea that we may not limit a person’s liberty for her good, then it seems that both the fact of the matter, the affecting person’s beliefs, third party beliefs, and the affected person’s beliefs should be included. Restrictions of the doctrine that exclude or are limited to some one of these different sources of reasons must explain why other reasons that similarly concern personal good are (not) included. If one rejects anti-paternalism, restricted versions are more reasonable only in the sense that they prohibit a smaller part of what should not be prohibited at all. The top candidate for exclusion is probably the judgment of the affected person; that a person judges an action to be in her own interest should reasonably provide some kind of reason for that action.
even if it does limit her liberty. If limiting liberty for such reasons is deemed acceptable, however, it becomes less clear why limiting liberty for other reasons, that similarly refers to a person’s good, is not.

Our intuitions concerning the reasonableness of different kinds of reasons for interference admittedly differ. Interfering based on one’s own judgment on what is good for another seems perhaps most problematic, while interfering with someone based on her judgment seems least problematic. There are important differences between these kinds of reasons. It is probably true that in general, a person is more likely than someone affecting her to know what is in fact good for her. It may also be true that in general, interference based on the affected person’s own judgment of her good is less severe than interference based on the affecting person’s judgment of her good – in the former case the person interfered with can at least identify with the motive or aim of the action. This does not mean, however, that reasons based on the affecting person’s judgment are not valid. It may also be that we mistakenly assume that limiting a person’s liberty in order to promote what she judges to be her good is not really limiting her liberty at all, since in a way she wants the action to take place. As I argued above, this is not so, since our decisions and our views of our good can diverge. However, the impression that there is this connection may still bias our intuitions.

4.3 Health Promotion and Moralism

In arguing against anti-paternalism I in no way intend to argue against anti-moralism. However, C.L. Ten (1971) and Heta Häyry (1992) have argued that these isms cannot be properly distinguished by the values or goods they aim to promote. Even limiting liberty in order to avoid physical harm, a rather straightforward personal good, involves a moral commitment, they claim, which leaves paternalism indistinguishable from legal moralism. Häyry states: ‘If legislators “know” – that is, are licensed to define – for ordinary people what is best for the individuals’ own physical good, then they presumably also “know” – are licensed to define – what is best for their moral good.’ (p. 196) It is not clear why Häyry takes this position. She refers to arguments by Devlin (1965) to the effect that defence of physical paternalism is necessarily founded on ‘rationally untenable distinctions’ (Häyry, p. 196). However, Hart’s (1963, pp. 30-32) distinction between (justifiable) paternalism and (unjustifiable) moralism, further developed by Feinberg (e.g. 1984, pp. 12-13), is quite clear. Legal moralism involves preventing the inherently immoral, regardless of whether it harms anyone. Paternalism involves preventing harm (or promoting good) to individuals. While it should be recognized that attributing negative value to physical harm is taking a moral stance, to attribute negative value to certain behaviour or lifestyles as such is another and distinct stance.

7 Ten and Häyry both hold that the distinction between paternalism and moralism can be upheld by restricting paternalism to interference with acts that are insufficiently voluntary. However, paternalism need not be so restricted and the point here is that Ten and Häyry claim that no distinction can be upheld between the kinds of values that are promoted by the respective isms.
Ten argues that what is valued by the typical paternalist is not absence of physical harm as such but rather the absence of immoral physical harm: ‘Intervention is only [thought] justified when the harm is caused by the commission of an immoral act.’ (p. 57) While surgical operations are generally accepted, sterilization is not, though both involve physical harm: ‘The difference between the two cases seems to be based on the fact that surgical operations are not generally regarded as immoral, whereas sterilization is so regarded by some religious groups.’ (Ibid.) Ten’s argument, then, seems to be that if the paternalist was really concerned to protect physical health, she would go after surgical operations as well. In fact, however she only favours intervention when physical harm is caused by something she finds immoral, such as sterilization.

In reproach, why should we think that surgical operations are accepted because they are morally impeccable, and not because they are conducive to physical health? It is true that an operation can be dived into steps which taken in isolation can be described as harmful (cutting someone’s stomach open etc.), but such arbitrary division can prove all sorts of acts to contain harmful or in other ways counter-intuitive parts. Surgical operations are performed with the aim of contributing to physical health and they most often do. To the extent that operations are questioned on normative grounds (such as with sterilization), this is exactly because they are atypical in that they might be seen to destroy rather than contribute to physical health. Immoral actions of course tend to be more susceptible to paternalistic intervention because they tend to be harmful. This does not show that it is their immoral nature that is the basis for intervention, rather than their harmfulness.

5. VALUES, REASONS AND PRINCIPLES

To my mind the two most important notions in moral philosophy are value and reason. What has value? What do we have reason to do? These seem to me the fundamental questions of morality and more generally of practical reasoning. Complex empirical relationships between actions and outcomes ensure that we cannot move easily from the first to the second. However, when something of value (i.e. something important) is affected by an action, this gives us a reason of some strength to perform or avoid the action. In other words, if we can effectively protect or promote (or diminish) some value, we always have reason (not) to do so. To oppose this common sense assumption is to muddle the distinction between considerations that are in a basic sense relevant, such as that a person will be harmed, and considerations that are utterly irrelevant, such as that a person will sneeze. Granted this assumption, many doctrines in moral philosophy can only be understood as influence-regulating principles that prevent reasons from having influence on the moral status of actions according to their strength. I propose that moral inquiry would benefit from rejecting all such principles and focusing on the two fundamental questions.

