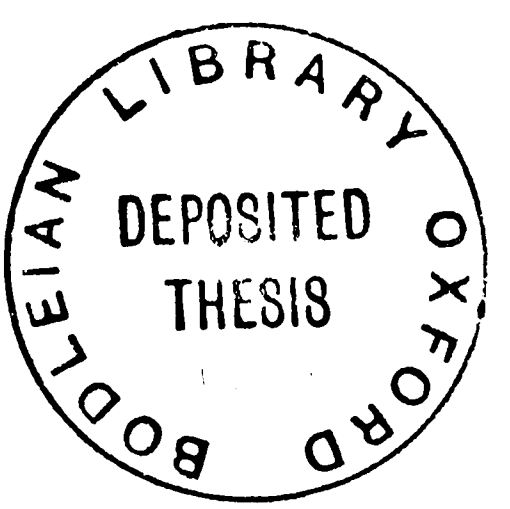
State Paternalism and the Neutrality-Perfectionism Debate

Simon R. Clarke
Balliol College

Submitted to the Faculty of Social Studies in fulfilment of the requirements for the degree of Doctor of Philosophy in Politics, University of Oxford, Hilary 2001.



Thesis Abstract
State Paternalism and the Neutrality-Perfectionism Debate
Simon Clarke, Balliol College, D.Phil., Hilary 2001

The purpose of this thesis is to compare the paternalistic implications of two opposing political theories: neutrality and perfectionism. Neutrality holds that the state must not appeal to conceptions of the good as the justification for its decisions. Perfectionism rejects this constraint.

Applied to paternalism, perfectionism makes it permissible for the state to appeal to a conception of the good when it acts paternalistically, that is, interferes with a person for his own good. Neutrality requires that paternalism must instead be guided by considerations that are in some sense neutral between various conceptions of the good. Perfectionism supports forcing people into worthwhile conceptions of the good. It provides motivation for a wider range of paternalistic policies than neutrality does. A number of perfectionist strategies for resisting this conclusion are examined.

The first is to appeal to the value of autonomy as a component of well-being. Paternalism, some perfectionists argue, undermines the value of living an autonomous life. I try to show how various arguments for the value of autonomy, even if accepted, fail to rule out a wide range of paternalistic policies. A second strategy is to appeal to the endorsement constraint. According to this, a person's well-being cannot be advanced by forcing her into activities that she does not endorse as valuable. I argue that the endorsement constraint is not plausible in its strong form, and that whilst a weaker form is plausible, it allows a wide range of paternalism. A number of other strategies, such as appealing to the value of activity, claiming that many worthwhile activities require people to approach them with the right intentions for those activities to be for their own good, and that paternalism undermines trust in the government are examined and criticised.

These difficulties do not mean that perfectionism should be rejected. But they do support setting aside conceptions of the good when the state acts paternalistically, whilst not necessarily ruling out perfectionism in non-paternalistic state action. This conclusion is strengthened in two ways. First, by taking Rawlsian contractualism as a method of elucidating neutrality, it is shown that neutrality supports a plausible principle of paternalism. Second, a number of recent attempts to set out necessary conditions for justified paternalism, such as that liberty must be balanced against well-being, that the consent of the paternalised is needed, and that the conduct must be nonvoluntary, are examined. The arguments for these conditions all suggest that neutrality is a necessary condition for justified paternalism.

The conclusion of the thesis is two-fold. A conclusion about the neutrality-perfectionist debate is that neutrality is required for paternalistic state action whilst perfectionism may be acceptable in the non-paternalistic sphere. A conclusion about state paternalism is that it is justified only if guided by neutral considerations.

Acknowledgments

I am grateful to my supervisor Joseph Raz for his helpful comments and questions on all of the chapters, and for his guidance throughout the entire process. I have learnt much from him but can hope to only approach his depth of understanding of political philosophy.

Michael Freeden supervised my first term in Oxford and gave helpful advice. Jerry Cohen and David Miller asked probing questions during their examination of some early chapters. Others who gave me comments on drafts of chapters were Halvor Nordby, Jean-Frederic Lafaille, Nir Eyal, Matthew Liao, Paul Markwick, Andrew Moore, Daniel McDermott, Gorana Djoric, and Martin Wilkinson. Drafts of some chapters were presented at several sessions of the Graduate Political Theory Workshop at Oxford University, a graduate conference in Warwick in May 2000, and a seminar in New York in fall 1999. I am grateful for the comments of participants.

For financial support I must thank the New Zealand Vice Chancellors' Committee for awarding me the Gordon Watson Scholarship and the William Georgetti Scholarship. Without those awards I would not have been able to even begin the thesis. I am also grateful to the British government for the support provided by an Overseas Research Student Award and to Balliol College for various graduate bursaries and book grants. Finally I also thank my parents for their support, which was not just financial.

Table of Contents

		Page
Title		i
Abstract		ii
Acknowledgments		iii
Table of Contents		iv
INTRODUCTION		1
I	DEFINING PATERNALISM, NEUTRALITY, AND PERFECTIONISM	
1.1	Introduction	9
1.2	Defining Paternalism	9
1.3	Neutrality and Perfectionism	21
1.4	How Paternalism Relates to Neutrality and Perfectionism	34
II	PERFECTIONIST AUTONOMY AS PART OF WELL-BEING	
2.1	Introduction	38
2.2	The Perfectionist Conception of Autonomy	43
2.3	Appeal to Intuitions	46
2.4	Rational Deliberation and the Exercise of Agency	50
2.5	Giving Meaning to Life, Having Character, and Responsibility	63
2.6	The Social Forms Argument	81
2.7	Autonomy and Worthless Options	94
2.8	Conclusion	101
III	RAZ ON COERCION AND PATERNALISM	
3.1	Introduction	103
3.2	Coercive Paternalism	104
3.3	Two Arguments	108
3.4	The Trust Argument	113
3.5	Conclusion	122
IV	THE ENDORSEMENT CONSTRAINT	
4.1	Introduction	123
4.2	Endorsement	127

4.3 Three Objections 130
4.4 The Intuitive Argument 133
4.5 Performance and Endorsement 143
4.6 Integrity and Endorsement 146
4.7 Conclusion 156

V HURKA ON ACTIVITY AND INTENTION
5.1 Introduction 157
5.2 Hurka's Arguments 158
5.3 Lifestyle Paternalism 161
5.4 Aesthetic Paternalism 170
5.5 Relationship Paternalism 175
5.6 Conclusion 179

VI NEUTRAL PATERNALISM
6.1 Introduction 182
6.2 Rawlsian Contractualism 185
6.3 Rawls on Paternalism 189
6.4 Richards on Paternalism 203
6.5 A Revised Principle 206
6.6 Conclusion 212

VII SOME ALTERNATIVES TO NEUTRALITY
7.1 Introduction 213
7.2 The Balancing View 214
7.3 The Consent Principle 219
7.3.1 The Certainty Principle 219
7.3.2 Future and Hypothetical Consent 221
7.3.3 Actual and Present/Prior Consent 226
7.4 The Voluntariness Principle 233
7.5 Conclusion 239

CONCLUSION 241

BIBLIOGRAPHY 246

Introduction

State paternalism occurs when the state restricts an individual's freedom for his own good, usually against his will. Laws against drugs, prostitution, pornography, gambling, and euthanasia, insofar as they are intended for the good of those whose freedom is restricted, are examples of state paternalism. The purpose of this thesis is to examine the implications for paternalism of two different understandings of the theoretical foundations of the state. In particular, I assume that there are or ought to be principled limits to state paternalism and I try to determine which of the two theories of the state can best account for this resistance against paternalism.

The first of these theories is sometimes thought to be the essence of the liberal state. It sees the state as providing a neutral framework of rules within which citizens may pursue their own goals and activities as they see fit. No particular vision of how people ought to live their lives is to be privileged by the laws and policies of the state. So long as a person does not harm others or take what justly belongs to them, she ought to be left free by the state in deciding how to live her life. The second theory rejects this ideal of neutrality as an unattractive mirage. All states do and ought to, it says, help citizens to live valuable and worthwhile lives. The state need not confine itself merely to enforcing justice and preventing people from harming others. It can and should help create conditions under which people can live fulfilling lives. There may be limits on the means that the state ought to use in pursuing this goal, but in principle the state's laws and policies may be based on such considerations. In this vision of the state, it does not stand above the fray in the debate of how people ought to live their lives, but rather takes sides in that debate in order to help people lead good lives.

These two theories are called neutralism and perfectionism respectively. I describe them more precisely in the first chapter. But from these bare sketches alone, the issue that this thesis focuses on is immediately clear. The second theory, perfectionism, has a paternalistic ring to it. If the state is to help people live good lives, it may have to restrict their freedom if they are mistaken about their own good. The perfectionist state seems committed to a wide range of paternalistic policies. In recent years, defenders of perfectionism have attempted to show that their theory of the state is not committed to as much paternalism as first appearances may suggest. Most of the thesis focuses on their arguments. I try to show that the arguments are ultimately unsuccessful. The perfectionist state is committed to a wide range of paternalistic policies. Neutralism on the other hand, as I also try to show, is not. First appearances in this instance are the correct ones.

This conclusion can be understood in different ways. Of the two theories, neutralism is less paternalistic than perfectionism. Insofar as we seek a theory of the state that can account for the widespread suspicion of paternalism, we have reason to favour the neutrality ideal. We can test the theories by examining their implications, and with respect to paternalism at least, I will argue, neutrality is the more plausible of the two. But the conclusion can also be understood as a statement about the justifiability of state paternalism. By accounting for beliefs about paternalism, neutralism gives these beliefs structure and a general explanation. A principle for when paternalism is justified follows: state paternalism is justified only when it respects the neutrality constraint. Certain grounds are ruled out as the basis for justified state paternalism and as a result, certain laws and policies which would have only those grounds as their rationale, are ruled out as illegitimate state action.

The thesis consists of seven chapters. In the first, I define paternalism more precisely. Above, I said that paternalism restricts freedom for a person's own good, usually

against his will. I explain how paternalism need not be contrary to its target's will and that some cases of paternalism do not restrict individual freedom. The essence of paternalistic behaviour, I will argue, is that it attempts to restrict a person's options, or choose for a person when he cannot choose for himself, for that person's own good.

A large part of the first chapter is also devoted to explaining the theories of neutralism and perfectionism more fully. The major point of difference between the two theories is a disagreement over the grounds that may legitimately be appealed to in the justification of state action. Neutrality excludes a certain type of rationale for laws and policies, those that appeal to some view of what would be a worthwhile way of life. Such views and components of those views are called conceptions of the good. The neutral state does not base its decisions on any conceptions of the good. This has practical implications: any law or policy that is not neutrally justifiable, that can only be justified by reference to a conception of the good, is illegitimate. Perfectionism rejects the neutrality constraint and holds that it is permissible for the state to base its decisions on a conception of the good. Justifications for state action may appeal to what is good or valuable in human life. As I suggested above, perfectionism seems to provide support for a wide range of paternalistic practices. Chapter 1 ends by spelling out this charge more fully.

Chapters 2 through 5 examine perfectionist responses to this charge. Perfectionists have a number of arguments that attempt to show that a perfectionist state need not be overly paternalistic. Perhaps the most common argument is to appeal to the value of personal autonomy as a component of the good life, and chapter 2 examines this argument. Since people's lives go well, so the argument goes, when they autonomously direct their own lives, and since paternalism undermines this aspect of well-being, perfectionism is antithetical to rather than supportive of paternalism. I concede that paternalism usually conflicts with autonomy. But this is an argument against paternalism only insofar as autonomy is valuable. I examine a number of

arguments for the value of autonomy and conclude that they are compatible with a significant range of paternalistic measures. I also give reason for thinking that autonomously choosing activities has no value when those activities are worthless ones, which reduces the appeal of autonomy as a means for resisting paternalism.

One defender of perfectionism who emphasises the value of autonomy is Joseph Raz. He has a number of other arguments relevant for paternalism and chapter 3 addresses some of these. Raz argues that an accurate perfectionist account of well-being contains components that would be undermined by coercion. Since much state paternalism, particularly the most objectionable cases, are of a coercive nature, his arguments give reasons against paternalism. I try to show that a perfectionist ban on coercion must allow for some exceptions, such as in order to prevent harm to others, and that once these exceptions are allowed, there are no good grounds for a perfectionist to not also allow coercion in the paternalistic sphere. Raz also argues that coercive paternalism must be accompanied by an attitude of trust on the part of the citizens towards their government for it to be legitimate. I show that contrary to Raz's claim, such an attitude can occur even when the state acts paternalistically.

In chapter 4, I examine a strategy of perfectionism that appeals to another thesis about well-being. This is that in order for an activity to advance a person's well-being, it is necessary that she believes in the value of that activity. A person's life cannot be improved, so the claim goes, against her conviction that it is not. This sets an endorsement constraint on the acceptability of paternalism: since most intervention would not be accompanied by the paternalised's endorsement of the activities she is forced into, paternalism may not usually proceed since it would not improve her well-being. I examine a number of arguments for the endorsement constraint. I try to show how these arguments actually support a weaker, but more plausible, thesis about a person's attitude towards aspects of her life, one that is compatible with a wider range of paternalistic measures.

The last chapter examining the perfectionists' response to the charge of paternalism, chapter 5, considers a range of arguments put forward by Thomas Hurka. Hurka's two main arguments are that well-being contains an important aspect whereby people must engage in activities themselves rather than be helped by others, and that they must engage in activities with the right sort of intentions. These arguments are similar to, but distinguishable from, those addressed in chapters 2-4. But they face similar problems. Even if Hurka's main claims are true, they fail to rule out a significant range of paternalistic measures. This is because a state that has a wide number of paternalistic policies can nevertheless allow scope for people to engage in activities themselves, and with the right intentions.

The conclusion of these chapters is that perfectionism is unable to account for the unacceptability of many cases of paternalism. The neutral state on the other hand is more likely to since it excludes the motivation for paternalism that is central to perfectionism. This conclusion is strengthened in chapter 6 by setting out what a neutralist approach to paternalism would look like. Two defenders of neutrality, John Rawls and David Richards, have proposed principles of paternalism which, they argue, respect the neutrality constraint and rule out most paternalism but allow it in some carefully limited cases where it ought to be allowed. But I argue that their principles still allow too much paternalism. Whilst they do rule out most of the paternalism that a perfectionist state would support, they allow paternalism that is guided by a person's own conception of the good. This, I try to show, may be too much. I present my own principle which, I argue, avoids this difficulty whilst at the same time capturing the plausibility of Rawls's and Richards's proposed principles.

The final chapter, chapter 7, tries to strengthen the support for neutrality in the sphere of paternalism in another way. Three possible alternatives to the neutrality constraint are examined. These are possible alternatives in the sense that they provide principles

for the restriction of paternalism to only certain types of cases. If any of them are true, then they may provide a means whereby perfectionism, by embracing them, can limit its own paternalistic motivations without having to concede the case to neutrality. Perfectionism with the addition of any of these principles, rather than neutrality, could account for the wrongness of much paternalism. I show how none of the three possible alternatives can help the perfectionist case. They are all in fact elucidations of neutrality or are only plausible in conjunction with the principle of neutrality. For perfectionism to embrace them would be for it to embrace neutrality.

I do not believe my arguments are the final word on the matter, so in the Conclusion I point out some possible weaknesses in the argument and responses that could be made that I have not had space to cover. I also explain some of the practical implications of the argument.

Some of the fundamental assumptions I make should be noted at the outset. One is that certain activities, goals, projects, and pursuits that people engage in are valuable, that is, contribute to their well-being, whilst others do not. People sometimes choose worthless activities because of misguided beliefs. Some people's conceptions of the good are worthwhile, in other words, whilst some are not. People can be mistaken about what is worthwhile, and these mistakes are what motivate paternalism. This assumption is essential to the perfectionist view and without it, few if any of the arguments in this thesis could proceed. If the assumption is false, then the debate between neutrality and perfectionism is pointless. In making this assumption, I hope to avoid taking sides in any (or at least the major) controversial questions about the nature of value. The assumption does not, I think, imply a thesis of strong anti-relativism about value. It is possible that the value of activities is determined by shared social practices and understandings and that the value of a given activity varies across societies because of different practices and understandings. Even assuming this view of the nature of value, it is still possible for people to be mistaken about the

value of activities. They could fail to appreciate the shared practices and understandings of their society. The assumption implies nothing more controversial, I hope, than that people may be mistaken about their own good.

The other major assumption I make is that there are clear hypothetical cases of unacceptable state paternalism, such as laws forcing couch potatoes to live more active lifestyles and forcing people to experience great art. I also assume that there are some cases of paternalism that are relatively unobjectionable, although I place less weight on these. I make these assumptions about the wrongness and rightness of cases of paternalism without explaining why they are wrong or right. That is what we should hope either of the two theories, neutrality and perfectionism, can explain. A plausible theory should give us the right results and explain why they are right. I will argue, however, that perfectionism does not even give us the right results. If so then we have reason for thinking that the correct explanation for why the cases of paternalism are unacceptable lies elsewhere. It may be that there is no correct explanation and that the state is not prohibited from implementing the paternalistic policies that seem unacceptable. But we should be reluctant to embrace this conclusion, only doing so if we cannot find a plausible theory that can explain our convictions. Neutrality, I will argue, provides such an explanation.