---

Perhaps these questions should be complemented with one or other agent-focused question, such as: What kind of person do I want to be? Nowhere in this thesis are such virtue ethical issues considered, unfortunately.
5.1 Reasons without influence

John Broome (2004) defines reasons in terms of their role in explanations of ought facts. Modifying his former view to allow for deontic principles, Broome states that these explanations may be merely ‘potential’, not ‘actual’ (pp. 40-41). In other words, reasons are entities that would explain ought facts if it were not for deontic principles. I prefer to say that what reasons do is that they influence the moral status of actions. That is, they do so unless their influence is blocked, typically by deontic principles.9

It is a rather common idea in moral philosophy that certain considerations should be disregarded on normative grounds and that certain other considerations have a special status that goes beyond their relative strength as reasons. Indeed, one could argue that this is the essence of a deontological approach to ethics. For example, Jonathan Dancy (1993), building on and adapting from John McDowell, argues that reasons can be ‘silenced’ by the presence of other reasons, or by other circumstances (pp. 47-58).10

Interestingly, for Dancy silencing does not occur in accordance with principles, but is particular to each situation. Apparently at the opposite end of the particularism-principalism debate, Thomas Scanlon (1998) argues concerning reasons in general that certain ‘considerations’ may make it the case that a reason is not ‘relevant’, which is distinguished from reasons being ‘outweighed without losing their force or status as reasons.’ (pp. 50-51) More specifically on moral justification, “being moral” involves seeing certain considerations as providing no justification for action in some situations even though they involve elements which, in other contexts, would be relevant.’ (p. 156) Or, more briefly, ‘certain reasons for action [...] are morally inadmissible.’ (p. 201)

Scanlon says that considerations lose their status as reasons. Dancy at one point says that a reason that is silenced is in fact ‘no reason at all’ (p. 60).11 It is common to talk this way, to say that something that would influence the moral status of actions were it not for some normative blocking is not a reason. It is also common to say with Scanlon and Dancy that these things are not relevant. This is confusing. There is an interesting distinction to be made between such effects of actions as people being harmed and people sneezing. The distinction is due to the great difference in the importance or value of these effects. There is also an interesting distinction to be made between things that are effects of an action and things that are not. If we say that deontic principles or particularist ‘deontic circumstances’ make reasons irrelevant non-reasons, we would be hard pressed to invent a new vocabulary to express these important distinctions. Better to say that these deontic entities block reasons from having influence on the moral status

---

9 I prefer this terminology mainly because I want to understand deontic principles as regulating the influence of reasons, rather than dispensing reasons to the realm of the merely potential. However, this could perhaps be accomplished by staying with Broome’s terminology and making room for deontic principles in the explanation of ought facts. I also prefer my terminology because I prefer to avoid talk of facts in morals, and because I am not as comfortable as Broome is with the primitive notion of explanation.

10 In Ethics Without Principles (2004, p. 41) Dancy talks of ‘disablers’ of reasons that seem to have the same silencing function. The later work, however, includes few moral examples and those given seem motivated by practical constraints.

11 Dancy’s position has not changed in its fundamentals. In his entry on ‘Moral Particularism’ to The Stanford Encyclopedia of Philosophy (2008), he states that ‘what is a reason in one case may be no reason at all in another’.
of actions, though effects of an action that concern things of value are in a very basic sense relevant and when we can produce such effects we have some reason to do so. This is only a terminological matter, but since the notion of a reason is fundamental it is a relatively important one.

Dancy (1993) argues in support of moral particularism that the fact that an action causes pleasure normally counts in favour of performing it, but not if the pleasure is malicious; that the fact that we have borrowed a book normally counts in favour of returning it, but not if the book was stolen from the library; and that the fact that an action is a lie normally counts against performing it, but counts in favour of performing it if lies are required as part of a game we play (pp. 56, 60-61). It is not clear if lack of value or normative regulation of the influence of reasons is doing the work here. What is clear is that if we should disagree about these claims, it would be informative to learn whether the disagreement concerns value or influence. Myself I would say that it has no value to avoid lying in lying games, but that it has some value to return borrowed books, even to thieves, though perhaps more value to return them to their rightful owner. I would therefore be curious to know why exactly Dancy thinks that there is no valid reason to return the book.

5.2 Anti-paternalism as an Influence-Regulating Principle

There are many forms of anti-paternalism. Some are straightforwardly consequentialist, pointing to the likely counter-productive effects of certain institutional arrangements. Others are more sophisticatedly consequentialist, questioning the value of interference-generated personal good (such versions are briefly considered in paper two). This thesis is mainly concerned with the more deontic, more principled, normative anti-paternalism that I find in Mill’s liberty principle and Feinberg’s soft paternalism. Such anti-paternalism can most favourably be understood as the position that reasons that concern a person’s good are invalid for actions or effects that limit her liberty. That a reason for an action is invalid means, on my understanding, that the reason does not influence the determination of the moral status of the action. Invalidation of reasons is one form of influence-regulation. Other forms include side constraints and lexical orderings of reasons. The distinctions between these forms of regulation are discussed in paper two. That we sometimes need to consider not only actions but also various effects of actions is explained in paper one.

5.3 Against Anti-paternalism

The case against normative anti-paternalism is made most thoroughly in paper two and three of this thesis. One way to capture the general problem with anti-paternalism is to point to the difficulty in specifying the notion of limiting liberty in such a way that any reason that concerns a person’s good is invalid for any action that limits her liberty. A person’s good includes such major goods as survival and preservation of autonomy. That there should be some specification of limiting liberty such that these major goods do not provide any valid reasons for any action that limits liberty is simply unlikely. There are of course very narrow specifications that would do the job, but not without making anti-
INTRODUCTION

paternalism redundant (since the allegedly invalid reasons are always outweighed anyway). This difficulty implies at least three arguments against anti-paternalism.

First, there is no action type such that the fact that an action of that type promotes personal good is always morally problematic (this is the argument of paper two). Second, there are cases where it is obviously justified to limit a person’s liberty (partly) for her own good, including cases where the person is perfectly informed and rational and all agree on the relative strength of the relevant reasons as far as they concern her interests (this argument is developed in paper three). Third, drawing a line for when reasons that concern a person’s good are valid and when they are invalid gives rise to peculiar jumps in justifiability, such that one action is overwhelmingly justified while another, very similar action, is overwhelmingly unjustified (this argument too is developed in paper three).

Admittedly, we may have a special need to be suspicious of reasons that refer to our good. Such reasons are more likely than other reasons to be invoked in rhetorical attempts to persuade us: ‘It will benefit you too.’ Such reasons are also commonly invoked for actions that mould us in the interests of others, especially when we are young: ‘It is just for your own good.’ We do not have the same need to develop a mistrustful attitude towards reasons that refer to the interests of others. This may possibly explain intuitions that tend towards anti-paternalism. It does not, of course, provide an argument for the doctrine.

On a more general level, I find deeply problematic the idea that reasons to effectively promote something of value can be invalid. Sometimes we cannot consider all relevant reasons because of practical constraints, including limited time, information and processing capacity. This does not mean, however, that they should be disregarded on normative grounds. I propose that secular morality is based on judgements of value. For any moral demand, we can ask whether it has value that we meet it. This is not simply to endorse consequentialism, or if it is this is consequentialism in a very broad sense (Cf. Sen 1982). It may have value that we respect rights, do our duty or are virtuous. More to the point, that I respect a certain right or fulfil a certain duty or attain or display a certain virtue may have value regardless of consequences. However, it cannot be morally prior to some other moral imperative in any other way than having more value. Values will conflict of course, and one of the main tasks of moral philosophy is to investigate the nature of various values and how values relate to each other, in order to understand and if possible resolve such conflicts.

I boldly propose that any non-redundant doctrine of invalidation will confront problems analogous to those that face anti-paternalism – difficulty specifying the relevant action type, wrong answers to moral questions in certain cases, and peculiar jumps in justifiability. Furthermore, if a reason has the special power to invalidate other reasons, this has the unfortunate consequence of putting within brackets the fundamental question of the relative strength of different reasons (this is a further argument against anti-paternalism developed in paper two).

In addition to the substantial arguments against anti-paternalism developed in the papers, there are also methodological reasons to prefer value conflicts over influence-
regulating principles. Investigation of what things have value and how they generate reasons for action can contribute to our moral outlook regardless of what exact relative value we attribute to those things or what other things we think have value. This is true both on an abstract and on a very concrete level. Say that you and I argue about whether or not to recycle our household trash. You explicate the value of a clean environment and I the value of a clean kitchen. You may point to the deep feeling of harmony and connectedness that can only be experienced immersed in untouched wilderness, while I may point to the elegance and convenience of putting all the trash in one single container. Our respective arguments can contribute to each our worldviews even if we disagree on the relative importance of these values.

In contrast, if you say that we should recycle as a matter of principle (perhaps because reasons that only concern convenience or aesthetics are invalid) and I disagree, then this principle may tell me nothing at all. More generally, if there is a principled rule that we think is incorrect in one instance, then the rule is falsified and it is unclear how we should relate to it – whether we should, for example, amend it or give it up completely. The rule itself gives no guidance in such matters. This is especially problematic since moral principles always end up being very complex, in order to avoid unreasonable prescriptions (one can of course bite the bullet and for example hold with Kant that we should simply never lie, but to avoid giving the wrong answer to moral questions, more complex treatments are needed, such as Korsgaard’s (1986) qualified defence of this Kantian principle).

If following a rule is not a strict deontic requirement but simply something that has value, then the rule can find its place among other values. Its value can be contrasted with that of other rules and other things. Further specification of what exactly is valuable in following the rule will tell us whether it has value to follow it as often as possible or if it only strict compliance has value, or whatever.