But it is an explanation that is limited in its depth. It accounts for certain judgements about paternalism and places those judgements in a particular understanding of the state, but it does not give any deeper reason for accepting that understanding, other than that it is able to account for certain judgements about paternalism. Someone who is undisturbed by the paternalistic policies that I will argue are supported by perfectionism may think that the truth of perfectionism means that its implications must simply be accepted and ask 'what reason is there, one that is independent of the policies and laws the theory supports, for accepting the neutrality theory of the state?' My arguments in this thesis do not provide an answer to this question. But it is

significant that most defenders of perfectionism, especially those whose arguments are examined in this thesis, do not seem to think that the theory should be accepted regardless of any objectionable policies that it may support. They all make efforts to show that perfectionism does not support such policies. The conclusion that I will argue for, that neutrality rather than perfectionism is more plausible with regard to the issue of paternalism, is one that, if true, would worry most perfectionists.

Chapter One

Defining Paternalism, Neutrality, and Perfectionism

1.1 Introduction

Before investigating the paternalistic implications of neutral and perfectionist theories, it is necessary to explain the meaning of some important terms. There are a number of competing definitions of paternalism. Section 1.2 sets out what I believe is the most accurate definition, comparing and defending it against some of the competing accounts. Section 1.3 provides a basic sketch of the neutrality-perfectionism debate. There is a large and complex literature surrounding this debate, but I attempt to focus on what I take to be the core issue: a disagreement about the legitimate grounds for state action. The third and final section of the chapter sets out how paternalism is related to the distinction between neutrality and perfectionism, showing that whilst both types of theories can support paternalism, perfectionism seems to support more of it. With this conceptual framework in mind, the meaning and implications of the arguments in subsequent chapters may be more readily understood.

1.2 Defining Paternalism

1.2.1 A simple definition of paternalism is that it is the restriction of a person's freedom, usually against her will, for her own good. This definition captures many cases that are often thought to be paternalistic, such as motorcycle helmet and seatbelt legislation, preventing a person from unknowingly walking across a dangerous bridge, a doctor refusing to inform his patient of her terminal illness, requiring people to undergo blood

transfusions when needed despite their religious objections, and laws against drug use, certain sexual activities, and gambling. The definition also explains the motivation for paternalism -- to benefit the paternalised -- and suggests why it may be morally troubling -- because it involves a restriction of freedom, usually against the paternalised's will.

The simple definition, however, does not capture all that we mean by paternalism. A more precise definition is the following:

X behaves paternalistically towards *y*:

(1) only if *x* aims to close an option that would otherwise be open to *y* or *x* chooses for *y* in the event that *y* is unable to choose for himself; and

(2) to the extent that *x* does so in order to promote *y*'s good.

According to this definition, only behaviour that aims to close an option or that aims to choose for a person can be paternalistic, and such behaviour is paternalistic insofar as it is for the target's own good.

Paternalism is therefore concerned with the good of the person subjected to the intervention. Restricting a person's freedom solely to prevent him from harming others, or society in general, or to promote some non-person-affecting good (if there is such a thing) is not paternalistic. Many acts, however, may involve multiple rationales. Someone may intervene with my action both to protect others and prevent me from harming myself. The intervention is paternalistic with respect to the latter motivation. That is the meaning of the phrase 'to the extent that...' in condition 2 above. Moreover, paternalistic motivation may be a major or a minor reason for acting. A wife who tries to stop her husband from drinking primarily because she believes abstention is for his own good may do so also because she is aware that he becomes violent towards her when

drunk.[1] A person's main reason for interfering with another may be to prevent harm to others, but she may also do it to help the person being interfered with. Both both these cases involve paternalism to a greater or lesser extent.[2]

In the above definition of paternalism, *x* and *y* may range over groups or individuals. *X* may or may not be the state since paternalism is not limited to laws and public policy. A friend acts paternalistically when she hides my cigarettes from me, a firm acts paternalistically when it (unforced by the state) makes its employees put part of their earnings into a pension plan, and parents often act paternalistically towards their children. If (and only if) the intervener is the state, then the intervention is a case of state paternalism, and if (and only if) it uses the law to intervene, then it is legal paternalism.

Paternalism is characterised above as a form of 'behaviour'. The meaning of this is that it includes both acts and omissions. *X* withholding from *y* that *y*'s spouse has died to prevent *y* from experiencing distress when about to undergo surgery is rightfully described as an act of paternalism.[3] And so is a doctor withholding information from a patient for what he sees as that patient's own good. Thus an accurate definition of paternalism covers omissions as well as acts, so long as the omissions are intended to benefit the person.

---

[1] The example is from David Archard, 'Paternalism Defined', *Analysis*, vol. 50, no. 1, 1990, p. 38.

[2] Some authors hold that behaviour is paternalistic only when the 'primary or sole aim' (Donald VanDeVeer, *Paternalistic Intervention*. Princeton, New Jersey, Princeton University Press, 1986, p. 22) or 'main reason' (Archard, 'Paternalism Defined', p. 38) for the behaviour is to benefit the paternalised. Defining paternalism in this way creates two problems: how to classify intervention that has the target's good as a minor motivation; and defining the terms 'main' or 'primary'. These problems may be avoided by saying that behaviour is paternalistic *to the extent that* it is for the paternalised's own good. My definition shares this feature with John Kleinig's. See his *Paternalism*, Manchester, Manchester University Press, 1983, p. 12.

[3] VanDeVeer, *Paternalistic Intervention*, p. 22.

The paternalist aims to close an option of the paternalised, but may fail to do so.[4] Some people may obtain certain recreational drugs despite legal prohibitions on those drugs. Such prohibitions are nevertheless paternalistic insofar as they are intended for the good even of those who dodge them. Another way that paternalism may fail is that the paternalist may aim but fail to bring about the good of his target. Laws against prostitution may be intended to benefit those who would otherwise be prostitutes. But they may in fact end up worse off under those laws. Such legislation would be unsuccessful paternalism but paternalism nevertheless.

The above definition is morally neutral in the sense that it allows the possibility of both justified (whether simply permissible or also required) and unjustified paternalism. The definition allows us to point to examples of paternalism without already judging the moral question of their permissibility. Thus descriptive considerations are separated from evaluative ones. But the definition also suggests why paternalism is often thought wrong. There are various methods by which a person may close another's options that may be at least presumptively wrong. These could involve the use of force, coercive threat, deception, cheating, manipulation, killing, physically impairing, causing pain, depriving of freedom, opportunity, or pleasure, or promise-breaking. If presumptively wrong in one of these ways an act of paternalism requires justification either by demonstrating countervailing considerations in the presence of such a wrong, or showing that there is in fact no presumptive wrong to begin with.[5]

Since paternalism involves attempting to benefit another person, it is a type of benevolence. But not all benevolent behaviour is paternalistic. Giving food to a hungry person is not normally paternalistic. The above definition explains why, since such an act

---

[4] This point is from Archard, 'Paternalism Defined', p. 37.
[5] VanDeVeer, *Paternalistic Intervention*, p. 22.

does not close an option. The hungry person may always give away or abandon the food. No opportunity is closed that would otherwise be open to the person. We can imagine circumstances where an option is closed. A person may not only give food to a hungry man, but also force him to eat it. Such an action seems paternalistic, and my definition explains why.

1.2.2 There are competing accounts of what makes benevolent behaviour paternalistic. Gerald Dworkin defines paternalism as 'interference with a person's liberty of action' for the good of 'the person being coerced.'[6] However, as several commentators have pointed out, paternalism need not be coercive nor interfere with liberty of action. Consider a case where Jones tells a mother on her deathbed that her son is doing well when Jones knows the son has just been killed attempting to escape from prison after being found guilty of multiple rape and murder. Jones is acting paternalistically toward the mother, but no coercion is involved, he is not restricting her liberty of action, and is not even attempting to control her behaviour.[7] The case is one of paternalism on my definition. The doctor could give her the option to know the truth about her son. But he does not, and his omission thereby closes the option for her own good.

Another case of paternalism that does not restrict liberty of action or involve coercion is giving a blood transfusion to an injured, unconscious person whom we know to be opposed to blood transfusions.[8] Such a case is accounted for and is what motivates the second part of clause 1 of my definition. It is not covered by the first part. The options of the patient are already closed in virtue of the fact that she is unconscious. Her

---

[6] Gerald Dworkin, 'Paternalism', in Peter Laslett and James S. Fishkin (eds.), *Philosophy, Politics, and Society: a collection*, 5th series, New Haven, Yale University Press, 1979, p. 78.

[7] This example is from Bernard Gert and Charles M. Culver 'Paternalistic Behavior', *Philosophy and Public Affairs*, vol. 6, no. 1, 1976, p. 46.

[8] Gert and Culver, 'Paternalistic Behavior', p. 46.

unconsciousness prevents her from having the options to choose whether or not to have a transfusion. Hence the paternalist cannot close the patients options in this respect. But he can choose for the patient since she is unable to choose for herself. In doing so he would be acting paternalistically. (It would also be paternalism if the unconscious patient is not opposed to transfusions, although not as morally interesting a case as when she is.)

John Kleinig argues that in light of examples of paternalism that do not restrict liberty of action, paternalism should be thought of as covering the restriction of any freedom, including but not limited to restrictions of liberty of action. Paternalism, he says, may abrogate other freedoms, such as the freedom of thought and expression, and the freedom to be let alone.[9] But it is not clear that all examples of paternalism restrict freedom. Cases of manipulation, where *x* manipulates *y*'s preferences in such a way that *y* has no real choice even if *x* is not interfering with *y*'s choice in an overt way, do not obviously do so. If a father indoctrinates his daughter from birth not to want to marry until she's aged 30, he is denying her the option to choose for herself.[10] This seems as much a case of paternalism as if he simply forbade her from marrying before aged 30. Similarly, the state may aim to manipulate people's preferences for unhealthy activities, for their own good. Gerald Dworkin also suggests that his refusal to play tennis with someone because that person gets upset when he loses, which he usually does, is a case of paternalism that does not restrict freedom in any way.[11]

The plausibility of a definition that makes restriction of freedom necessary for paternalism depends upon what conception of freedom is used. Some conceptions may embrace the above examples of paternalism. Kleinig takes freedom to consist in the

---

[9] Kleinig, *Paternalism*, p. 6.

[10] VanDeVeer, *Paternalistic Intervention*, p. 19.

[11] Gerald Dworkin, 'Paternalism: Some Second Thoughts', in Joel Feinberg and Hyman Gross (eds.), *Philosophy of Law*, Belmont California, Wadsworth, 1995, pp. 219-20.

absence of *impositions* on a person. This leads him to propose the following definition of paternalism:

> X acts paternalistically in regard to Y to the extent that X, in order to secure Y's good, as an end, imposes upon Y.[12]

It is clear from the above definition and other of Kleinig's comments that he takes the idea of an imposition to be the same as that of restricting freedom. This is a mistake. If someone asks me the time of day, he is imposing upon me but not restricting my freedom in any way. Kleinig seems to make the mistake of thinking that since all restrictions of freedom are impositions, all impositions must be restrictions of freedom. This has the result of making his definition of paternalism too broad. If I remind a friend of the time of day so that she is not late for an appointment, this seems hardly an instance of paternalism, yet it fits Kleinig's definition since my action would be an imposition.

Kleinig admits that there are different conceptions of freedom. He concentrates on two. One is the view that freedom consists in the absence of constraints on doing what one desires. This, however, cannot be the sense in which freedom is restricted by paternalism since as is suggested below, some cases of paternalism do not conflict with the target's desires. The second conception holds that freedom is 'the nonrestriction of *options*.'[13] This, I suggest, is the sense of freedom in which paternalism almost always restricts freedom. The resulting definition of paternalism is almost equivalent to my own. But my definition avoids taking a stand on the controversial question of what freedom is. And paternalism only *almost* always restricts freedom in the sense of restricting options since as was seen in the blood transfusion case above, paternalism sometimes does not restrict options.

---

[12] Kleinig, *Paternalism*, p. 13. See p. 7 for Kleinig's claim that restrictions of freedom are impositions on people.

[13] Kleinig, *Paternalism*, p. 19. Emphasis in original.

There are several other ways that some authors have attempted to define paternalism so that it embraces more than restrictions on liberty of action but without making it so wide as to cover all benevolent behaviour. One is to hold that paternalistic acts or omissions are those that either are, or are believed by the paternalist to be, contrary to the target's will. Donald VanDeVeer holds such a view, since part of his definition of paternalism is that *x* believes that his behaviour is contrary to *y*'s 'operative preference, intention, or disposition at the time'[14] *x* performs the action or omission. Richard Arneson makes it a necessary condition of paternalism that the intervention be 'carried out either against [the paternalised's] present will ... or against his prior commitment'.[15]

Both definitions are however too narrow.[16] They do not cover seatbelt and motorcycle helmet laws, which are paradigm cases of paternalism. Laws requiring drivers and passengers to wear seatbelts while in motor vehicles may not be contrary to the operative preferences of many of those drivers and passengers. Similarly, laws requiring the wearing of helmets while riding on motorcycles are paternalistic even towards those who prefer to wear them. Requiring people to wear seatbelts and helmets for their own good seem instances of paternalism even if many of those people would voluntarily choose to

---

[14] VanDeVeer, *Paternalistic Intervention*, p. 22.

[15] Richard Arneson, 'Mill versus Paternalism', *Ethics*, vol. 90, no. 4, 1980, p. 471.

[16] VanDeVeer's definition is also too wide. Say I plant some roses near my neighbour's fence for his own good even though I know that my neighbour has a preference that I not do so. It seems inaccurate to describe that as paternalistic behaviour. The example is itself VanDeVeer's (see *Paternalistic Intervention*, p. 17) and it is clear he thinks that his definition avoids it. Nevertheless the example seems to fit his definition as an instance of paternalistic behaviour. This point does not apply to Arneson's definition since he also makes it a necessary condition of paternalism that it restricts liberty. Rose planting near my neighbour's fence does not restrict his liberty. VanDeVeer's definition contains no other conditions that can rule out the case as one of paternalism. Along with the condition quoted, that it be contrary to the target's operative preference, VanDeVeer thinks it sufficient for paternalism that *x* acts deliberately and for *y*'s own good (see his definition on p. 22 of *Paternalistic Intervention*). The rose-planting case seems to be thereby covered. My definition also avoids including the rose-planting case; planting roses in my garden does not close an option previously open to my neighbour. It was never his option whether there be roses there or not. It was only mine. Planting them in *his* garden does close an option of his. And it seems paternalistic if done for his own good.

do so anyway. Laws against prostitution and pornography intended for the good of those whose freedom is restricted are paternalistic even towards those who have no desire to engage in the prohibited activities. What is important is not that the act or intervention is contrary to the person's operative preference but that the option is closed, whether the paternalised wants it to be closed or not. The mistake these authors make, I suggest, is to concentrate on only the more interesting or morally problematic cases of paternalism. Whilst instances of paternalism that are not against the target's will may strike us as less interesting or morally troubling than others, this should not lead us to mistakenly believe that they are not cases of paternalism at all.

Similar remarks apply to another proposed definition. Bernard Gert and Charles M. Culver claim that 'an essential feature of paternalistic behaviour toward a person is the violation of moral rules (or doing that which will require such violations)'.[17] Gert and Culver's building of this feature into their definition of paternalistic behaviour does not mean that they think it always unjustified. Someone could have an adequate justification for violating a moral rule. But the Gert-Culver account is inaccurate since there are cases of paternalism that do not involve the violation of any moral rule. As Gerald Dworkin says, a husband who hides his sleeping pills from his suicidal wife violates no moral rule. They are his pills and he may put them wherever he wishes.[18] Again, the more interesting cases of paternalism may be those that violate a moral rule. But this is not a reason to make it part of the definition of paternalism and exclude as cases of paternalism those which do not violate moral rules.

---

[17] Gert and Culver, 'Paternalistic Behavior', p. 48. On pp. 49-50 they give a five-part definition of paternalism. Part three is that the person acting paternalistically must *believe* that his action involves violating a moral rule. This is different to the claim that his action actually does violate such a rule. It also seems implausible since we could think of many instances of paternalism where the person acting paternalistically does not believe his actions violate any moral rule.

[18] Dworkin, 'Paternalism: Some Second Thoughts', p. 220.

My proposed definition also contrasts with another recently put forward by Gerald Dworkin. As we have seen and as Dworkin has subsequently admitted, his original account is inadequate due to its limitation to only cases of coercion and interference with liberty of action. Dworkin's more recent proposal is that an act is paternalistic only if it is a limitation on *y*'s autonomy.[19] It is not clear what Dworkin means by autonomy. At one point he seems to suggest that it is equivalent to 'a usurpation of decision-making.'[20] But it is not clear that all such usurpations limit autonomy. According to one view, autonomy requires general freedom from coercion and manipulation, an adequate range of worthwhile and morally acceptable options, and identification with and loyalty to one's choices.[21] A particular usurpation of decision-making that is not coercive or manipulative, therefore, may not necessarily violate autonomy. But it may be paternalistic, which suggests that not all paternalism violates autonomy. Nor need paternalism usurp decision-making, since the paternalised may have no capacity for decision-making for the paternalist to usurp. Indeed, this could partly be what motivates the paternalist.

Moreover Dworkin's argument for restricting paternalism to behaviour that limits autonomy is a non sequitur. Dworkin considers the case of a father (a lawyer) who wants his daughter to be a lawyer. The daughter believes she would be such a good lawyer that she would be more successful than her father and that her success would make her father unhappy. She therefore decides not to become a lawyer. Other parts of Dworkin's definition, that it be contrary to the person's will and that it be for his own good, are satisfied in this case. Dworkin claims that since it is not a case of paternalism an

---

[19] Dworkin, 'Paternalism: Some Second Thoughts', p. 220. See also Gerald Dworkin's entry under 'Paternalism', in Lawrence C Becker (ed.), *Encyclopedia of Ethics*. New York, Garland, 1992, p. 940.

[20] Dworkin, 'Paternalism: Some Second Thoughts', p. 220.

[21] This view of what it means to be autonomous is taken from Joseph Raz, *The Morality of Freedom*. Oxford, Clarendon Press, 1986, pp. 369-90. The concept of autonomy shall be examined more fully in chapter two.

additional part is needed in the definition. He argues that it is not an instance of paternalism because the daughter does nothing to interfere with the self-determination of the father. From this he concludes that paternalism involves a violation of a person's autonomy.[22] However, his discussion could simply support something like part 1 of my definition without the need to claim that this involves a limitation on autonomy. My proposed definition can account for this case better than Dworkin's own. The daughter does not close any option that was previously open to the father. It is not his option what his daughter does, but hers (although he has a preference for how she should choose amongst her options).

The definition I have offered is similar to one set out David Archard. The main difference is that he adds another condition which says that the paternalist, *x*, must discount *y*'s belief that *x*'s behaviour does not promote *y*'s good.[23] This condition seems implausible. By 'discounting' Archard means that the paternalist ignores the paternalised's views, either believing that his target's view of the interference is one of disapproval and overriding it, or having no belief regarding his target's view of the matter.[24] Neither of these may be the case in many instances of paternalism. With laws requiring the wearing of seatbelts, for example, the state may think that many of those required to wear seatbelts believe that such laws do promote their good. The condition also implies that the paternalised has a belief regarding the intervention. But many people may hold no belief with regard to how they are affected by laws against prostitution, gambling and recreational drugs. Archard's rationale for the condition is not clear. He seems to think that if *x* believes *y* shares his view of what is good for *y*, then such a situation is one of benevolence but not paternalism. The case can be one of

---

[22] Dworkin, 'Paternalism: Some Second Thoughts', p. 220.
[23] Archard, 'Paternalism Defined', p. 36.
[24] Archard, 'Paternalism Defined', p. 39.

paternalism, according to Archard, only if *x* thinks *y* is wrong about *y*'s own good or *x* ignores *y*'s view of the matter.

However, the extra condition is not needed. The case where *x* believes *y* shares his view of the correctness of interference may not be a case of paternalism. But this, I suggest, is because *x*'s reason for acting is to give *y* what *y* desires. If so, then insofar as this, rather than a desire to bring about *y*'s good, is *x*'s motivation, the intervention is indeed not paternalistic. But my definition explains why, since in that case the second condition, that the behaviour is intended for *y*'s good, is not satisfied. *X*'s reason for acting is to bring about what *y* wants, rather than to bring about *y*'s good. Archard's extra condition is therefore unnecessary. A paternalist may believe that his target shares his view that the intervention is for his own good. He remains a paternalist insofar as his reason for intervening is the target's good rather than the target's desires.[25]

Given the definition of paternalism as the attempted closure of a person's options or choosing for a person for his or her own good, much paternalistic behaviour, perhaps most, would be morally unacceptable. Various suggestions have been made regarding its permissibility. Gerald Dworkin, for example, argues that it is justified only to preserve a wider range of freedom for the paternalised.[26] Another suggestion is that paternalism is acceptable only if the target has consented to the intervention or would do so if she were aware of all the relevant circumstances.[27] A third possibility is that paternalistic intervention is permissible only when the conduct involved is substantially nonvoluntary

---

[25] If the paternalist is a subjectivist about well-being then he may believe that in acting to bring about the satisfaction of the paternalised's desires, he will thereby advance her good. In that case he would be acting paternalistically. Hence the point in the text is more precisely made by saying that insofar as *x* acts to bring about the satisfaction of *y*'s desires *independently of the goodness for* y *of such satisfaction*, *x* is acting non-paternalistically. *X* may believe that he has reason to do something because *y* wants it even though he believes it would not be in *y*'s interest.

[26] Dworkin, 'Paternalism', p. 88.

[27] VanDeVeer, *Paternalistic Intervention*, p. 88.

or when temporary intervention is needed to establish whether it is voluntary or not.[28] My concern, however, is with a different limitation that applies specifically to paternalism by the state. Neutralist theorists argue for a constraint on the types of reasons that the state is permitted in acting upon. This constraint would apply to possible state paternalistic policies, ruling some of them out as illegitimate. It would make paternalism permissible only if it satisfies the neutrality constraint. The next section attempts to explain this constraint more fully.

1.3. Neutrality and Perfectionism

1.3.1 According to the ideal of state neutrality, the state should not favour any conception of the good. It should not take a stand on what is a worthwhile way of life or on what are valuable activities that people should pursue. Instead, the state should simply provide a neutral framework of laws and policies within which every person may pursue the good life as he or she understands it. The neutrality requirement is usually thought to restrict the state from, amongst other things, enforcing sexual morality, encouraging industriousness and discouraging idleness, forcing people into art galleries and museums, and inculcating a particular religion. It does not, however, entail that the state may do nothing. The state may, consistent with neutrality (or at least not obviously inconsistent with it), prevent people from being harmed by others, protect and enforce rights, and ensure a just distribution of resources. So long as it does these things without being biased towards a particular conception of the good, the state remains neutral.

Perfectionism rejects the requirement of neutrality. It holds that the state may take a stand on what is a worthwhile way of life and on what activities are of value, in order to

---

[28] Joel Feinberg, 'Legal Paternalism', *Canadian Journal of Philosophy*, vol. 1, no. 1, 1971, p. 113. Feinberg makes this suggestion with regard to legal paternalism only.

help people lead good lives. Laws and policies may favour ideals of the good life. Different theories of perfectionism construe the value of activities and the nature of the good life differently. Certain theories may condemn homosexuality, prostitution, and pornography and require that the state discourage these activities. Other theories may see nothing bad about them and thus not support such state intervention. Whilst perfectionist theories may have different conceptions of the good, they all share the feature that, according to them, the state does not have to be neutral towards these conceptions. It may favour those that are valid against those that are not. It is important to note that perfectionism holds that the state may favour only true conceptions of the good, not those that are simply thought to be true (by the state, or the majority or whatever).[29] Perfectionists would condemn a state that implements its view of the good life when that view is mistaken. Neutralists also of course condemn state implementation of false conceptions of the good but take the further step of also ruling out the implementation of true ones.

1.3.2 The above is a rough sketch of the notion of neutrality. There are several different conceptions of neutrality that try to make the notion more precise. The one that will be my focus is often termed justificatory neutrality.[30] According to it, the state's justifications for its actions may not appeal to the truth or falsity of any conception of the good. Arguments containing reasons, principles, values, etc. that invoke a conception of the good are excluded as illegitimate arguments for laws or policies of the state. They are excluded even if, in the absence of the neutrality requirement, they would be sound and strong arguments. A state is neutral, in the justificatory sense, if its actions are justified independently of any conception of the good.

---

[29] See for example Raz, *The Morality of Freedom*, p. 158 and Robert George, *Making Men Moral*, Clarendon Press, Oxford, 1996, pp. x, 7.

[30] The term is originally from Will Kymlicka, 'Liberal Individualism and Liberal Neutrality', *Ethics*, vol. 99, no. 4, 1989, p. 884.

Justificatory neutrality may be contrasted with other conceptions of neutrality, such as consequential neutrality. According to this, the state should ensure that its actions have neutral effects on the different conceptions of the good held by people. There are different accounts of what it means for the state to ensure that it is neutral in the consequences of its policies. One requires that state action must not make it more likely that people will endorse a particular conception of the good or be successful in their pursuit of the one they endorse, whilst another requires that the state should ensure everyone has an equal opportunity to pursue their chosen conception.[31] But they all take neutrality to be a property of the consequences of state action rather than of justifications for policies. A state could be neutral in the justificatory sense without being so in its consequences. Enforcing (neutrally justified) rights may have the effect of making some conceptions of the good much more difficult to pursue than others.

Justificatory neutrality also differs from the closely related notion of intentional neutrality. According to this, the state should not aim to promote any particular conception of the good. A state may be neutral in its intentions without being neutral in the justificatory sense. It may have policies which do not intentionally favour any conception of the good, but the state's supposed justification for these policies may be that it believes Quakerism is the true religion and that one of the principles of Quakerism is that the state should not aim to promote any conception of the good. The reverse could also be the case. A state may intentionally favour one religious view, thus violating intentional neutrality, but do so for the sake of stability and order, thereby remaining justificatorily neutral.[32] Some authors fail to distinguish justificatory from intentional

---

[31] These principles of consequential neutrality are taken from Raz, *The Morality of Freedom*, pp. 114-5.

[32] The points in this paragraph, and the accompanying examples, are from Richard J. Arneson, 'Neutrality and Utility', *Canadian Journal of Philosophy*, vol. 20, no. 2, 1990, pp. 218-9.

neutrality. Will Kymlicka, for example, contrasts consequential neutrality with a view which 'allows that government action may help some ways of life more than others but denies that government should act in order to help some ways of life over others.... [T]he desire to help one way of life over another is precluded as a justification of government action.'[33] Kymlicka calls this view justificatory neutrality, but whilst the second sentence seems to refer to that view as I defined it earlier, the first sentence indicates a different idea according to which the state is not to act with the intention of favouring some conception of the good even if that act may be neutrally justified.

Perfectionism rejects both justificatory and intentional neutrality. Perfectionists deny that state actions must be justified on neutral grounds only. The state may base its decisions on conceptions of the good. Perfectionists also claim that it is permissible for the state to aim to promote worthwhile conceptions of the good. Both perfectionists and neutralists usually reject consequential neutrality, at least in the sense that state action must have equal consequences on the different conceptions of the good held by citizens. Of the three conceptions of neutrality, justificatory neutrality is often taken to be the most plausible by both defenders and critics of neutrality. However, rather than consisting in a single correct conception from amongst the three described above, the general ideal of neutrality is probably made up of some combination of them. Justificatory neutrality is necessary for neutrality in general. It is necessary because intuitively, a state that justifies its policies on the grounds of some conception of the good cannot be thought of as neutral. But whilst justificatory neutrality is necessary, it cannot be sufficient. A state that enforces one religious view for the sake of the stability and cohesiveness of society

---

[33] Kymlicka, 'Liberal Individualism and Liberal Neutrality', pp. 883-4. George Sher also runs the two ideas together (see his *Beyond Neutrality: Perfectionism and Politics*, Cambridge, Cambridge University Press, 1997, p. 4), as does Thomas Hurka (see his *Perfectionism*, New York and Oxford, Oxford University Press, 1993, p. 159). Stephen Wall separates them (see Wall, *Liberalism, Perfectionism, and Restraint*, Cambridge, Cambridge University Press, 1998, pp. 3, 32).

could not plausibly be described as neutral. It must violate the ideal in some other way, perhaps by intending to help one conception of the good, or perhaps by being non-neutral in its consequences on different conceptions, or perhaps both.

Contrary to many theorists' beliefs, the ideal of neutrality may include a principle of consequential neutrality of some form. Some writers are too quick to reject consequential neutrality. They argue that since it is impossible or undesirable for that state to ensure that every conception of the good fares equally well or that every person does equally well in fulfilling his or her conception of the good, the state should not attempt to achieve neutrality in its consequences but instead, if neutrality is a legitimate aim at all, adhere to justificatory or intentional neutrality.[34] But whilst some versions of consequential neutrality may be impossible or undesirable to implement, others may not be. Equality of opportunity to pursue one's conception of the good (or at least equal opportunity to pursue conceptions that do not conflict with the requirements of justice), for example, avoids the objection.[35] Pointing out the non sequitur that critics of consequential neutrality make, however, does not establish that consequential neutrality is a part of the ideal of neutrality in general. Such a condition may go some way toward explaining why a state that enforces a religion, even for neutral reasons, is nevertheless not a neutral state. But intentional neutrality may provide the explanation instead, since such a policy also obviously aims to help a particular conception of the good.

I will not, however, set out the conditions for neutrality in general. Rather, justificatory neutrality will be the focus of this thesis. My concern is with investigating justificatory

---

[34] For this argument see Kymlicka, 'Liberal Neutrality and Liberal Individualism' p. 884; Sher, *Beyond Neutrality*, pp. 3-4; Charles E. Larmore, *Patterns of Moral Complexity*, Cambridge, Cambridge University Press, 1987, pp. 43-4.

[35] Rawls claims his theory achieves such equal opportunity. See John Rawls, *Political Liberalism*, New York, Columbia University Press, 1993, pp. 192-3.

neutrality's and justificatory perfectionism's implications for paternalism. This is because it seems a plausible candidate for what neutrality requires (even though not a complete explanation) and because of the attention it has been given in the literature.

1.3.3 That justificatory neutrality has been given much attention by both critics and defenders of neutrality can be easily shown. Some theorists use different terms for the idea but the notion seems the same. Amongst its critics, Joseph Raz rejects the idea that

> governments should be blind to the truth or falsity of moral ideals, or conceptions of the good.... [N]either the validity, cogency or truth of any conception of the good, nor the falsity, invalidity or stupidity of any other may be a reason for any governmental action.[36]

Raz does not call this neutrality, reserving that label for consequential neutrality only, but the idea nevertheless seems the same.

Ronald Dworkin is another critic in at least one of his major works.[37] He distinguishes two views of political morality. One holds that conceptions of the good be

> bracketed or set aside on political occasions. Political life is to proceed prescinding from ethical issues of that sort; whatever one might think about the structure or character of the good life is simply not relevant, one way or the other in one's political activity or decision.[38]

On the contrary view, which I am calling perfectionism, all

---

[36] Raz, *The Morality of Freedom*, p. 108.

[37] In earlier works he was a defender of the idea that 'political decisions must be, so far as is possible, independent of any conception of the good life, or of what gives value to life.' See Ronald Dworkin, 'Liberalism', in his *A Matter of Principle*, Cambridge, Massachusetts, Harvard University Press, 1985, p. 191.

[38] Ronald Dworkin, 'Foundations of Liberal Equality', in Grethe B. Peterson (ed.), *The Tanner Lectures on Human Values*, IX, Salt Lake City, University of Utah Press, 1990, p. 21.

> ethical convictions are available in politics.... On this view, ethics and politics are intertwined so that some of the most far-reaching questions about the character of the good life are political questions too.[39]

Dworkin calls these views the discontinuity and continuity strategies, respectively. Again, the idea of discontinuity seems to be justificatory neutrality under a different label.

Thomas Hurka is another critic of the idea that 'the state must never ... have as its justification for acting that some ways of life are intrinsically preferable to others.'[40] Stephen Wall criticises the idea that political arguments should not be based 'on reasons or considerations that are controversial or not publicly accessible.'[41] Wall writes here about ideas that are controversial or publicly inaccessible rather than conceptions of the good. Whilst this makes an important difference for the arguments he considers, he seems to believe that most conceptions of the good fall into this category. What is significant here is that the notion he has in mind is one that places a restriction on the range of justifications available in political morality.

Justificatory neutrality also has its defenders. Charles Larmore rejects consequential neutrality and supports

> a constraint on what factors can be invoked to justify a political decision. Such a decision can count as neutral only if it can be justified without appealing to the presumed intrinsic superiority of any conception of the good life.[42]

---

[39] Dworkin, 'Foundations of Liberal Equality', p. 21.
[40] Hurka, *Perfectionism*, p. 159.
[41] Wall, *Liberalism, Perfectionism, and Restraint,* p. 32.
[42] Larmore, *Patterns of Moral Complexity*, p. 44.

Brian Barry argues that 'at the point where the basic principles and rules are being drawn up ... nobody is to be allowed to assert the superiority of his own conception of the good over those of others'[43] as a reason for those principles and rules. Principles of basic political morality (but not every law and policy) have 'to be argued about in ways that avoid invoking any conception of the good.'[44] His own theory of justice, Barry claims, is neutral in the sense that it 'does not rest on any particular conception of the good.'[45]

John Rawls seems to have the notion of justificatory neutrality in mind when he argues that his conception of justice is political and not comprehensive. He claims that one of the features of a political conception of justice is that accepting it 'does not presuppose accepting any particular comprehensive religious, philosophical, or moral doctrine.'[46] A doctrine is comprehensive 'when it includes conceptions of what is of value in human life, as well as ideals of personal virtue and character.'[47] This seems to imply that a political conception of justice is one that can be given a justification that does not appeal to a conception of the good. Since Rawls thinks that only political conceptions are appropriate for modern democratic societies, this would entail the claim that such societies should be justificatorily neutral.[48]

1.3.4 Respecting the neutrality constraint has a significant effect on the way political decisions are arrived at. To see this more clearly, consider the example of a law that

---

[43] Brian Barry, *Justice as Impartiality*, Oxford, Clarendon Press, 1995, p. 160.
[44] Barry, *Justice as Impartiality*, p. 143.
[45] Barry, *Justice as Impartiality*, p. 123.
[46] Rawls, *Political Liberalism*, p. 175.
[47] Rawls, *Political Liberalism*, p. 175.
[48] Some of Rawls's other comments seem to imply that he does not accept justificatory neutrality. For example, he says that 'we want a political conception to have a justification by reference to one or more comprehensive doctrines' (*Political Liberalism*, p. 12). However, that some policies may have perfectionist justifications does not mean that they can not also have neutral justifications. It is only if a policy's *only* justifications are perfectionist ones, and no possible neutral justification exists, that such a policy would violate neutrality. What Rawls seems to want is a neutrally justifiable conception of justice which is also supported by perfectionist reasons in order to ensure stability and 'overlapping consensus'.

prohibits the use and possession of certain types of drugs. Such a law has various reasons and arguments for and against it. Some of these reasons and arguments, on both sides of the issue, appeal to conceptions of the good and some do not. Rejecting neutrality would mean that all reasons are available to the state in justifying its policies. Suppose that even in the absence of the neutrality constraint there is a neutral reason against such a law that outweighs the reasons derived from conceptions of the good that support the prohibition. The outcome in this case, that there ought not be such a law, is the same as when the neutrality constraint is applied. But the reasoning is different. In the one case, where all reasons are available, perfectionist reasons are outweighed by neutral ones. In the other, where neutrality applies, perfectionist reasons are excluded by the neutrality constraint; neutral reasons are the only relevant ones for deciding the issue.