The methodological case against principles is even stronger in politics. As individuals it has value that we interact with each other on equal terms. If the value of equal standing is very great, it may entail something akin to deontic principles regulating how we may behave towards each other and what considerations we should and should not take into account. However, society is not a moral agent that needs to foster a sense of its own proper footing. While it is certainly of value that people are not coerced, for example, it is hardly of further value that certain institutions do not coerce. It may be a bad that a person is an agent of coercion, independently of the effect that someone is coerced. Even failed attempts to do bad may possibly taint a person’s character. However, institutions have no character. It is not bad in itself that the state is involved in coercion. If, implausibly, designing certain state institutions to attempt to coerce people would have some positive side effect, yet not have the effect that people were coerced, nor the agents of coercion, and if nothing else of value was affected, then the quasi-coercive nature of these institutions would not be a reason against this order. Individual deontic restrictions are, from the perspective of political philosophy, like practical constraints, part of the background against which institutions are designed.
My claim that principles are especially ill suited for politics may be surprising, given that they are so common in that area. This is partly because I do not accept the popular premise that state actions stand in need of some special justification (legitimacy) that is not required of private actions. I take it that, ceteris paribus, coercion by the state is not worse than coercion by some private party (Cf. Pettit 2002). Partly, the explanation is that I am abstracting from practical constraints and empirical circumstances such as the tendency of power to corrupt etc.

5.4 Constant Overriding

If, for a certain type of action, con reasons of a certain type are always stronger than any possible pro reasons of a certain type, the con reason will always override these pro reasons. Such regularities can perhaps be codified into principles. For example, for anti-paternalist constant overriding, the con reasons would concern liberty and the pro reasons would concern personal good. The effect would be very similar to that of invalidation – paternalism would be prima facie wrong. If my arguments against invalidation above and in the papers are convincing, a principle of constant overriding may seem an obvious alternative.

There are two important weaknesses with a principle of constant overriding. First, such a principle has nothing to say about cases where we have both personal good reasons (etc.) and other reasons for limiting liberty (etc.). All the principle says in such cases is that there are con reasons stronger than the personal good pro reasons. We do not know how much stronger and we do not know what will be the balance of reasons when we consider also other pro reasons. This weakness in itself is severe enough to make constant overriding a rather uninteresting interpretation of anti-paternalism, or principle more generally.

Second, it is simply difficult to see how one kind of reason can always be stronger than another kind, given that they are both valid. I may hold, for example, that in general liberty is more important than life and health, but how can it be that even the most trivial limiting of liberty will override the most enormous damage to health? If I can save you from certain death by pushing you out of the way of an approaching bus, though this will limit your liberty to move about without being pushed, how can the former reason be overridden by the latter? If your standing in the way of the bus is a deliberate suicide attempt, my pushing you might interfere with your right to decide over your own life and death. This right may be a particularly important aspect of liberty. However, being free to move around without being pushed is arguably another aspect of liberty and the one the example is concerned with. Assume therefore that your standing in the way is not a suicide attempt.

One might try to define the overriding reason type in such a way that reasons of this type are always very strong. The reason we have not to push people around is perhaps not a significant liberty reason, while such things as the freedom to vote or to marry do provide significant liberty reasons against interference. This strategy of course makes for a rather narrow doctrine, leaving many liberty-limiting actions outside of its domain. Similarly, other narrowly defined values could give rise to other prima facie
wrongs. One example may be the destruction of one culture for the expansion and material gain of another – call this colonialism. If colonialism is prima facie wrong this is not because of some special relationship between the value of the non-destruction of a culture and the value of expansion and material gain, but simply because anyone’s reasons against destroying a culture are always stronger than her reasons for achieving expansion and material gain.

Even more narrow principles are hard to formulate in ways that make them reasonable, unless phrased in terms of great disvalues such as the destruction of a culture. Liberties of kinds that are normally very important may on occasion be relatively unimportant. Suppose that you lack money for urgently needed medical attention and you plan to marry a certain person on a whim. You reason that you won’t live long anyway because of your predicament. Now someone (a demon perhaps) pays me a large amount of money to stop you from marrying this person. I do so and use the money to save your life. This seems to me justified or at least close enough to being justified to raise doubts that non-limiting of the freedom to marry always overrides health concerns. There is of course no end to how narrowly the principle may be defined. At the extreme, anti-paternalism would be the principle that personal good reasons are always overridden by our reasons not to eradicate every single trace of self-determination from a person’s life. Somewhere short of this point anti-paternalism will no longer merit the name of principle, being simply a (short) list of certain reasons that override certain other reasons.

5.5 Liberalism without anti-paternalism

Anti-paternalism is an integral part of the liberal tradition. However, one may well affirm the values of liberty and autonomy without embracing anti-paternalism. For one thing, it is not obvious that these values are best promoted or protected by non-interference. My reason for limiting your liberty may well be that if I succeed, you will have more valuable options to choose from in the future and be a more self-reliant person (or whatever similar properties are taken to embody liberty and autonomy). This point has been made again and again in the literature (since it was made by Dworkin 1972). Paternalism does not entail a net loss of liberty over time; it merely entails a limiting of liberty in some more specific sense, a sense that is not easily distinguished (as argued in paper three).

Let us assume for a moment that limiting liberty can be clearly distinguished so that it implies a loss of value. Now, why should we not simply compare this loss or disvalue to the positive value that will result from limiting liberty? Stopping you from taking great risks may involve a disrespect of your autonomous will and your decision to take those risks. This is a bad thing. That you stay alive and healthy, on the other hand (and that your long term autonomy is thereby preserved) is arguably a good thing. How the situation should be evaluated apparently depends on such things as the significance the choice of taking the risks has for you, how my stopping you will affect your autonomy and sense of self-respect, what your life will be like if I save you etcetera. These specific matters may indeed depend on more general considerations concerning such things as the value of respect for autonomous choice and the value of human life.
However, there is no apparent need for the liberal to simply disregard some aspects of the situation.