Given this effect on political reasoning, even though justificatory neutrality is a property of the justifications for state action rather than consequences of laws and policies, it has practical implications. According to it, the state must have neutral justifications for its actions. If no neutral justification of sufficient strength for a policy can be provided, then according to neutrality, that policy should not be undertaken. Any law or policy that can be justified only on perfectionist grounds is illegitimate. For example, if there is no sufficiently strong neutral justification for state funding of the arts, then such a policy must not be implemented even if there is a strong perfectionist justification for the funding.[49]

George Sher argues that without a clause requiring that the justifications at least be plausible, the account of justificatory neutrality is inadequate. Any given law or policy

---

[49] For suggested neutral arguments for state funding of the arts, see John Rawls, *A Theory of Justice*, pp. 331-2, Ronald Dworkin, 'Can a Liberal State Support Art?' in his *A Matter of Principle*, Harvard University Press, Cambridge, Massachusetts, 1985, and Harry Brighouse, 'Neutrality, Publicity, and State Funding of the Arts', *Philosophy and Public Affairs*, vol. 24, no. 1, 1995.

that can be shown to be neutral by simply supporting it by some neutral argument, no matter how plausible that argument is.[50] However, making plausibility a condition of neutrality results in difficulties. It should be possible to say that the state's justification for a policy, even though implausible, nevertheless does meet the criteria of neutrality. Imagine that two people are arguing about public policy. One might say to the other "I agree that the state's justification is neutral, but it is nevertheless implausible." According to Sher's proposal, that person has misunderstood the notion of neutrality.

For this reason, we should reject that proposal. But we are faced with Sher's objection that neutrality is too easily achieved. However, it may not be as easy to achieve as he thinks. Many normative premises that appear to be neutral may presuppose a conception of the good. For example, claiming that government policy will advance economic growth may presuppose a belief in the value of economic growth that stems from a materialistic conception of the good. This would violate neutrality if there is no other way of supporting the value of economic growth that does not presuppose a conception of the good. There probably is such a way (e.g. that economic growth is generally beneficial to everyone no matter what their conceptions of the good), but the example illustrates the point that what may seem like a neutral argument may in fact not be.

The force of Sher's objection is not what it seems for another reason. It should be remembered that neutrality is not the only requirement that a neutralist may set on legitimate state action. A further requirement could be that the arguments be sound ones. This makes soundness external to neutrality, rather than internal as Sher believes, but nevertheless a requirement on state action.[51] Once this requirement comes into play, the

---

[50] Sher, *Beyond Neutrality*, p. 25.

[51] Other requirements could be that the arguments could not be reasonably rejected by anyone (Barry's view in *Justice as Impartiality*) and/or that they be able to be the subject of an overlapping consensus in society (Rawls's view in *Political Liberalism*).

force of the neutrality requirement becomes more apparent. The soundness requirement will eliminate many of the arguments for a given state action. Of those that are left, some will appeal to a conception of the good and some will not. The neutrality requirement will rule out the former. This will have the effect of leaving some laws and policies without legitimate justifications if the only sound arguments for them are nonneutral ones. Seen in this light, the neutrality requirement has more force that Sher's objection implies. Hence, a state satisfies the requirement of justificatory neutrality so long as its justifications for its policies are neutral, regardless of the soundness or plausibility of those justifications. Justificatory neutrality is necessary for legitimate state action, neutralists claim, but they need not claim (nor would it be plausible to do so) that it is sufficient for it.

1.3.5 There are numerous questions concerning the scope of justificatory neutrality (henceforth, simply neutrality). A first set of questions concern what neutrality requires neutrality towards. As stated above, it does not require neutrality towards all values or normative beliefs, only a class of them known as conceptions of the good. The requirements of neutrality will thus turn upon what counts as a conception of the good. A conception of the good is a view about what lifestyles or activities it would be valuable or worthwhile to pursue. Such a view may be carefully planned (a step-by-step plan for one's career) or mostly unreflective (spending most of one's life surfing). It may be life-encompassing (full time devotion to a religion) or only involve decisions for a small part of one's life (attending church service every Christmas). It may be substantive (appreciation of Van Gogh paintings, being a carpenter) or highly formal and abstract (contemplation of beauty, leading one's life autonomously).[52] Hence, conceptions of the

[52] This account of conceptions of the good is similar to the account of 'ideals of human flourishing' given by Steven Wall in *Liberalism, Perfectionism, and Restraint*, p. 12. For other accounts of conceptions of the good, see Rawls, *Political Liberalism*, pp. 19-20, Raz, *The Morality of Freedom*, p. 135, and Sher, *Beyond Neutrality*, pp. 37-43.

good do not include all normative beliefs (claims about justice and rights are not included, although some of them may be based on a conception of the good) nor even all beliefs about the good (primary goods, goods that are useful in the pursuit of a wide range of conceptions, such as power, wealth, opportunities, and health, are not included, though beliefs about how one should rank preferences for them may be and a belief that one should attempt to maximise one's primary goods is).

Some views of neutrality do not require neutrality towards all conceptions of the good, but only controversial conceptions.[53] If a conception of the good is generally shared by all or most in society, according to the latter view, justifications for state action may legitimately appeal to it. Differently (and possibily in conjuntion), the neutralist could be neutral towards only substantive conceptions but non-neutral towards formal conceptions. For example, the state may justify its policies by appealing to a view that conceives of the good as preference satisfaction. Preference satisfaction is a formal notion that does not determine the substantive aims of people.[54] The state would then be non-neutral towards formal conceptions of the good since it favours one of them, but remain neutral towards other conceptions so long as it does not attempt to justify its policies by appeal to more substantive values. Neutrality in the widest sense would prohibit the state from relying on any perfectionist considerations whether controversial or not, and whether formal or substantive.

---

[53] Charles Larmore defends neutrality towards controversial conceptions of the good. (See Larmore, *Patterns of Moral Complexity*, p. 67, and Charles Larmore, *Morals of Modernity*, Cambridge, Cambridge University Press, 1996, chapter 6.) In *Justice as Impartiality*, Brian Barry seems to defend neutrality towards all conceptions of the good whether controversial or not.

[54] This is the view of neutrality that Arneson argues for. See 'Neutrality and Utility', pp. 220-2. Neutrality between substantive but not formal conceptions of the good is labeled 'comprehensive anti-perfectionism' by Stephen Mulhall and Adam Swift in their *Liberals and Communitarians*, Blackwells, Oxford, second edition, 1996, pp. 252-5. Such a position, they say, holds that justifications for state action should not be based on any substantive conception of the good but justifies this anti-perfectionism by appeal to a conception of the good. (p. 253.) As Mulhall and Swift suggest, this latter conception must be formal not substantive since otherwise the position would be self-contradictory.

Another question of scope is whether the requirement of neutrality applies to all societies or only a subset of them. Rawls, for example, confines his arguments to democratic societies,[55] whilst Barry thinks his case for neutrality is 'universally valid.'[56] Steven Wall claims that neutrality confines itself to modern western societies (and similarly restricts his perfectionist political morality that rejects it).[57]

A further question is whether neutrality applies to all political actions or only certain classes of them. Again, Rawls claims that neutrality is a requirement only with respect to 'constitutional essentials and matters of basic justice' but does not apply to 'many economic and social issues that legislative bodies must regularly consider.'[58] With regard to the latter, political decisions may invoke a conception of the good in their justification. Barry also admits that neutrality applies to only some but not all political decisions (although he thinks a different type of neutrality applies to the others).[59]

It is also unclear whom the neutrality requirement applies to. Does it impose duties on citizens as well as duties on the state? Dworkin and Wall take it to constrain people on how they may legitimately vote and argue for political decisions.[60] This need not be the case. One could accept state neutrality but think that citizens should be free to appeal to their conceptions of the good when voting, participating in local politics, etc. Laws or policies that result from these actions could be struck down by a judiciary that enforces state neutrality.

---

[55] See for example pp. 3, 11, and 13 of Rawls, *Political Liberalism*.
[56] Barry, *Justice as Impartiality*, p. 3.
[57] Wall, *Liberalism, Perfectionism, and Restraint*, p. 3.
[58] Rawls, *Political Liberalism*, p. 230. See also pp. 214-5.
[59] Barry, *Justice as Impartiality*, pp. 143-5, 160-1.
[60] Dworkin, 'Foundations of Liberal Equality', p. 20; Wall, *Liberalism, Perfectionism, and Restraint*, pp. 31-2.

Investigation of these questions would take up too much space here. This is partly because answers to many of them depend upon arguments about neutrality. Arguing for neutrality on the grounds of scepticism about the good, for example, would seem to require neutrality in all the state's actions, whilst distrust of the coercive imposition of conceptions of the good would confine neutrality to coercive measures.[61] I hope to arrive at partial answers to some of them when examining arguments in subsequent chapters. Whilst there is not space here to answer all these questions, I will assume that neutrality requires the state to refrain from appealing to conceptions of the good in most of its decisions.

4. How Paternalism Relates to Neutrality and Perfectionism

With these definitions in mind, we can now see how paternalism relates to neutrality and perfectionism. Perfectionism need not be paternalistic. As I argued in section two, the essence of paternalism is that it (1) aims to close an option or make a choice for a person and (2) it is intended for the person's own good. Some perfectionist policies do not satisfy one or both of these criteria. State action motivated by perfectionist considerations that expands rather than restricts a person's options, such as making opera available to those who cannot afford it, is a type of non-paternalistic perfectionism. Another type, where the second clause is not met, is state restriction of freedom for the perfectionist good of others, such as imprisoning people for blaspheming against the true religion. Both conditions are not met when the state makes funds available for artists (thereby increasing their options) for the good of the artists' audiences. Perfectionist paternalism occurs when the paternalist intervenes with the paternalised for the latter's own good and the paternalist's justification for the intervention appeals to a conception of

---

[61] This point is discussed more fully by Jeremy Waldron in 'Legislation and moral neutrality', in Robert E. Goodin and Andrew Reeve, *Liberal Neutrality*, Routledge, London and New York, 1989.

the good. Examples could include laws against pornography, prostitution, euthanasia, drugs, incest, listening to heavy metal, watching TV gameshows, and not attending weekly character-improvement sessions.

Whilst perfectionism does not have to be paternalistic, it provides motivation for a wide range of paternalistic policies. Since people would be better off having worthwhile conceptions of the good rather than worthless ones, a perfectionist state could force people into worthwhile activities and lifestyles for their own good. Neutrality does not face the same problem. The neutral state may have some paternalistic policies. Neutral paternalism intervenes for the target's own good, but does so without appealing to a conception of the good. There seem to be two main ways this can occur.[62] The first is when the paternalised's good is conceived in a general way that is compatible with a wide range of conceptions of the good. Hygiene regulations for food manufacturers are an example of this since consumers lose the option of buying food not subject to such regulations (which would be at a lower price), but few if any conceptions of the good would require not avoiding contaminated food. Other examples could include government safety regulations for motor vehicles and machinery, water fluoridation, and requirements for labeling ingredients and nutritional information on food products. These policies all close some option for people's own good. But the goods that these policies protect are compatible with a wide range of conceptions of the good. The paternalist's rationale does not refer to a conception of the good.

The second type of neutral paternalism appeals to more specific goods, but these have to be goods accepted by the paternalised. They are part of his conception of the good.

---

[62] The two types of neutral paternalism are suggested by Rawls's discussion of paternalism in *A Theory of Justice*, Oxford, Oxford University Press, 1971, pp. 248-50. There are conceptual difficulties with both types, concerning whether they are truly neutral or even types of paternalism. I try to address these difficulties in chapter six.

Justification of the intervention may not appeal to a conception that he does not accept. Instead, the justification is that the paternalist is helping the paternalised to achieve his own conception of the good. A father's forcing his daughter to attend the best law school, when the daughter believes being a lawyer is the best career for her, in order to assist her realisation of her career goal, is a case of this type of neutral paternalism. Paternalistic policies could be perfectionist and neutral at the same time. Laws requiring the wearing of helmets while riding on motorcycles are probably neutral towards riders whose conceptions of the good do not include tempting death but perfectionist towards those who value feeling the wind in their hair. If some way can be found to apply the laws only to the former, then they would be cases of neutral paternalism.

Whilst neutrality may be compatible with some paternalism, it does not face the same difficulty that perfectionism does, namely supporting the forceable imposition of conceptions of the good. Because perfectionism holds that some ways of life are better than others, and holds that the state may have such judgements as justifications for policies, it provides a rationale for paternalistic policies that force people into worthwhile conceptions of the good. Many such policies, such as forcing people to experience great art or to attend weekly character-improvement sessions, seem objectionable. Neutralist theories do not face this difficulty since they rule out perfectionist paternalism along with all perfectionism. The neutrality constraint claims that all perfectionist paternalism is illegitimate. The only acceptable form of paternalism, according to it, is neutral paternalism. By excluding a wide range of reasons for state action (those that appeal to conceptions of the good) a neutral state simply does not have the motivation that a perfectionist one does. This is a considerable difficulty for perfectionism. Insofar as it is desirable for a theory of the state to account for instances of paternalism that seem unacceptable, we have some reason to reject a theory that supports unacceptable paternalism. It would not be a decisive objection against perfectionism, since our

intuitions about the wrongness of paternalism could simply be mistaken. But it is a difficulty that defenders of perfectionism have thought worth attempting to rebut. They have a number of strategies to resist the paternalistic implications of their theories. One is to appeal to the value of autonomy as a component of the good life. Another is to rely on the notion that no person's life can be improved against his considered judgement that it is not. These and other strategies are examined in the next four chapters.

Chapter Two

Perfectionist Autonomy as Part of Well-Being

2.1 Introduction

One of the most common strategies used by perfectionists to resist the charge against their theories of being overly paternalistic is to appeal to the value of autonomy. They claim that autonomy is an important feature of personal well-being and that paternalistic interventions would undermine this feature. Hence, the argument goes, such interventions would fail to improve the value of the lives of the paternalised, thus making paternalism self-defeating. This chapter examines arguments from the value of autonomy. One obvious question about the arguments is whether it is true that paternalism always undermines autonomy. The answer depends upon what autonomy consists in. If, say, it is simply freedom from coercion then paternalism need not interfere with autonomy since as we saw in the first chapter, paternalistic interventions need not be coercive. I set out the perfectionist conception of autonomy below. Whilst it does not rule out all paternalism, autonomy does I think, conflict with the most objectionable forms of paternalism.[1] I instead focus on a separate question: given that paternalism undermines autonomy, why does this matter? It only matters if autonomy has value and is valuable to such an extent as to make paternalism unacceptable. Only then will perfectionism be able to resist its

[1] For a discussion of the incompatibility of paternalism and autonomy, see Douglas Husak, 'Paternalism and Autonomy', *Philosophy and Public Affairs*, vol. 10, no. 1, 1980. Husak argues that the two are not necessarily incompatible as is widely believed. But he does not show what would be more problematic for someone trying to rule out paternalism on grounds of autonomy, namely that the more objectionable cases of paternalism do not conflict with autonomy, or that the value of autonomy is not significant enough to rule out these cases.

paternalistic implications. Hence this chapter is concerned with arguments for the value of autonomy.

There is a huge philosophical literature on autonomy, with many different understandings of the concept and arguments for its importance.[2] But the scope of the chapter is narrowed by focusing only on a particular conception of autonomy and examining only a particular type of argument for its value. I will be concerned only with the perfectionist conception of autonomy, which is set out in the following section. And I will examine only arguments that attempt to show that autonomy contributes to the well-being of the person who has it. Such arguments hold that the autonomous life is an ideal way of living, that self-direction enhances one's personal well-being.

Isolating the understanding of autonomy that is part of the perfectionist project and focusing only on that understanding is a useful strategy for several reasons. Firstly it avoids the difficulty that a perfectionist appeal to autonomy could be self-defeating because autonomy entails neutrality. For the state to respect people's autonomy, so an argument could go, it must refrain from basing its decisions on conceptions of the good. If it does so, it risks imposing a conception of the good on those who do not accept it and this undermines autonomy, at least to the extent that autonomy partly consists in leading one's life according to one's own reasons and values. The argument is given added weight by the fact that autonomy seems to play an important role in the arguments of several neutralist theorists. Rawls, for example, claims that situation of the hypothetical contractors in his theory, which he believes supports neutrality, expresses the ideal that people are 'primarily moral persons with an equal right to choose their mode of life'[3] and possess 'full autonomy'[4]. David Richards also

---

[2] For a useful starting point, see Gerald Dworkin, *The Theory and Practice of Autonomy,* Cambridge, Cambridge University Press, 1988, especially chapter one.

[3] John Rawls, *Theory of Justice*, Oxford, Oxford University Press, 1971, p. 563.

[4] John Rawls, *Political Liberalism*, New York, Columbia University Press, 1993, p. 77.

argues that neutrality among a wide number of visions of the good life has its source in the value of autonomy.[5] If these theorists are correct, then a perfectionist appeal to the value of autonomy would be self-defeating since it supports abandoning perfectionism in favour of neutrality.

The solution to this difficulty lies in distinguishing between two different understandings of the normative significance of autonomy, only one of which may imply neutrality. One understanding of autonomy sees it as an ideal for how people should live their lives. Autonomous self-direction, on this view, makes the life that contains it a better one. Autonomy is valuable for this reason. On the other understanding, autonomy places a constraint on how people may be treated. Autonomous self-direction, on this view, must be respected independently of whether doing so improves the goodness of the lives of the autonomous. Complying with the moral rule forbidding interferences with people's autonomy may have the result that they are better off, but this is not the reason for so complying. I follow Sher in calling these notions the value of autonomy and respect for autonomy, respectively.[6]

Respect for autonomy may support a neutrality constraint of some sort. It requires the state to ignore certain reasons in some sphere. The state must respect people's autonomous choices regardless of the quality and correctness of those choices. Respecting autonomy requires ignoring the truth or falsity of a person's conception of

---

[5] See David A. J. Richards, 'Human Rights and Moral Ideals', *Social Theory and Practice*, vol. 5, nos. 3-4, 1980, p. 474; 'Rights and Autonomy', *Ethics*, vol. 92, no. 1, 1981, especially p. 9; *Sex, Drugs, Death, and the Law*, Rowman and Littlefield, Totowa, New Jersey, 1982, especially chapter 1, p. 9; *Toleration and the Constitution*, Oxford University Press, New York and Oxford, 1986.

[6] George Sher, *Beyond Neutrality*, Cambridge, Cambridge University Press, 1997, p. 16. As Steven Wall points out, there are several other understandings of autonomy that are not of concern here. One is the idea of metaphysical autonomy concerning the nature and possibility of free will. Another is Kantian autonomy, concerning the conformity of the will with true moral rules of conduct. Wall also separates the idea of what Sher calls respect for autonomy from the perfectionist conception of autonomy. See Steven Wall, *Liberalism, Perfectionism, and Restraint*, Cambridge, Cambridge University Press, 1998, pp. 127-8. For further taxonomies of conceptions of autonomy, see John Christman, 'Constructing the Inner Citadel: Recent Work on the Concept of Autonomy', *Ethics*, vol. 99, no. 1, 1988 and Joel Feinberg, *Harm to Self*, New York and Oxford, Oxford University Press, 1986, chapter 18.

the good as a reason for deciding how to treat him, at least when his autonomy is at stake. If perfectionism were to rely on the argument of respect for autonomy, it may be embracing neutrality (at least in the sphere in which respect for autonomy applies).[7]

But this not so if the perfectionist appeals to the value of autonomy. That understanding of autonomy does not have the exclusionary nature of the respect for autonomy understanding. Appealing to the value of autonomy does not require the state to ignore conceptions of the good in its decisions. On the contrary, it assumes that autonomy is one conception of the good that the state should be guided by. And it is compatible with autonomy that the state should be guided by other conceptions of the good as well as autonomy. The claim is that self-direction makes a person's life more valuable and is of such value that it allows perfectionism to resist objectionably paternalistic implications. The arguments I will examine try to show that the value of autonomy either outweighs other values or is a condition of other values. But they do not claim that autonomy excludes other values which is what neutrality requires. Hence, perfectionism may appeal to the value of autonomy without being self-defeating since the value of autonomy does not entail neutrality.

Another advantage of perfectionism's reliance upon arguments for the value of autonomy rather than respect for it is that if such arguments succeed, they defeat paternalism at its source. One necessary condition for justified paternalism is that it at least be successful. Interventions for a person's own good are permissible only if they succeed in advancing that person's good. If they make the person worse off, then despite their good intentions, the interventions are unwarranted.[8] Arguments for the

---

[7] Here I am not claiming that the argument for neutrality from respect for autonomy is a sound one, just that the argument has the right form to support neutrality. An argument from the value of autonomy does not even have that. For arguments that autonomy, under any conception, fail to establish neutrality, see Andrew Mason 'Autonomy, Liberalism and State Neutrality', *Philosophical Quarterly,* vol. 40, no. 161, 1990.

[8] There may be some exceptions to this rule, say where it was extremely unlikely that the intervention could have been thought to fail to make the person better off. Freakish circumstances made the

value of autonomy as part of well-being attempt to show that paternalism fails even this minimal condition. They try to show that in virtue of autonomy's contribution to well-being, paternalistic interventions make people worse, not better off. They thereby rule out paternalism at its source. Other types of arguments for autonomy against paternalism face a greater burden. They have to show that paternalism is unjustified even if it makes the paternalised better off. For example, appealing to arguments that rely upon the notion of respect for autonomy requires showing that the rule prohibiting interferences with people's autonomy is always or usually sufficiently strong to overcome the gain in well-being that the interference might produce. The arguments I consider below do not face this added burden.

Connected with this is another reason why perfectionists should rely upon the value of autonomy rather than respect for it. The value of autonomy is itself a perfectionist claim, a claim that autonomy is, or is part of, a worthwhile way of life. Appealing to it thus appeals to a perfectionist reason for or against state policy. Respect for autonomy on the other hand, is not itself a claim about the good. It posits a rule that should be followed irrespective of it whether doing so furthers people's well-being. Appealing to it thus appeals to a non-perfectionist reason against paternalism. The task of this and the following three chapters is to see what reasons against paternalism can be provided from within perfectionism itself. The more perfectionism relies upon non-perfectionist reasons for its conclusions, the less it is a distinctive theory. Relying on respect for autonomy to resist paternalism would make perfectionism less of a distinctive theory.[9] The alternative (other than trying to derive respect for

---

paternalism unsuccessful, where it would normally be certain to succeed. We may nevertheless want to say that the intervention was justified. I set aside these exception cases here.

[9] For arguments against paternalism that appeal to respect for autonomy, see Donald VanDeVeer, *Paternalistic Intervention*. Princeton, New Jersey, Princeton University Press, 1986 and Joel Feinberg, *The Moral Limits of the Criminal Law, volume three: Harm to Self*. New York and Oxford, Oxford University Press, 1986. Both books are discussed and criticised by Dan W. Brock in 'Paternalism and Autonomy', *Ethics*, vol. 98, no. 3, 1988. Others who argue against paternalism from respect for autonomy include Rawls and Richards. See the references in footnotes 3, 4, and 5. Some non-perfectionist arguments against paternalism that perfectionists could use are considered in chapter 7.

autonomy from perfectionist considerations), appealing to the value of autonomy, is a perfectionist reason. The arguments I consider below all contain such appeals.

In the next section, I set out the perfectionist conception of autonomy out more fully. Sections 2.3 through 2.7 examine the arguments for the value of this understanding of autonomy. I argue that none of them succeed, or at least not in a way that enables perfectionism to avoid objectionable paternalism. Section 2.8 tries to bolster this by showing that autonomy is only valuable when exercised in the pursuit of worthwhile activities. This means that the value of autonomy does not rule out paternalism that blocks the pursuit of worthless activities.

2.2 The Perfectionist Conception of Autonomy

Before proceeding to the arguments, it is useful to set out the perfectionist conception of autonomy more fully. Thomas Hurka captures the essence of the notion when he defines it simply as 'free choice from many life options'[10]. Joseph Raz gives a description of the underlying ideal. According to Raz, personal autonomy is the idea that

> people should make their own lives. The autonomous person is a (part) author of his own life. The ideal of personal autonomy is the vision of people controlling, to some degree, their own destiny, fashioning it through successive decisions throughout their lives.[11]

Similarly, Steven Wall claims that personal autonomy is

> the ideal of people charting their own course through life, fashioning their character by self-consciously choosing projects and taking up commitments from a wide range of eligible alternatives, and making something out of their lives according to their own understanding of what is valuable and worth doing.... They neither drift through life, aimlessly moving from one object of desire to another, nor adopt projects and pursuits wholesale from others.[12]

---

[10] Thomas Hurka, *Perfectionism*, New York and Oxford, Oxford University Press, 1993, p. 148.

[11] Joseph Raz, *The Morality of Freedom*. Oxford, Clarendon Press, 1986, p. 369.

[12] Wall, *Liberalism, Perfectionism, and Restraint*, p. 128.

These descriptions make it clear that to be autonomous, a person must have certain capacities and characteristics. She must be aware of the various options she faces and how choosing amongst them would affect her and her projects and activities.[13] She must be able to 'absorb, remember and use information, reasoning abilities and the like'[14] to some degree. She must also have the capacity to form intentions, plan their execution, make decisions, and be loyal to those decisions once made.[15] She must be able to pursue projects and activities, form personal attachments, and maintain relationships.[16] The autonomous person must also have a degree of independent-mindedness. She forms her own judgements of how to live rather than simply following others' lead without reflection.[17]

A person must exercise these capacities in order to lead an autonomous life. Two conditions are needed for a person to be able to exercise his capacities for autonomy. The first is that he must have an adequate range of options to choose from.[18] Someone who has only trivial and short-term options negligible in their effects does not have an adequate range of options. But a person whose every decision is shaped by her prospects of survival also lacks an adequate range. To be adequate, the option set must have 'long term pervasive consequences as well as short term options of little consequence, and a fair spread in between.'[19] Adequacy also depends on the variety of options available rather than simply their number. A choice between hundreds of identical options is inadequate. Another requirement is that there be some morally acceptable options available to a person. If a person chooses one option because all others would have immoral consequences, that person has not acted autonomously.

---

[13] Raz, *The Morality of Freedom*, p. 371, Wall, *Liberalism, Perfectionism, and Restraint*, p. 132.
[14] Raz, *The Morality of Freedom*, p. 408.
[15] Raz, *The Morality of Freedom*, 372, 408, Wall, *Liberalism, Perfectionism and Restraint*, pp. 132-3.
[16] Raz, *The Morality of Freedom*, p. 408.
[17] Wall, *Liberalism, Perfectionism and Restraint*, pp. 136-7.
[18] The description of adequacy given in this paragraph draws on points made by Raz in *The Morality of Freedom*, p. 373-9 and Wall, *Liberalism, Perfectionism and Restraint*, p. 141-3.
[19] Raz, *The Morality of Freedom*, p. 374.

The option set must also include options that allow a person to exercise the capacities that human beings have an innate drive to exercise and options that allow a person to develop all her capacities or to choose not to. Additionally, the choice should not be dominated by the need to protect one's life or one's important goals or projects. A choice dominated by the need to protect my career is not an autonomous one. An option set is adequate to the extent that it meets these criteria, and a person faced with this option set has an important requirement for being able to live an autonomous life.

The second condition needed for a person to exercise his capacities for autonomy is independence, that is, freedom from coercion and manipulation.[20] Coercion occurs whenever a person threatens another with some harm unless that other person performs some action or refrains from doing so. Manipulation, instead of interfering with a person's options, interferes with the capacities of a person to reach decisions, form preferences, or adopt goals. Whilst coercion and manipulation may invade autonomy by reducing the adequacy of options, they also invade autonomy independently of their effect on the option set. A person may be coerced or manipulated not to pursue one option while being left with an adequate range of others to choose from. Coercion and manipulation invade autonomy by subjecting the will of one person to that of another.[21] By doing so, according to Raz, they express 'disregard or even contempt for the coerced or manipulated people.'[22]

These conditions are not themselves sufficient for an autonomous life. Even if the requirements are met, a person may fail to exercise her capacity for autonomy. Such a

---

[20] For discussions of independence, see Raz, *The Morality of Freedom*, pp. 377-8 and Wall, *Liberalism, Perfectionism, and Restraint*, pp. 133-6.

[21] Raz, *The Morality of Freedom*, p. 377. See also Wall, *Liberalism, Perfectionism and Restraint*, p. 134.

[22] Raz, *The Morality of Freedom*, p. 378. Wall's explanation of how invasions of independence reduce autonomy differs from Raz's. For Raz, coercion and manipulation invade autonomy simply because of the type of treatment they involve, regardless of their consequences. For Wall, the type of treatment they involve is likely to give rise to feelings of resentment which will damage the victim's 'sense of self-worth, thereby undermining or weakening his resolve to pursue his projects and carry on with his life.' (Wall, *Liberalism, Perfectionism, and Restraint*, p. 135.)

person fails to live an autonomous life despite having the capacity and the environment in which she may do so. An autonomous person uses her abilities to choose, free from coercion and manipulation, amongst her range of options and thereby shape her own life.

Many instances of paternalism would undermine the autonomy of the paternalised. Paternalism could damage the capacities required for autonomy, or reduce the set of options below adequacy, or use coercive or manipulative methods. Some cases of paternalism may not do any of these things. Taxes on cigarettes, for example, do not. But if all invasions of autonomy are impermissible, then appealing to it would provide an effective method by which perfectionism can avoid many, though not all, of its paternalistic implications. Whether all such invasions are impermissible, however, depends upon the value of autonomy. If self-direction does not contribute to the value of one's life, then autonomy itself does not provide any reasons against (autonomy-invading) paternalism. Hence we need to examine perfectionists' arguments for the value of autonomy in order to assess its plausibility as a defence against paternalism.

2.3 Appeal to Intuitions

The first argument to be considered is the most straightforward. Wall claims that many intuitions about particular cases are best explained by the value of autonomy. His first set of cases concerns the independence aspect of autonomy, which prohibits coercion and manipulation. Wall asks us to suppose that you are wise and have an excellent understanding of what would be for my own good and are concerned to further my well-being. Would my life go better of I let you take control of it? Wall answers in the negative, saying that my life would not be a fully good one. This, he says, is because it is intrinsically good for people to take charge of their own lives.[23]

[23] Wall, *Liberalism, Perfectionism, and Restraint*, p. 146.

Wall considers a second case. Person A chooses projects that suit his nature and reflect his understanding of what is worthwhile and he leads a good, morally decent life. Person B finds decision-making irksome and lets a friend, who has the requisite wisdom, knowledge, and good will to make these decisions for her, take over her affairs. She also leads a good, morally decent life. Do A and B lead equally good lives? Wall says no. Even if B leads a better life than someone who does a worse job of running their own affairs, we think B lacks something important that A has. This is because autonomy has intrinsic value.[24]

The two cases are subject to the following objection. It will hardly ever be the case that someone else could run my life better than I. We make the judgement we do in the above case because we are influenced by the judgements we would make in more realistic situations.[25] We think my life does not go better in the first case above, so the objection goes, not because autonomy has intrinsic value, but because you will not run it as well as I. The explanation for the case lies in John Stuart Mill's claim that people are the best judges of their own interests[26] rather than in the value of autonomy. The difference between the two explanations is that the intrinsic value of autonomy would provide greater protection against paternalism than the best-judge principle. According to the former, people's lives go better if they are left alone to lead autonomous lives even if they end up making worse choices for their lives than others would make.

Wall considers this objection with regard to his first case. He does not accept it, but does not say why. He does not, however, seem to realise that the objection also applies to the second case with equal force. It will hardly ever be true that B can let a

---

[24] Wall, *Liberalism, Perfectionism, and Restraint*, p. 147.

[25] Wall, *Liberalism, Perfectionism, and Restraint*, pp. 146-7.

[26] John Stuart Mill, *On Liberty*, Norton Critical Edition (David Spitz, ed.), New York, Norton, 1975, ch. 5, paragraph 18, p. 101: 'there is no one so fit to conduct any business, or to determine how or by whom it should be conducted, as those who are personally interested in it.'

friend decide for her and lead a better life than she would have had she made her decisions herself. This could be why we are led to make the judgment that Wall says we would make, rather than because personal autonomy has intrinsic value. So there is a plausible explanation for the judgements we make in the cases above that does not depend upon an assertion of the intrinsic value of autonomy.

One reason why we might think that the best-judge explanation cannot be the right one is that it seems to imply that when, as in the two cases, someone else could run my life better than I could then they should do so. And we may be reluctant to accept this. But our reluctance is not necessarily explained by the thought that autonomy makes our lives better. A different explanation of the intuition could be given. Wall claims that the intuition is that A's life goes better than B's. Perhaps it is a slightly different one: that A's situation is better than B's. This could be because certain liberties are not being respected in B's case that are in A's. B's friend could be deciding where B will live and work and who she will spend her time with without giving B any freedom to decide these matters. We may think that this means that B's rights are not being respected. This is worse even though not worse for B, that is, it may not reduce his well-being. Even if B has voluntarily consented to the interference, we may think that the rights being violated are inalienable, or we may think B is in danger of being exploited by A, or we may doubt that B's rights have been relinquished with real voluntariness. The judgement we make may not involve the thought that B is worse off, has less well-being, than A. We may think that B's situation is worse than A's, worse, that is, in terms of rights not being respected or the likelihood of exploitation rather than worse in terms of well-being. Thus we have a different explanation for our intuition in this case, one that does not imply that autonomy has any intrinsic value.

Wall has a separate intuitive argument for the value of the aspect of autonomy that stresses the importance of having a wide range of options from which to choose. He

asks us to imagine a case that compares the goodness of a person's life in two situations. In the first situation Sally has two different valuable options, *x* and *y*, and chooses *x*. In the second Sally has eight different valuable options including *x* and *y* and again chooses *x*. Wall claims that we would normally think that Sally is better off in the second situation than in the first, but since she chooses *x* in both situations this cannot be because the second allows her to choose an option that better suits her. Thus there must be something independently good about having a wider range of options.[27] The difference between the explanations is again important with regard to paternalism. If having options is valuable over and above providing something better to choose, there would be a further perfectionist rationale against paternalism.