It may seem that I am unnecessarily creating a theoretical problem. In practice, there must be some point where we should let people take their own risks. But this is not the issue. The issue is what grounds there are for allowing people to take their own risks. Why should people sometimes be free to take their own risks? A plausible answer is that they should be so when the value of (some kind of) liberty outweighs the ensuing disvalue of harm or risk of harm. Anti-paternalism, however, implies that in cases of conflict between values, we should first consider whether some person's liberty is limited (in some sense). If it is, we must remove reasons that concern the good of that person from consideration. Only after this weeding out of personal good reasons may we move on to consider the reasons that are still valid, in order to reach an all things considered judgment. The need to make practical judgments does not in itself call for weeding out reasons in this way.

Absolute prohibitions may be appropriate as rules of thumb, as constitutional restrictions on government, as laws or as codes of conduct. If they are, however, this is because people cannot be trusted to weigh values properly, so that the consequences of insisting on such prohibitions are better than the alternatives, in terms of the overall promotion and protection of values. Our general inability to properly consider reasons should not, however, circumcise theoretical discussions of morality, whether it concerns private action or public policy.

5.6 **More General Theories**

Anti-paternalism, as understood here, can be a consequence of various more general normative positions. State anti-paternalism could be a requirement of state neutrality or anti-perfectionism in the sense that no government action should be justified with reference to the good of anyone. Another form of anti-paternalism follows from libertarianism, with its opposition to infringement of ownership rights, whether benevolent or not (though it is presumably consistent with libertarianism for the owner of a road to require that people only drive on it with seatbelts, and helmets, and gum shields, or for the owner of an apartment complex to require that the tenants submit to drug testing, or adopt strict diets and exercise regimes). Another ground for anti-paternalism is a Kantianism according to which treating others as ends requires that benevolence should only manifest itself as aid to people in promoting their ends as defined by themselves, assuming that duties to oneself may not be enforced by others (e.g. Baron 1997, pp. 13-15) (the central notion here is ‘aid’ rather than ‘as defined by themselves’, since the latter is consistent with some forms of paternalism, as noted above).

With the exception of some exotic versions, consequentialism is incompatible with anti-paternalism on the level of rightness, since its central claim is that all consequences of an action that have value should be considered. However, consequentialism can incorporate anti-paternalism as a rule of thumb if the consideration of some reasons for some actions is typically counter-productive or simply not worth the
Anti-paternalism and Public Health Policy

Effort. More univocally, anti-paternalism is incompatible with many forms of communitarianism. For example, a person’s identity may be understood to necessarily involve certain ideas of the good, implying that it makes little sense to give priority to non-interference over promoting those goods (Taylor 1989, pp. 26-27).

In the literature commented on in the first two chapters, the detailed discussion on the proper definition and moral status of paternalism has to some extent been kept separate from more general ideas on morality and politics (though commonsense consequentialist arguments are common). Perhaps this is unavoidable in dealing with any more specific moral problem. In our case the tendency is strengthened by the fact that important figures such as Mill (in On Liberty) and Feinberg do not bother much with the most abstract level of rightness, but are more concerned with values or reasons on an intermediate level – between final value and purely empirical considerations – and the principles that allegedly govern them. Mill of course was a utilitarian and there has been much discussion on how his anti-paternalism fits into his more general understanding of morality. I tend to side with C.L. Ten (1980) in thinking that Mill’s liberalism and his utilitarianism simply do not add up, but that he took them both very seriously. Feinberg (1984) explicitly avoids questions of final value, calling his four volume work on the moral limits of the criminal law an ‘extended essay in applied moral philosophy’ (p. 4).

This thesis follows the custom of keeping some distance between the local issue of paternalism and global issues of morality and politics. On its most theoretical level, the thesis concerns the proper understanding of paternalism and the reasonableness of rejecting it on normative grounds. More general theories are not considered in the papers. Assume that anti-paternalism follows from some more general theory. This is how I would understand the relationship between the arguments against anti-paternalism presented here and such a theory: 1) If there are decisive arguments for the theory independently of the reasonableness of anti-paternalism, then my arguments are in a sense irrelevant, though they may still provide some insight and may contribute to undermine the decisiveness of the decisive arguments in some (future) contexts. 2) If the theory gains support from the alleged reasonableness of anti-paternalism, my arguments, to the extent that they are sound, weaken this support, possibly undermining the theory. To exemplify: It could be argued against my approach that everyone has an absolute right to control her own body and property, and so that it can never be justified to limit that control for whatever reason other than to protect the like rights of others. If the premise of this claim is some idea of ownership that is independent of intuitions or arguments to the effect that people should not be interfered with for their own good, then the claim must be evaluated according to whatever standard is appropriate and if sound will imply anti-paternalism. However, if the idea of absolute ownership is to some extent based on anti-paternalist intuitions and arguments, then to that extent it is undermined by the arguments presented here, if sound.
6. **Liberty or Health?**

With its focus on arguing against anti-paternalism, this thesis is largely critical, though in interpreting paternalism and anti-paternalism, and in proposing aggregation rules for individual into group consent, as well as in taking a stand on certain public health policies, it is more constructive. In any case, something more positive should be said about the central values of health and liberty. This last proper chapter of the introduction, therefore, contains some rather undeveloped thoughts on the nature of these values. In connection, I present some views on libertarian paternalism, which has received so much attention lately.