It is difficult to see what the intuition is that Wall appeals to. The most obvious reason for thinking that eight options are better than two is that the extra options provide greater opportunity for choosing something better. But this reason is unavailable in the above case since Sally chooses *x* when faced with two options or eight. I believe, contrary to Wall, that there is no improvement in the worth of my life in having extra options once this reason is set aside. If I spend a lifetime in a career that is enjoyable, fulfilling, and successful should I think of that life as better if, at the start, I had other options available to me, than if I had not? The life with options seems no better than the one without, once the possibility of having something better to choose it ruled out.

Consider another example. Say that I am shipwrecked on an island in the Pacific. There is not much to do on this island, but luckily for me, it has an abundance of healthy and delicious food, the sun is always shining, the natives are friendly and companionable and I can spend the rest of my life relaxing in this paradise. Assume that I do so and die a happy man. Now vary the example slightly and assume that a ship comes to rescue me and return me to civilisation. Its arrival supplies me with an

[27] Wall, *Liberalism, Perfectionism, and Restraint*, p. 149.

adequate range of options that I was lacking. Say I refuse to leave the island, content to stay where I am. Imagine that the ship comes back every few years, giving me the same offer, thus maintaining my range of options. However, I never take up the offer and remain on the island, again dying a happy man.

Is my life better in the second situation with the occasional offer of more options? Surely the two lives are equally good. They contain equal enjoyment, fulfillment, and success. How can the mere presence of more options, options that in the pacific island example above, are all for something worse than what I now have on the island, make my life better? The relative goodness of my life turns upon what there is in it, compared with what might have been. Remaining on the island is the best my life could have been compared to the other options the ship's arrival offers me. It seems difficult to see why it would be any better merely with the presence of options for worse lives.

In presenting these hypothetical examples, I do not claim to have shown that autonomy has no value. I merely hope to have at least cast some doubt on the intuition that autonomy improves the value of people's lives, over and above providing the opportunity to choose something more worthwhile. My examples, which I believe to be at least plausible, were intended to demonstrate that the appeal to intuitions alone cannot solve the issue one way or another. More substantive argument is required. Whilst Wall does provide several substantive arguments, I will first examine those of Thomas Hurka.

2.4 Rational Deliberation and the Exercise of Agency

2.4.1 Hurka's perfectionist theory begins with an account of the good life, characterising this as realising what is central to human nature.[28] When extended to politics, his theory holds that 'the best political act, institution, or government is that

[28] Hurka, *Perfectionism*, p. 3.

which most promotes the perfection of all humans.'[29] Because perfectionism holds that some ways of life are better than others regardless of whether people want or would choose them, Hurka admits, it may be thought to favour state coercion to force people into excellence.[30] One way he tries to resist this charge is by arguing that autonomy itself is a perfection. Autonomy, according to Hurka, is an intrinsic good because it is a form of human perfection and any restrictions on it are prima facie forbidden.[31]

As Hurka explains it, autonomy fits into his account of perfection and is thereby of value in two ways. One connects autonomy to *rational deliberation*, whilst the other connects it to the *exercise of agency*. Both rational deliberation and the exercise of agency, according to Hurka, are elements of human perfection. Regarding the connection between autonomy and rational deliberation:

> Someone with many life options can reflect upon these options and in so doing exercise his rational powers. He can weigh their respective merits and defects and reach a reasoned judgement about them.[32]

This, Hurka admits, is not enough to establish autonomy since rational deliberation does not require that options actually be open to one. A slave may evaluate possibilities as much as he wants without ever having the opportunity to bring about any of those possibilities. But since people rarely deliberate about options they cannot choose, by giving people alternatives the state 'encourages reasoning that would have no practical point if options were closed.'[33]

The other way in which Hurka defends autonomy is by arguing that it allows people to exercise agency. According to Hurka, human perfection

---

[29] Hurka, *Perfectionism*, p. 147.
[30] Hurka, *Perfectionism*, p. 147.
[31] Hurka, *Perfectionism*, p. 148.
[32] Hurka, *Perfectionism*, p. 150.
[33] Hurka, *Perfectionism*, p. 150.

> involves expressing intentions in the world and determining what it does and does not contain. The autonomous agent, by virtue of her autonomy, more fully realizes this ideal. When she makes choices, she has two effects: realizing some options and blocking others.[34]

These choices have value only when the person makes them knowingly says Hurka. Her rejection of the others must appear in her mind so that she intends their non-realisation. Moreover, agency is realised even more when choices follow deliberation. Reflecting on the desirable and undesirable properties of options creates more possibilities for realising intentions in the world. Choosing *a* and not *b* because of their properties means that not only does the agent bring about *a* and not *b* but also intends to bring about the properties associated with *a* and not the properties associated with *b*.[35]

Hurka admits that this argument for autonomy cannot provide an absolute protection for liberty.[36] Whilst it makes state paternalism that restricts liberty prima facie forbidden, paternalism could be justified if interference would improve overall excellence by enough to outweigh the loss of excellence when autonomy is invaded. This could occur by promoting other excellences such as creativity and knowledge sufficiently to outweigh the loss of autonomy. Many activities involve little rational deliberation and exercise of agency, thus making them susceptible to paternalistic intervention with little loss of valuable autonomy. But Hurka does not find this implication unacceptable. He thinks the argument from autonomy (along with his other arguments for liberty, to be discussed in chapter 5) is enough to rule out all paternalistic policies except those he finds acceptable such as laws against addictive drugs[37], seatbelt legislation, compulsory medical insurance, and perhaps laws

---

[34] Hurka, *Perfectionism*, p. 150.
[35] Hurka, *Perfectionism*, p. 151.
[36] Hurka, *Perfectionism*, pp. 149, 152.
[37] Hurka, *Perfectionism*, p. 156.

discouraging smoking.[38] His theory, however, supports much more paternalism than he seems to think.

2.4.2 To see this consider three types of paternalism, each illustrated by a specific case. The first may be termed *lifestyle paternalism* since it attempts to guide people away from worthless lifestyles into more worthwhile ones. An example of this could be a policy of discouraging couch potato lifestyles. Some people spend the most part of their lives lying on the sofa all day every day drinking beer and watching soap operas and game shows on television or living similarly idle lives. They may feel contented in their lifestyles and there could be no indication that they will ever change. Paternalistic state policies could encourage such couch potatoes to take up more worthwhile activities. These policies could be of a sort that is only slightly or not at all objectionable, such as publicly-funded advertising campaigns that aim to encourage more worthwhile behaviour. Assume however that such policies are not effective and the state takes more direct action. It legislates against idle lifestyles, punishing them with heavy fines. The state names its policy the Campaign for Quality of Life (hereafter CQL).

An initial difficulty with this policy that may be suggested concerns its enforceability. How can the state find out who are the couch potatoes? The state could encourage citizens to inform on couch potatoes, in much the same way that it encourages people to report other crimes committed in the home such as domestic violence and child abuse. I do not wish to trivialise these crimes by putting them on a par with activities like watching television and drinking beer. My point is simply to show that there are ways of making the paternalist policy enforceable.

Once the couch potatoes are discovered, how can state officials know that they are wasting their lives rather than quietly achieving perfection? The fact that they are not

[38] Hurka, *Perfectionism*, p. 158.

actively employed or involved with the community, if it is a fact, does not necessarily mean that they are not achieving Aristotelian perfection, as interpreted by Hurka. They could be writing novels, painting pictures, and mastering musical instruments in private. However, this does not prevent the state from enforcing its CQL since people who are writing novels, etc. could use that as a defence when charged with life-wasting. And they could exhibit the products of their activities in order to prove that defence. The real couch potatoes, on the other hand, could not.

I take it that CQL is a case of objectionable paternalism. Couch potatoes should be free to pursue their chosen lifestyle unhindered by the state. It would be unacceptable for the state to prosecute people for watching television and drinking beer too much. But Hurka's appeal to the value of autonomy cannot explain why. On the contrary, some of his arguments would seem to support the CQL.

Hurka may claim that such a policy would interfere with individual autonomy and is therefore prima facie objectionable. However, just considering the value of autonomy itself, Hurka admits that sometimes restricting a person's autonomy at the present will do more to increase it in the future by giving more options or a greater capacity to choose.[39] This is especially evident with couch potatoes. By living a life of beer-drinking television-watching, their mental and physical abilities would become dulled and they would thus lose many capacities that would enable them to do other things. Forcing them out of their lifestyles would help ensure that they retain their abilities and thus their range of options.

Moreover, as we have seen, Hurka views autonomy as only one good that may be outweighed by others. This is especially true for couch potatoes since their choice of lifestyle arguably has few of the qualities that Hurka thinks makes autonomy valuable, thereby making it easily outweighable. It is conceivable that couch potatoes may

[39] Hurka, *Perfectionism*, p. 148.

rationally deliberate about which television soaps and game shows to watch and which type of beer to drink. Assume however that most do not do this. They watch whichever ones are on without giving any careful thought to the matter and drink whichever beer is available in the shops without ever becoming beer connoisseurs.

Whilst couch potatoes do exercise agency by 'expressing intentions in the world and determining what it does and does not contain'[40], they do so only in a trivial way. If being a couch potato is option *a* and the alternatives are *b, c, d,* etc. then a couch potato makes *a* the case and is also responsible for not-*b*, not-*c*, not-*d*, etc. If the state prohibits couch potatoism then people can instead choose another option, say *b*, and still be responsible for not-*c*, not-*d*, etc. Bringing about *b* would be an equally valuable exercise of agency as bringing about *a*.[41] So the only loss of value is the elimination of bringing about not-*b*. And people who would not have been couch potatoes anyway only have their responsibility for not-*a* eliminated by the paternalistic policy. Thus, in terms of the effects on the exercise of agency, the policy would be trivial. All that is lost is the chance not to bring about one option among many others also not brought about.

Moreover, it may be assumed that couch potatoes don't do what they must, according to Hurka, to make the exercise of agency valuable, that is 'make in the fullest sense a choice among [options]: a choice for one option and against others, so [their] rejection of the others appears in [their] mind[s].'[42] We can assume that they have no more than a vague notion that there are other possibilities that they are rejecting by choosing the lifestyle they do. It can also be assumed that couch potatoes do not deliberate about the qualities possessed by the options they choose, which means that they have no or only vague intentions about which properties they are bringing about by choosing

---

[40] Hurka, *Perfectionism*, p. 150.

[41] This is on the assumption, which I see no reason to doubt, that *b* has no fewer intentions in and around it than *a*. (See Hurka, *Perfectionism*, p. 151.)

[42] Hurka, *Perfectionism*, p. 150.

what they do and which properties they do not bring about by not choosing what they do not.

Hence, couch potatoes' autonomous choices (if they can count as autonomous choices) of lifestyle have little value, on grounds of autonomy, on Hurka's account. Interference to force them into more worthwhile activities is therefore not prohibited by the value of autonomy in Hurka's theory. Hurka's appeal to the value of autonomy does not succeed in ruling out the CQL. On the contrary, autonomy supports it. The CQL would protect and further couch potatoes' autonomy by making them rationally deliberate and exercise fully conscious agency. It will enable them to achieve more perfection by performing activities that affect themselves and others in significant ways and that involve a more complex hierarchy of goals. Similar arguments could be made, I believe, to show that autonomy on Hurka's account does not rule out and may support other instances of lifestyle paternalism.

2.4.3 Another type of paternalism that seems barely threatened by the value of autonomy, which could be called *aesthetic paternalism*, occurs when the paternalist tries to force people into aesthetically more worthwhile activities. Imagine that the state prohibits the broadcast, sale, and distribution of recordings of a genre of music that is of little value. The music has instant superficial appeal but no lasting value. Many people choose it over more worthwhile music due to its instant appeal, even though in the long run they would gain a deeper appreciation of music and more enjoyment overall if they listened to something else. In our hypothetical example, the state suppresses bad music in the hope that people will listen to more worthwhile music instead. It calls this policy Improving Musical Appreciation.[43] The policy would be enforceable. There would of course be black markets in bad music, but these can be dealt with by effective policing.

---

[43] While a more accurate title for the policy would perhaps be something like the 'Campaign Against Bad Music', governments tend to avoid negative terms for their programmes.

Whilst the above example is hypothetical, similar policies have been implemented by several governments. Most governments censor individual songs or records[44] but some target entire genres. Prior to the collapse of the Soviet empire, Western pop music was considered 'degenerate and immoral' by authorities in Eastern Europe. East German leader Walter Ulbricht thought it 'spiritually deadening' and Bulgarian leader Todor Zhivkov called for it to be eradicated 'with a red hot iron.'[45] Russian musicians who played in this style were criticised and restricted, although not arrested or exiled. In the Soviet Union, the government run recording company, which had a monopoly, refused to release recordings of 'underground' bands until the 1980s. In Burma, the authoritarian state calls for the prevention of 'decadent' pop music. The government also proscribes the recording and performance of pop music but not informal concerts. In Indonesia in the 1950s, rock and roll was banned from the radio, with President Sukarno calling it 'decadent.' Local groups playing Western rock were harassed or jailed. In the 1970s, *dangdut*, a genre that is a fusion between Malay ensemble music and Western pop, was condemned by politicians as 'loud, vulgar, and inherently corrupting of public morals.' Its leading performer, Rhoma, was banned from government radio and television for a decade and was not permitted to perform in Jakarta. In Malaysia in the late 1980s, Parti Islam, the major Islamic opposition party, called for the banning of pop music as immoral. In 1995, the Parti Islam-controlled government in the east coast state of Kelantan banned public performances of pop music involving singing and dancing.

---

[44] For the U.K. see Martin Cloonan, *Banned! Censorship of popular music in Britain 1967-92*. Aldershot, Arena, 1996.

[45] All of the quotations and the rest of the information in this paragraph are from Craig A. Lockard, *Dance of Life: Popular Music and Politics in Southeast Asia*. Honolulu, University of Hawai'i Press, 1998, pp. 28-9, 35-6, 77, 96-101, 248. It is not always clear that the grounds for censorship in these cases is paternalistic. Sometimes the thought seems to be that the music will cause its listeners to harm others or society in general. But part of the rationale at least in many of the cases sounds paternalistic. For more information and articles on the censorship of pop music see *Index on Censorship*, issue 6 of 1998.

We are, I believe, rightly suspicious that Improving Musical Appreciation (IMA) would be a legitimate state policy. People should be free to listen to whatever they wish no matter how aesthetically bad the music. Hurka's arguments for the value of autonomy, however, do not prevent the state from adopting IMA.

The argument from autonomy cannot rule out IMA because autonomy is not absolute in Hurka's theory. Whilst free choice among options is an intrinsic good, it may be restricted for sufficient increases in other goods. This is especially true in the case of bad music since the freedom to buy and listen to it has few of the characteristics that Hurka claims make autonomy valuable. With regard to listening to bad music on radio or television, very few if any who do so rationally deliberate upon its merits and defects to reach a reasoned judgement about it. On Hurka's account a person exercises agency by determining what the world does and does not contain, and if this is a fully conscious decision then it has value. But this exercise of agency does not occur when people listen to bad music on radio or television. Whenever a bad song is played, listeners have not determined what the world does and does not contain in any significant sense, and they have not made a fully conscious choice among options. Rather, most simply listen in a passive way to whatever is played (although they may exercise some agency by changing stations if what they hear is beyond the limits of their toleration). Of course, those who make the decisions regarding what will be played on radio or television are exercising agency in the way Hurka describes. They make a fully conscious decision to bring about a situation where bad songs are broadcast. But if bad music were banned they would choose to play something else instead (since there is still a wide range remaining open to them) and exercise as much agency. If the state prevents the broadcast of bad music it does not undermine any valuable exercise of agency.

There is probably some rational deliberation whenever a bad music record is bought. But if they were prohibited there is no reason to think there would be any less rational

deliberation whenever people do whatever they do instead, say, buying another type of music. There is a valuable exercise of agency whenever someone buys a bad music album. The buyer has expressed his intention. The world contains one more owner of a bad record and one less unsold bad album. And the purchaser's decision would usually be fully conscious and may follow deliberation about the properties of a bad album and available alternatives. But IMA would not undermine the exercise of agency in any significant way. The same argument that was made with regard to lifestyle paternalism applies here. Say buying the bad album is option *a*. If the state prohibits the sale of bad albums then the purchaser can instead choose another option *b* and still be responsible for not bringing about the other alternatives. Bringing about *b* would be an equally valuable exercise of agency as bringing about *a* (assuming that *b* has no fewer intentions in and around it than *a*). So the only loss of value is the elimination of bringing about not-*b*. And people who would not have bought bad albums anyway only have their responsibility for not-*a* eliminated by IMA. Thus, in terms of the effects on the exercise of agency, IMA would be trivial. All that is lost is the chance not to bring about one option among many others also not brought about.