In essence, I propose that there are many kinds of liberty and of health and that they are all important. As with the concept of paternalism, I think that interesting explications of these notions are generally normative. I follow Rawls (2001) in preferring an account of liberty where ‘no priority is assigned to liberty as such, as if the exercise of something called ”liberty” had a preeminent value and were the main, if not the sole, end of political and social justice.’ (p. 44) The liberty values I shall soon identify, however, are rather more abstract and general than Rawls list of basic liberties. They are more on the level of abstraction of Berlin’s (2002 [1969]) concepts of positive and negative liberty, of which he states: ‘Both are ends in themselves’ (p. 42).

6.1 **Health**

In public health policy, what the government and public health authorities are aiming to promote (if uncorrupted) is not people’s good per se, but their health. Health may be defined biomedically as the absence of disease and infirmity. This is the common definition in medical practice. The biomedical account could be complemented with the idea of survival and reproduction as basic functions of organisms (Boorse 1975). Alternatively, health may be defined biopsychosocially, which is common in theoretical contexts. The constitution of the World Health Organization (1946) famously states that health is ‘a state of complete physical, mental and social well-being’. Several recent definitions of health aim to avoid the somewhat utopian character of the WHO definition and to shift focus from outcome to opportunity, by defining health in terms of potential, ability, or resources, rather than well-being (Bircher 2005; Law & Widdows 2007). On these wide definitions of health, the distinction between health and good more generally is not sharp, though such things as finances and achievement that could be considered part of the good are presumably excluded.

Health is obviously a contested concept and the debate over its nature is in part normative – what kind of well-being or functioning, if any, has value for all people. This is an important question which does not necessarily have only one answer. There could be several concepts of health which all have value for all, or most, people. I believe that both well-being and potential have value, for example, and perhaps even normal functioning independently of the first two. In any given context, health can quite properly be as broad a concept as is coherent with some particular policy promoting health so defined, as long as it captures something of value.
6.2 Liberty

Anti-paternalists seem most eager to protect the freedom of the capable (informed and rational) to do what they want to do. They tend to think that this value must be balanced against other values, including the freedom of other capable people to do what they want and the prevention of harms to others, but that it trumps (other) personal goods or values in the lives of the capable. While I agree that the freedom of the capable is a value, I do not believe in trumping, as should by now be clear. I also believe that the freedom of the capable is not a value in itself but that it can be derived from two other values. First, it has value that people are free to do what they want to do, here and now. This is important for anyone, also for less capable judgers such as minors, the ignorant, the immature, the intoxicated and the emotionally upset. It is important even for small children and the outright insane, at least in some areas of their lives. Second, it also has value that people preserve their integrity or coherence as they exist over time. In the interest of this latter value, what a person wants to do, here and now, can sometimes be overridden by consideration of what she would have wanted to do if she had been more sober, emotionally stable, rational, mature, informed etc. These two values will of course conflict more often for less capable people. Note that the value of integrity can motivate overriding the present will of a person, regardless of whether or not harm threatens. We may have reason to stop a person from making a fool of herself in public, simply because this would radically change the impression other people have of her and so how they relate to her and so her existence as a social being. This may be true even if the change would not be harmful, but rather good for her, because, for example, it would cause her to be more relaxed, to ‘lighten up’.

The freedom to do what one wants to do is not, on my understanding, restricted to those things that are under one’s legitimate control. It has some value for me to use your property or to decide what you should do. *Ceteris paribus*, it is better that I can live in your house and drive your car. If you are away and don’t mind me using your things (with or without your permission), there are no reasons against my doing so and some for. However, quite often (though not as often as most would expect) it is more important that people have long term control over certain resources (their bodies, property and belongings), in order that they be able to shape their life according to plans and goals, and perhaps in order to stimulate economic activity and development.

Identifying as a value this freedom to do what one wants to do, quite generally, avoids the need to attribute value to self-determination in the sense of freedom to do what one wants to do in some more limited sphere of one’s life (the self determining the *self*). What each person should have control over, all things considered, depends on some balancing of this general liberty value as it is manifested in each person’s life, and against other values. I am aware that many have held that freedom of this general form can have no value, but I disagree. However, the value of some options is negligible and that of

---

12 This bias in value is distinct from the important point made by Arneson (1989; 2005, section V) that anti-paternalism has unfortunate distributional effects since bad choosers have more to gain from paternalism than good choosers.
INTRODUCTION

others great. Therefore, it may make sense to distinguish as independent values, in a Rawlsian manner, doing what one wants in particular areas (worship, movement).