2.4.4 Consider a final type of paternalism, which may be termed *relationship paternalism*. Here, the paternalist tries to improve the quality of people's relationships. Imagine that the state aims to promote and support healthy adult relationships. It believes that such relationships are part of the good life and tries to ensure that they are maintained by those in them. To help achieve this end, the state employs relationship advisors whose tasks are to regularly call upon couples, whether married or not, and attempt to identify and offer help with relationship problems. They make personal visits to the homes of couples once per month. The state makes it illegal for them to be turned away, although they have to find a visiting time suitable for the couple. Whilst the state makes it illegal to refuse to meet with and hear advice from the advisors, there is no requirement to follow the advice. What happens within the relationship is entirely voluntary (or as voluntary as these things

normally are) and the state does not force people to remain in a relationship. The advisors' endearing personal manner enables them to overcome the resistance they will naturally and inevitably initially face, thereby making it more likely that their advice will be listened to. Due to their training and experience the relationship advisors help to improve and save many relationships. The state names this policy the Supporting Relationships (SR) programme.

The British government is proposing something similar. In its consultation document *Supporting Families*[46] it expresses the desire to 'help couples make a success of their adult relationships and to encourage people experiencing difficulties to seek help.'[47] Following the birth of a child, the government 'sees advantage in asking midwives and health visitors to identify and offer help with relationship problems during the post-natal period.' It will ask the National Family and Parenting Institute (an 'independent charity' established and funded by the government, receiving £2 million over the next three years) to consider how best to develop new ideas and services that would help adult relationships. It also seeks to use an already existing requirement, that couples attend an individual information meeting at least three months before starting divorce or separation proceedings, to determine whether the marriage may be saved. Other proposals include encouraging marriage registrars to make available a marriage preparation pack, requiring both partners to attend the register office to give notice of marriage, and requiring couples to give a minimum of 15 days' notice of their intention to marry. These proposals are not straightforward cases of paternalism since sometimes the stated rationale is to benefit and protect children rather than to further parents' own good. However, even if protecting children is the ultimate goal, the proposals are paternalistic in that they do seek to further the good of parents (the good

---

[46] Home Office, *Supporting Families; A Consultation Document*, London, HMSO, 1998. For a brief analysis and critique of *Supporting Families*, see Irma Reilly 'Supporting families like a rope supports a hanging man', *LM* (Living Marxism), number 116, December 1998/ January 1999, pp. 8-9.

[47] Home Office, *Supporting Families*, p. 36. All subsequent quotations are from pp. 4, 34-6 of this document.

of those subject to the intervention) even though only as a means to the ultimate goal. Whilst there are similarities between Supporting Families and my hypothetical SR, the latter is more clearly objectionable than the former, due to the greater intrusion involved.

Hurka's perfectionism holds that certain relationships are valuable. They create common goals for a partnership and require coordinated acts to achieve them. In a good relationship, 'a pair's aims stretch across persons, uniting them in a single intended and accomplished state.'[48] Since relationships usually contain long term goals that involve both people at all times in the future, and since those goals fit into a single hierarchy and are intended both intrinsically and instrumentally, then relationships have perfectionist value.[49] This means that the state has a motive for SR if it adopts Hurka's theory. However, I take it that such a policy would be unjustified. State officials should not try to forcibly tell people how to live their lives in this manner. Hurka's arguments, however, fail to provide any reason for why the state should not implement that policy.

A monthly visit from a relationship advisor would do little to invade autonomy since it closes off only one option (that of not being visited monthly) whilst leaving many others open, and as has already been noted, any loss may be outweighed by some other perfectionist gain. This would especially be true given the criteria that according to Hurka makes autonomy valuable. The possibility of rational deliberation would be enhanced rather than diminished by sessions with the relationship advisor. She could help make clear the possible options and their merits, and encourage rational deliberation on them. Moreover, the valuable exercise of agency could also be enhanced. Since the advice given is not compulsory, no options are blocked. Hence people will still determine what the world does and does not contain with

---

[48] Hurka, *Perfectionism*, p. 134.
[49] Hurka, *Perfectionism*, p. 135.

respect to their relationships, as before. To be valuable, these choices must be fully conscious ones. The relationship advisor can help ensure this is the case. She can encourage the making of fully conscious choices intended to bring about certain properties. The value of autonomy supports rather than limits paternalism in this case.

Thus the value of autonomy, on Hurka's account, fails to rule out objectionable instances of these three types of paternalism, many instances of which seem highly objectionable. Hurka considers three possible moves which would give greater protection to liberty.[50] One possibility is the idea that autonomy could be lexically prior to other perfections so that any loss of autonomy would outweigh any gain in other goods. Another is to hold that some minimum of autonomy is lexically prior to other goods, so that losses below this minimum but not losses above it outweigh gains in other goods. Finally there is the idea that autonomy, rather than being one good among others, is a condition of goods, so that activities only have value if they are autonomously chosen. Whilst these possibilities would help his view, Hurka finds them implausible. Instead, he relies upon non-autonomy based arguments to resist the illiberal features of his theory, arguments that will be examined in chapter 5.

Perhaps, however, Hurka is too quick to dismiss these strategies for giving autonomy greater protection. He asks us to imagine that Mozart was forced into music as a child, so that his life lacked autonomy, whether as a good that can be weighed against others or as a condition of goods. According to the three views of autonomy above, 'Mozart's life, despite its great musical achievements, contained less perfection than if he had been given freedom in his youth and had autonomously chosen a life of suntanning.'[51] Hurka believes this claim is implausible so that we have to reject the three possibilities and accept that autonomy is a good that may be weighed against others.

---

[50] Hurka, *Perfectionism*, p. 149.
[51] Hurka, *Perfectionism*, p. 149.

In 2.6, I examine an argument for autonomy being a condition of goods which seems to avoid the implications Hurka thinks it has. With regard to lexical priority, it does seem implausible to give autonomy lexical priority over other values. But it may be possible to argue for some position that lies between the view Hurka accepts (the one that allows trade-offs) and the lexical priority view. Such a position could give the value of autonomy great weight so that it would avoid most of the paternalistic trade-offs that Hurka's theory seems to allow, but at the same time not give the lexical priority to autonomy which entails that a life containing an infinite amount of other values is less good than one containing only a drop of autonomy. An argument for this position would have to show that autonomy, whilst being one value amongst others, is of a special nature such that most, if not all, trade-offs are ruled out. The next section examines some such arguments.

2.5 Giving Meaning to Life, Having Character, and Responsibility

The arguments on this section are all put forward by Steven Wall. Wall does not think of them as establishing any lexical, or almost-lexical priority. He believes that their success simply shows that autonomy is valuable, without showing that the autonomous life is more valuable than other ways of life.[52] For that, he relies upon an argument to be considered in the next section. But he believes that the arguments I examine in this section add to the one in the next to build a composite case for autonomy. The arguments here are also important for another reason. Their nature suggests that they could establish a high priority for the value of autonomy, one which may rule out the possibility of trade-offs between autonomy and other ideals.

2.5.1 Wall has separate arguments for the independence and adequate-option-set aspects of autonomy. His argument for independence is that it is good for people to lead their own lives because by taking charge of their affairs and charting their own

[52] Wall, *Liberalism, Perfectionism, and Restraint*, p. 162.

course through live, people give meaning to their lives. They impart significance to it. Others' lives may have meaning but it is not something they give to it themselves. It is valuable for people to be independent because it is valuable for them to give meaning to their own lives.[53] Wall also has an explanation for why it is valuable to give meaning to one's life. He claims that giving meaning to one's life is essential to having a character. Since a slave has no part in fashioning who he is, Wall argues, he has no character that is his own. A slave's character is determined by what happens to him rather than by what he does. In contrast, by engaging in projects and taking on commitments people can make themselves one type of person rather than another. Having a character in this active sense, says Wall, is intrinsically good. This is why it is valuable for people to give meaning to their lives, and hence why independence is intrinsically valuable.[54]

It is not obvious that this argument need have the structure that Wall gives it. The second step of the argument may be superfluous since giving meaning to one's life seems a compelling value that having character in the active sense may add nothing to. Or having character in the active sense may alone be sufficient to establish the value of independence without the need for an intermediary step referring to giving meaning to one's life. Alternatively, the two steps of the argument could also be reversed. The value of independence perhaps rests on it enabling people to have character in the active sense and the value of having such character may be explained by it helping people to give meaning to their lives.

Whatever the structure of the two ideas in the argument, the argument is not sound anyway since acceptance of either of the two ideas fails to establish the value of independence. This can be demonstrated by recalling the examples of paternalism discussed in the last section. Lifestyle paternalism, illustrated by laws against couch

---

[53] Wall, *Liberalism, Perfectionism, and Restraint*, p. 147.
[54] Wall, *Liberalism, Perfectionism, and Restraint*, pp. 147-8.

potato lifestyles, involves forcing people out of worthless ways of life. Such policies clearly invade independence, but do not seem to undermine the possibility of people giving meaning to their lives or having character in the active sense. Laws against the couch potato lifestyle leave ample room for people to 'take charge of their affairs and [chart] their own course (sic) through life, [thus giving] meaning to their lives.'[55] Moreover, couch potatoes would have little meaning in their lives, let alone have given them meaning, and their lives are more likely to have meaning and perhaps be given meaning by them if they are forced out of their lifestyle. So the rationale would seem to support lifestyle paternalism even though it invades independence. Similarly, lifestyle paternalism also leaves room for a person to have character in the active sense, to '[fashion] who he is', to 'make [himself] one type of person rather than another.'[56] Couch potatoes, in fact, have no character in the active sense, so laws discouraging them would help further the achievement of it, by blocking inactive lifestyles and allowing people to choose amongst active ones.

Similar arguments can be made about aesthetic paternalism. Censoring a genre of worthless music would not prevent people from giving meaning to their lives, or from shaping their own character. Relationship paternalism is more problematic. There is the danger that the relationship advisors imagined might run the relationship rather than the parties to it, which would mean that although the relationship might have meaning in the lives of the parties to it, it is not meaning that the parties have themselves given. And by not fashioning the relationship completely alone, they may have less character in the active sense. But these consequences are not unavoidable. So long as the advisors conduct themselves appropriately, the parties will still be able to give meaning to their relationship and determine their own character, even though they are legally required to listen to relationship advice. The two values need not be undermined by a tactful relationship advisor.

---

[55] Wall, *Liberalism, Perfectionism, and Restraint*, p. 147.
[56] Wall, *Liberalism, Perfectionism, and Restraint*, p. 148.

Wall does not set out how and why the two values may be undermined. Some circumstances that do undermine them include having others run your affairs (either as a slave or with your consent) or simply drifting through life without making decisions. The explanation for why the two values are not instantiated in these circumstances is that the people in those cases *themselves* play no part in fashioning their character. *They* do not chart their own course through life or impart significance to it. But this explanation has an important consequence. It makes the source of the restriction that undermines the two values irrelevant. Whether the source of the restriction is the intended interventions of others or circumstances that are unintended by anyone, the result in both cases could be to leave people less able to give meaning to their lives and have character in the active sense.

This means that blocking the options of engaging in a particularly worthless lifestyle or devoting much time to listening to a worthless genre of music for paternalistic reasons does not undermine the two values any more than if the options became unavailable, not due to intentional intervention, but say because the lifestyle is economically impossible for people to sustain and the genre of music is unable to sustain itself in the market. If the paternalistic policies undermine the two values, does that mean that the state would have to maintain the availability of the options if they are economically unsustainable in order to ensure people are able to give meaning to their lives and have a character in the active sense? Even if the state should protect these options, would it be for *that* reason? It seems implausible that bad music should be supported by the state in order to help people give meaning to their lives and have character.

Similarly, if the two values rule out relationship paternalism, then they also seem to require that the state ensure that the possibility of having bad relationships continues to exist, so that people may avoid them and thereby give meaning to their lives and

have character in the active sense. The continued existence of bad relationships might be threatened by the availability of advice books on relationships. Such books are similar to relationship paternalism in that both offer information that may be helpful in running the relationship while still allowing the parties enough control over it to ensure that whatever meaning it has in their life is given to it by themselves. Since it seems so implausible that the two values are threatened by a reduction in the possibility of having bad relationships caused by widespread availability of advice books, we should accept that the two values are similarly unthreatened by relationship paternalism.

It might be objected here that there is a difference between acts and omissions. Whilst the state should not act in a way that undermines the two values, it need not act in order to protect them when they are threatened by unintended circumstances. But this distinction ought to be of little importance to a perfectionist. If the two values are undermined, people's lives will be worse. Not acting to prevent this would be a betrayal of perfectionist ideals.

Wall suggests a reply to the objection that people's independence may be invaded without sacrificing either of the two values. He states that autonomy, and the lack of it, come in degrees. This, he says, means that to the extent that people's lives are run by others, then to that extent 'they do not have character in the active sense.'[57] If so, then even if the paternalistic policies described above leave some room for the two values in Wall's argument to operate, the policies still lessen them to some degree. Each person affected will have a character in the active sense to a lesser degree. (Wall is unclear on whether the person will also have given meaning to her life to a lesser degree, but I believe that is what the argument entails.)

---

[57] Wall, *Liberalism, Perfectionism, and Restraint*, p. 148.

One problem with this suggestion is that it seems to re-introduce the possibility of trade-offs between values. Losses in the two values could be outweighed by other perfectionist gains, thus giving more scope for paternalism. Moreover, even if giving meaning to one's life and having a character in the active sense can occur in varying degrees, I doubt that the paternalistic policies described above lessen these values *in any way*, so long as there are plenty more and more worthwhile options available to choose from. Even if they do, Wall's suggestion faces the following difficulty. For the argument to be of any use, the undermining of the two values has to be significant enough to provide a reason against the objectionable paternalistic policies. Perhaps this undermining has to be large to do so or perhaps any undermining no matter how trivial provides such a reason. Whatever the size of the reduction required, however, if their reduction provides a reason against the examples of lifestyle, aesthetic, and relationship paternalism I described, then the same reductions also provide reasons for ensuring that the options continue to remain available when they are threatened not by intentional intervention but because of unintended circumstances such as the economic unsustainability of a lifestyle, or the unpopularity of bad music making it fade away, or increased understanding and perception making the possibility of bad relationships remote. If the paternalistic policies that invade independence reduce the two values, then the values are also reduced, that is people are less able to give meaning to their lives and have a character in the active sense, when the options become unintentionally unavailable, even if there are plenty more, and more worthwhile, options still available to them. I assume, however, that the reduction in the two values, if there is any, when the options disappear through unintended circumstances does not provide a reason to ensure those options remain available. (If there are reasons to preserve options, they are not those.) Hence, similar reductions do not provide reasons against the paternalistic interventions.

2.5.2 Wall's argument for the value of the independence aspect of autonomy is therefore unable to do much work in showing how perfectionism can resist

objectionable forms of paternalism. But he has a separate argument for the value of the wide-range-of-options aspect. The argument attempts to show that having a wide range of options has value over and above the fact that it allows a person to choose something that better suits her. If the argument succeeds, it would provide more protection against paternalism. It would provide a case against paternalism which restricts options even when the options restricted are less good for the paternalised than those left open. Wall claims that Sally is worse off in the situation where she faces only two options than in the situation where she faces eight, even though she chooses the same option, *x*, in both. He states that in the second situation Sally is *more responsible* for the state of affairs that obtains than she is in the first situation. This is because in the second situation she is responsible for the non-realisation of seven alternatives whilst in the first situation she is responsible for the non-realisation of only one alternative. This greater responsibility reflects a greater exercise of self-determination, where the latter is determined by both what we choose to do and choose not to do. Having a wide range of options, says Wall, 'makes us more responsible for our choices in the sense that it makes our choices more our own.'[58] Hence autonomy (or at least this aspect of it) has intrinsic value due to its connection with responsibility.

This argument, however, confuses two senses in which a person may be said to be more responsible. First is the sense that in making a decision a person is less subject to the influence of other people and/or of circumstances in general. Other things equal, she would be more liable for blame or praise due to her greater responsibility. In this sense, for example, someone who chooses an occupation free from financial constraints is more responsible for that choice than someone who chooses an occupation only because economic circumstances forced him to do so. Coercive threats also undermine this sense of responsibility. The second sense in which someone may be said to be more responsible is when a person affects the state of

[58] Wall, *Liberalism, Perfectionism, and Restraint*, p. 149.

affairs in a more significant way. For example, the person who gives a large sum of money to charity may be more responsible for relieving poverty in society than someone who gives a small sum. A person could be more responsible in the first sense but less in the second; he could make a decision significantly free from influence but not bring about much. And vice versa; her decision could have important and far-reaching effects but have been highly influenced or even coerced by others.

It is clear that if Sally has greater responsibility in the second situation it is in the second sense of responsibility described above. She is more responsible for the state of affairs involving the non-realisation of seven alternatives than she is for the state involving the non-realisation of only one alternative because her decision has more consequences. However, Wall's argument switches to the first sense of responsibility and assumes that she is more responsible in that sense, when he claims that Sally's greater responsibility in the second situation reflects a greater exercise of self-determination. It is only in this sense of responsibility that our choices are made more our own; our choices are more our own when we are not coerced or manipulated by others or forced by circumstances. It is implausible to think that upon the other, second, sense of greater responsibility, which turns upon the significance of the effects one's decisions have, one's decisions are made more one's own. Sally's decision in the first situation to choose x over y is no less her own choice than her decision in the second situation to choose x over the seven alternatives.[59] By confusing these two senses of responsibility, Wall's argument fails to establish the intrinsic value of the wide-range-of-options aspect of autonomy.

---

[59] It is assumed by hypothesis that she is equally responsible (in the second sense of responsibility) in the two situations, that is, that she is no more subject to the influence of others or of circumstances in either situation. This is assumed since in the example all other things must be assumed equal in order to focus only on Wall's concern in the present argument: the presence or absence of a wide range of options.