The anti-paternalist insistence on non-interference naturally invokes the value of negative liberty – freedom from constraints. However, in explaining why we are nevertheless justified in interfering with people who act insufficiently voluntarily, anti-paternalists tend to invoke the hypothetical, more informed and rational self. Since they oppose benevolent limiting of liberty on principle, they tend to argue that restraining a person acting insufficiently voluntarily is not really restraining her at all. Indeed, this is ‘no more illiberal than interference to prevent him from harming or offending an unwilling second party.’ (Feinberg 1986, p. 12) This amounts to an appeal to that most suspicious form of positive liberty – the real will or real self, quite separate from the will of the actual person. Berlin (2002 [1969]) rightly warned against this ‘monstrous impersonation, which consists in equating what X would choose if he were something he is not, or at least not yet, with what X actually seeks and chooses’ (p. 180). It is tempting to reject such tendencies completely and attribute value only to the will of the present, actual self. However, there is no escaping considering the person as a more or less abstract entity existing over time. For how else would we attribute some value to treating people in sleep or unconsciousness according to their presumed wishes, wishes that they may not explicitly have formulated but that are easily inferred from their earlier wishes, choices and behaviour. It is partly to account for such situations that I suggest that integrity or coherence over time is a value. If this value is acknowledged, we need never appeal to the will of the person’s more informed and rational self (if we still do that self will often sanction the same kind of involvement).

If we shall respect and preserve the integrity of a person, we better think hard on what this is exactly. There is an obvious risk that we take a person’s everyday social persona to be her core self. Are people most truly such as they are most often? Or are their core selves better revealed in some circumstances than others? These questions seem to invite an outside, second person view of who a person most truly is, which may of course be influenced by normative assumptions of what aspects of the person are most valuable. One option is to go with a subjectivist account, according to which what is most central to any particular individual’s self is what she herself identifies with. However, here we have the definiendum in the definiens. So we would have to specify what aspect of her it is that identifies who she most truly is. One way or other, we need some more objective account to answer what it means to preserve integrity.

I have focused on those liberty values that I take to be most central to paternalism. This far from exhausts the complete list. There is of course a rich literature on liberty and freedom that may inform our investigation of liberty-values. Some focus on available options or capabilities (Raz 1986; Sen 1992). Others focus on what causes us to have a certain bundle of options or capabilities, making distinctions between options being constrained by non-human nature or by human action, perhaps in turn divided into intentional constraint and unintentional (Berlin 2002 [1958]), or culpable and non-culpable (Nozick 1974), or according to whether or not someone can be held responsible for the constraint (Kristjánsson 1996). Yet others focus on the possibility that others will
interfere to limit our options (Pettit 1997). To me, it seems important that we have a wide range of good options to choose from, particularly when it comes to life-defining choices such as what careers to pursue and where to live and with whom. It does not seem important what has caused us to have some rather than other options, or whether someone is culpable or responsible for us having the options we have. Leaving be is not leaving free. I am undecided on whether it is important that others cannot interfere, independently of whether or not they actually do. Questions of what has value are, however, notoriously difficult to argue.

There is also a partly overlapping discussion on autonomy and its value. I believe that autonomy is an (or perhaps several) independent value(s) in its own right. Dworkin (1988) defines 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 the light of higher-order preferences and values.’ (p. 20) I agree with Dworkin that this is an important value. While it is important that we can do what we want, and choose from a wide and rich range of options, it is equally important that we can reflect on our wants and, to some degree, change them. Mill’s anti-paternalism is in part based on his conjecture that the doctrine will promote the value of individuality, in the sense of living an original or even eccentric life, performing ‘experiments of living’ (1991 [1859], chapter III). Perhaps there is a special value in this too (no doubt it serves to enrich our view of the possible, which is perhaps Mill’s main reason to favour it).

Capabilities can be interpreted as a super liberty value, encompassing all other liberty values. We should distinguish, however, between what is important in itself to have or be, and what is important that one has the option of having or being. The concepts of liberty and autonomy identified are of the former kind and so have value independently of their possible incorporation under capability.

**6.3 Libertarian Paternalism**

Thaler & Sunstein (2003a, 2003b) have established the term libertarian paternalism for the promotion of healthy choice through non-intrusive influence. Libertarian paternalism is paternalism in the sense that ‘it attempts to influence the choices of affected parties in a way that will make choosers better off.’ (2003b, p. 1162) It is libertarian in the sense that ‘people should be free to opt out of specified arrangements if they choose to do so.’ (p. 1161) That is, it is libertarian in the weak sense that it does not conflict with the freedom of choice of the parties. Thaler & Sunstein accurately and importantly note that such freedom is consistent with a great deal of influence, since, because of the great impact of default rules, starting points and framing effects, there is no way to track people’s preferences as they exist independently of choice situations. These observations support the attribution of value to being free to do what one actually wants to do, rather than to realize some more hypothetical kind of preference. People often do not have well-formed preferences independently of choice situations, but once they are in one, they usually quickly form a preference, and then it is important that they be free to act on it. Since Thaler & Sunstein are not particularly impressed with natural ownership rights,
and since their doctrine does not positively promote libertarianism (one could imagine influence on choice that would promote autonomy, self-reliance, or a libertarian lifestyle of non-interference), it might perhaps more appropriately be called *non-intrusive paternalism*.

I think that it is obvious that we should, when possible, ensure that the choice situations that we face are as good as possible. This means in part that the alternatives available should be good ones in the sense that they tend to promote what has value (but variation also has value – independently of outcomes). If there must be an ordering of the alternatives such that some are more easily accessible than others, as is most often the case, this ordering too should promote what has value, such as health. I therefore subscribe to a version of the doctrine of libertarian or non-intrusive paternalism. Choice situations should in large part be designed so that we tend to make good choices. Such design is an important task for public health policy.

Thaler & Sunstein’s most basic example is the placement of the desert in a cafeteria, where the director simply has to make a choice one way or other where to place the desert – early or late in the line. However, they do not shrink from much more far-reaching conclusions:

But governments, no less than cafeterias (which governments frequently run), have to provide starting points of one or another kind; this is not avoidable. As we shall emphasize, they do so every day through the rules of contract and tort, in a way that inevitably affects some preferences and choices.’ (p. 1165)

Concrete examples include US labour regulation such as the 40 hour working week and the requirement that employees be discharged only for a cause, or else be compensated (p. 1187).

Policy makers must normally choose, or act as if they had chosen, between regulating and not regulating. Regulation affects people directly and through the regulation of other-regarding behaviour such as desert placement in cafeterias. Concerning the latter, not regulating means that people will make other-regarding choices based on their own self-interest, or on some other reason or influence. In the cafeteria case, the director might just place the desert where it is usually placed, or roll a dice. I propose that we should not passively stand by and let the society and environment we live in be shaped by habit, chance and special interests. We should actively design our environment to further the kind of life that we value. If we disagree on what has value, this is a reason for discussion and compromise. Policy makers should, directly and indirectly, promote good choice situations. If some choice-affecting policy is of great consequence, this is no reason not to enact it, though it is perhaps a reason to submit it to public scrutiny and debate.

Thaler & Sunstein’s focus on employee savings is instructive. Governments can regulate employers to ensure that their employee policies are conducive to good health. Governments can also enact welfare programs, impose taxes, offer subsidies, and design public spaces and the infrastructure more generally in ways that promote public health.
Sometimes popular resistance to regulation is an obstacle to public health policy. Such norms are partly shaped by advertising and private propaganda, which can be countered by regulation. The elimination of advertising that stimulates unhealthy behaviour and resistance to health-promotion would likely have very positive effects on public health and would most likely be non-intrusive.

Thaler & Sunstein recognize that people may prefer not to choose and that requiring them to choose may therefore be paternalistic (in some sense, presumably not their own). However, they do not consider the possibility that people may more generally prefer certain choice situations, even if they have no settled preference for how they would choose in that situation. There is a tendency in Thaler & Sunstein to assume that liberty is protected as long as people have a chance to opt out of any program. Opt-out freedom is liberty in a very narrow sense. Thaler & Sunstein’s focus on promoting welfare should be balanced against other liberty values as well.

7. Overview of Papers

‘The Normative Core of Paternalism’ defines anti-paternalism and ‘Anti-paternalism and Invalidation of Reasons’ rejects it. ‘Paternalistic Interference’ drives home the point that paternalism is not prima facie wrong. The focus of the first paper is on the complex interconnections between reasons and actions that must be sorted out in order to provide the most generous understanding of anti-paternalism. The second paper builds on this definition to characterize anti-paternalism as an instance of an influence-regulating principle of invalidation of reasons, and three arguments against such principles are presented. In the third paper, five accounts of the action component of paternalism are surveyed and found inadequate as specifications of something that it should be prima facie wrong to do out of benevolence. These three articles are the most general, most abstract and to my mind most fundamental to this thesis. It is no accident that I have most frequently referred to them in this introduction.

‘Liberalism, Altruism and Group Consent’ is an attempt to investigate how liberalism, in the spirit of anti-paternalism, can deal with regulation of groups where some members consent to regulation and some do not. In contrast to the rest of the thesis, I am here completely devoted to interpreting and refining anti-paternalism, not to criticise it. However, in this process its limits become apparent, including the impossibility of justifying the regulation of groups by partial consent of its members only, with no regard for their good. The proper regulation of divided groups seems to me a central issue in public health ethics, as opposed to traditional medical ethics.

This issue is carried over to the first of three more practically oriented articles, where my theoretical understanding of anti-paternalism is put to use in various policy areas. ‘Anti-paternalism and Public Health Policy: The Case of Product Safety Regulation’ investigates how anti-paternalism can be interpreted in that specific context, and suggests that the liberal does not need it. ‘Responsibility, Paternalism and Alcohol Interlocks’, co-authored by Jessica Nihlén Fahlquist, investigates the empirical and economical issues relevant for assessing a policy of mandatory alcohol interlocks in all
cars, as well as considers possible liberal resistance to such a policy based on the idea of individual responsibility as well as on anti-paternalism, finding the objections unconvincing. Finally, ‘Epistemic Paternalism in Public Health’, co-authored by Sven Ove Hansson, considers the case for withholding of information of uncertain threats to public health – a paternalistic policy designed to protect people from allegedly unnecessary anxiety and depression. Based not on principled anti-paternalism, but on broadly consequentialist considerations, such a policy is found unwarranted and counter-productive.

In writing this paper, I was responsible for the overall structure and coherence of the text. I also wrote most of the section on paternalism, while Jessica wrote most of the section on responsibility. That said, we were both actively involved in all parts of the essay.

While I did most of the writing for this paper, both authors contributed equally to all parts.
REFERENCES


Parfit, Derek. *Climbing the Mountain*. Manuscript.