Perhaps, however, something can be rescued from the argument. Wall's argument involves two steps. The first claims that having a wide range of options makes people more responsible. The second step tries to show that this greater responsibility reflects greater self-determination since when it occurs, our decisions are made more our own. I argued above that each of these two steps involve different notions of responsibility and that Wall's argument collapses the two. But the first step alone may be enough. Having a wide range of options makes people responsible for more consequences, namely the non-realisation of more alternatives. This may be valuable even if it does not make people's decisions more their own. Exercising agency may be valuable in itself. This, as we have seen, is Hurka's view.[60] He would think that it would be better for Sally to face the eight options in the second situation than the two in the first situation even though she chooses *x* in both situations, because she would exercise more agency through greater non-realisation of options. Thus, on this view, a wider range of options is valuable independently of what is chosen. But as we have seen, even if the exercise of agency is always valuable, this view faces the difficulty of allowing trade-offs against other values, especially with regard to activities which involve little exercise of agency. The value of the exercise of agency fails to rule out many objectionable paternalistic measures.

Another possible way of rescuing something from Wall's argument is to focus on its second step. Wall claims that the greater responsibility of Sally in the second situation reflects a greater exercise of self-determination in that her choice is more her own. I argued that the sense of responsibility involved in the notion of making our choices more our own is different to the one involved in Sally's case which means that the former sense of responsibility cannot be used to support the value of a wide range of options. But perhaps it can be used to support the independence aspect of autonomy. A person is independent when substantially free from the coercion and manipulation of others. The more independent a person is, the more she may be said

[60] See section 2.4 above and Hurka, *Perfectionism*, p. 150.

to be responsible for her choices in the first sense of greater responsibility described above. Her choices may be said to be more her own. If this sense of responsibility is valuable (as Wall assumes it is in his argument) then we have an argument for the value of the independence aspect of autonomy rather than the wide-range-of-options aspect.

This argument, however, faces a similar objection that Wall's other argument for independence faced: that if invasions of independence undermine the value in question, then other circumstances would seem to undermine it just as much. For example, the fact that a lifestyle is economically impossible to engage in means that a choice not to engage in it is less my own choice than it would be if it were economically possible, in a way similar to if it were blocked from me by the state. Whether the source of the restriction is the intentional actions of others or economic circumstances seems irrelevant. In both cases, with respect to responsibility, the effect is the same; the choice is less my own. Hence, upon this argument, if the value of responsibility provides a reason for the state not to invade independence paternalistically, then it also provides a reason for the state to protect and ensure the availability of worthless lifestyles from disappearance caused by economic unsustainability. By doing the latter, the state would ensure that people's decisions not to engage in such lifestyles are more fully their own. Similar reasoning would apply to the censoring of a worthless genre of music and not protecting it from disappearance due to unpopularity. The argument would also, I think, give the state a reason to ensure that the possibility of having bad relationships does not become remote since without such possibilities, the goodness of people's relationships would be less their own in that part of the responsibility for the good relationships would instead lie in the fact that the possibility of having bad ones is remote. Imagine that the remoteness is due to widespread use of advice books on relationships. There is a difference between being forced to listen to relationship advisors and deciding to read a self-help book. The latter involves greater responsibility since the person chooses to

gain more knowledge on relationships rather than being forced to receive it. But there is still nevertheless a reduction in responsibility when people follow the advice they receive from a book rather than work things out for themselves. Since it would be implausible to think that the value of responsibility means that the state should discourage the use of advice books, the value of responsibility cannot explain why the state should not act paternalistically by blocking worthless options and forcing people to listen to relationship advisors.

The value of responsibility does rule out some paternalism. Paternalism of a global nature, interfering with each and every major decision in a person's life would leave no space for the person to exercise responsibility.[61] But more local paternalism, whilst remaining objectionable, does little to undermine the value of responsibility in such a way as to make paternalism unacceptable.

2.5.3 Wall also argues that autonomy is valuable because it facilitates the achievement of the ideal of self-development. Self-development is achieved when people fully realise their talents and potentialities. They choose and successfully pursue projects that fit their nature, agree with their innate proclivities and bring to fruition their skills and capacities.[62] Wall does not claim that autonomy is an indispensable condition for self-development. The latter may be achieved without autonomy. Moreover, autonomy does not always lead to self-development. Someone who is autonomous may or may not achieve it. But Wall claims that '[o]ther things equal, autonomous people are more likely to achieve self-development than their non-autonomous counterparts.'[63] Wall has two arguments for this claim. I will try to point out some problems with each argument before raising a deeper problem with Wall's claim that autonomy is conducive for self-development.

---

[61] The term global paternalism is from David Archard, 'For Our Own Good', *Australasian Journal of Philosophy*, vol. 72, no. 3, 1994, p. 287.
[62] Wall, *Liberalism, Perfectionism, and Restraint*, p. 130.
[63] Wall, *Liberalism, Perfectionism, and Restraint*, p. 130.

The first argument attempts to support the independence aspect of autonomy. According to Wall, people generally are in a better position than others to know what projects would best develop their talents. Whilst people do not have perfect self-knowledge, most know more about themselves than others do. Other things being equal, says Wall, people who are independent are better able to develop their talents than those who are told what to do or who follow others or drift through life.[64]

There are at least two types of exceptions to this argument, cases that is, where autonomy would not facilitate self-development. Wall recognises the first type of exception. Even if people are the best judges of what is conducive to their self-development, they may not be motivated to choose it. Coercion or other restrictions of independence could make them choose it.[65] Wall thinks that though there may be something to this objection, it is not conclusive because as well as motivating people, coercion will close off routes to self-development. Hence coercion may make self-development less rather than more likely.[66] But this seems true only for clumsy coercion. Coercion that is careful would avoid this. For example, fines for couch potatoes would motivate them to do something that would develop their talents and abilities and would only close off the option of being a couch potato.[67] No routes to self-development would be closed off assuming, plausibly, that being a couch potato is not a route to self-development. The example may be generalised. The state could coerce people away from choices that do not develop their talents and abilities. In many cases this could be done without interfering with options that lead to self-development.

---

[64] Wall, *Liberalism, Perfectionism, and Restraint*, p. 152.

[65] Wall, *Liberalism, Perfectionism, and Restraint*, p. 154. Wall attributes this objection originally to James Fitzjames Stephen (*Liberty, Equality, Fraternity*, Liberty Fund Inc, Indianapolis, 1993, p. 31) in the latter's criticism of Mill's *On Liberty*.

[66] Wall, *Liberalism, Perfectionism, and Restraint*, p. 154.

[67] To close off only the option of being a couch potato, such a policy would probably have to not be punish too severely. If it were it may close off other options, such as being a television critic, by creating the fear of being mistaken for a couch potato.

The second type of case that is an exception to Wall's argument are those where it is not true that a person is generally the best judge of what projects would best develop his or her talents and abilities. Even if it is true for most people there are some who will not be the best judges of what options are most likely to lead to self-development. They mistakenly choose a lesser option in the belief that it is the one that best develops their talents and abilities. Hence, for them independence is not conducive to self-development and the latter may be better promoted by limiting their independence. Moreover, even if people are generally in a better position than others to know what options would lead to self-development they are nevertheless not perfect judges in this respect. Hence even if they are the best judges, they sometimes get it wrong. If so, then someone else, so long as they are reasonably competent at knowing which options would not lead to self-development, could coerce or manipulate someone away from activities that will not develop her talents or abilities and by doing so make self-development more likely. Here again, independence does not facilitate self-development.

It may be thought that these exceptions need not worry Wall since he does not claim that self-development necessarily requires autonomy and admits that there are cases where the former is achieved without the latter. However, the exceptions are important since the more of them there are, the weaker the argument for autonomy. The exceptions are also important in another way. If self-development provides a justification for independence in cases other than these exceptions, then it provides a justification for not respecting independence in the exception cases. That is, if the state should respect and ensure people's independence for reasons that derive from the value of self-development, then for those same reasons it should invade or violate their independence in the exception cases. If we find it implausible that such intervention would be justified in those cases, then we should be suspicious of the claim that self-development justifies independence in the other cases. The

explanation for the wrongness of intervention would lie elsewhere. Of course there could be other reasons why the state should not interfere in the exception cases, reasons which override those supporting intervention. But if those reasons are sufficient to justify not interfering with independence, overriding the reasons for interference that derive from the value of self-development, then those reasons would also be sufficient to justify independence in the non-exception cases. It then seems difficult to see what work if any the argument from self-development is doing.

Wall's other argument for the claim that autonomy is instrumental for self-development focuses on the aspect of autonomy that makes it (autonomy) depend on having a wide range of options. Having access to a wide range of options facilitates self-development, according to Wall, because it is likely that one of the additional options will better suit and develop one's talents and abilities. If a person is motivated to develop his talents and abilities and if he is the best judge of what activities would best develop them, then giving him access to a wide range of options would facilitate his self-development.[68]

This argument is unconvincing, even given the assumptions it asks us to make. Even assuming the person is motivated to achieve self-development and even if he is the best judge of what is most conducive to this, it is not necessarily the case that widening the range of options he faces will make self-development more likely. This will be the case only if the extra options are conducive to the development of the person's talents and abilities. But self-development will be less likely if the additional options are ones that involve less development of his talents and abilities since there may be a greater chance that he would choose one of these options. This would be true so long as he is not a perfect judge of what is best for his self-development, for if he were then adding options that involve less self-development would simply result in those options being ignored. If he is less than perfect then he sometimes gets it wrong

[68] Wall, *Liberalism, Perfectionism, and Restraint*, pp. 155-6.

and may mistakenly choose one of the additional options. If so then the state, so long as it is reasonably competent, could eliminate options that would not develop his talents and abilities and thereby make self-development more likely. Extra options may be just as likely to hinder as to further self-development.

There may be a way around this objection. Perhaps the value of self-development, even though it does not support a wide range of options as such, supports a wide range of options all of which are conducive to self-development. This may be enough to rule out many objectionable instances of paternalism where the state eliminates options. As Wall says (following Mill), people differ in talents, dispositions and preferences to such a degree that making the same small set of options open to them would not be conducive to everyone's self-development. Individual diversity requires that there be a wide range of options involving different ways of achieving self-development, open to all.[69] But this reply is problematic. Even if individual diversity means that the same options are not equally conducive to the self-development of different people, it does not follow that *all* the different options should be open to *everyone*. Careful state paternalism could block options *a, b*, and *c* (the options less conducive to their self-development) for some people, while leaving those options open to others (where those options are conducive to their self-development) but block the set of options *d, e*, and *f* from those others (because those options are not conducive to their self-development) while leaving *d, e*, and *f* open to those who have *a, b*, and *c* blocked. In this way, the state could block different options for different people, only allowing options to each person that will further their self-development. These remaining options need not be a wide range. On the contrary, they could be intentionally narrow in order to ensure self-development. Thus the fact of diversity combined with the value of self-development does not support the wide-range-of-options aspect of autonomy.

---

[69] Wall, *Liberalism, Perfectionism, and Restraint*, p. 156.

Moreover, like the previous argument, the assumptions that people are motivated to choose options that lead to self-development and that people are generally the best judge of what is most likely to result in self-development may not always hold. Giving those people for whom the assumptions do not hold access to a wider range of options would not facilitate self-development and may make it less likely. Someone who wants to do activities that do not develop her talents and abilities will find that difficult if her range of activities is narrowed so that it is dominated by those that encourage self-development. And someone who is a poor judge of what is best for self-development is more likely to achieve it when his range of options is similarly narrowed.

The arguments so far cast doubt on the claim that autonomy facilitates self-development. They show that self-development may be furthered by restricting autonomy both in its independence and wide-range-of-options aspects. This may not worry Wall. He admits that self-development may be achieved without autonomy.[70] However there is a deeper problem with Wall's argument which stems from an ambiguity with the concept of self-development. There are two different ways of interpreting this concept. On the one hand self-development may refer to the full development of talents and capacities that one has. On the other, it may be stressed that the notion involved is that of *self*-development, so that the development of talents and capacities must be brought about by oneself and not others or circumstances in general. Hence on this second interpretation, given two options x and y where x will lead to a greater development of talents and capacities than y, a person may nevertheless achieve more self-development if he chooses y than if he is forced to choose x. There is less development but more *self*-development.[71] (Of course even

---

[70] See for example Wall, *Liberalism, Perfectionism, and Restraint*, pp. 130, 150, 152-3.

[71] This aspect of the concept is stressed by Jon Elster in his discussion of the closely related notion of self-realisation. See his 'Self-Realization in Work and Politics: The Marxist Conception of the Good Life', *Social Philosophy and Policy*, vol. 3, no. 2, 1996, p. 101. Interestingly, Elster argues (p. 107) that self-realisation is valuable because of the value of autonomy, rather than the other way around.

more self-development would be achieved if he freely chooses *x*, but we can assume that this is not going to happen.)

The first interpretation allows more exceptions to the general claim that autonomy facilitates self-development. As discussed above, some people are not motivated to achieve full development, and some are not the best judges of what is conducive to it, either generally or on specific occasions. Hence there would be many cases where the state could intervene to restrict autonomy to promote self-development. Thus on the first interpretation, self-development provides only weak support for the value of autonomy and weak protection against paternalism.

The second interpretation provides more. In those cases where the state can further the development of a person's talents and capacities by forcing him into some activity, doing so would not further his *self*-development. For that to happen the person must do choose the activity himself. So the state must respect his autonomy. Thus this interpretation of self-development does a better job in supporting autonomy. But this interpretation also gives the argument a question-begging nature. If we ask the further question of what is important about self-development in that sense, we want to know not what is important about having talents and abilities that are developed, for that would only support self-development according to the first interpretation, which as we have seen does little to justify autonomy. We want to know, rather, what is important about developing one's talents and abilities *oneself* and not have it done for you by others or by circumstances. This question is so similar (although perhaps not identical) to the question of what is valuable about autonomy that answering it (the former question) will do most of the work in answering the latter. Saying that autonomy facilitates self-development (in this sense) amounts to saying that choosing for oneself facilitates choosing for oneself that which develops one's abilities. Maintaining that the former is valuable because the latter is does not explain much.

So the dilemma faced by Wall is that he can interpret self-development according to the first interpretation in which case the support for autonomy is weak or he can take the second interpretation which provides more support for autonomy but at the cost of seeming to simply assert rather than argue for the value of autonomy.

One way to resolve the dilemma is to adhere to the second interpretation of self-development and then provide a sound argument for thinking that ideal valuable. But the reasons that Wall gives for thinking self-development valuable support it in its first interpretation rather than in its second. He provides three such reasons. The first is that by achieving self-development a person can lead a life that contains pursuits that give it depth. He claims that by developing one's talents a person can accomplish valuable things.[72] An example of this, as I understand the argument, could be that by developing one's ability for playing the piano, one could play Chopin concertos which is a valuable accomplishment. However this is true if one's talents are developed not by myself but by others or circumstances. It makes no difference how the developed talents come about. Whether developed by me or by others perhaps forcibly, they allow me to accomplish valuable things. Accomplishing these valuable things may be more likely when one is forced to develop one's abilities than when one is free to decide whether to do so. Hence the argument supports the first interpretation of self-development rather than the second.

The second reason for why self-development is valuable, according to Wall, is that people are constituted so as to derive enjoyment and satisfaction from the development of their talents.[73] The same reply applies here. The claim is that people derive enjoyment from having their talents developed not that they enjoy developing them themselves. Thus the argument supports the first interpretation of self-

---

[72] Wall, *Liberalism, Perfectionism, and Restraint*, p. 157.

[73] Wall, *Liberalism, Perfectionism, and Restraint*, p. 157. Wall says that this is the 'Aristotelian Principle' discussed by Rawls in *A Theory of Justice*, p. 426.

development rather than the second. If it were amended to say that people enjoy developing their talents themselves then this would support the second interpretation. But this empirical claim seems less true generally than the first claim. Some people do enjoy developing their own talents themselves, but some find it a frustrating experience and may prefer to be made to develop their talents by others whilst other people would prefer to leave them undeveloped.

Wall's third reason is that everyone benefits when people develop their talents and capacities. As he says,

> When an artist creates a beautiful painting, a scientist discovers a new truth about the physical world, an entrepreneur invents a new product or an athlete develops his or her prowess, all can share in their accomplishments.[74]

Even if this is true, however, it again also supports only the first interpretation of self-development. For these external benefits to occur, all that is needed is that people create, invent, discover, etc. with their talents. Whether they develop their talents themselves or are forced to develop them by others or by circumstances makes no difference. More of these public goods will occur when everyone has their talents and abilities developed to the highest degree. This may be more likely to occur when people are forced into such development when they are not inclined to choose it themselves when free to do so.

Thus none of the three reasons that Wall gives for the value of self-development support it in its second interpretation rather than its first. Hence we should conclude that the value of self-development is unable to support the value of autonomy in such a way as to show how it can serve as the basis for ruling out the paternalistic implications of perfectionism.

2.6 The Social Forms Argument

[74] Wall, *Liberalism, Perfectionism, and Restraint*, p. 158.

All of the arguments considered so far have taken the value of autonomy to be one good amongst others, in one way or another. The arguments have also all failed. Perhaps a successful argument, therefore, may be one that conceives of the value of autonomy slightly differently. According to an argument put forward by Raz and developed by Wall, the value of autonomy lies in its being a condition for participation in projects and activities that constitute well-being. Without autonomy, one's access to the things that determine well-being is blocked. The implication is that autonomy is an 'essential ingredient of the good life so that anyone's well-being suffers if his autonomy is incomplete'.[75]

The argument contains several steps. According to Raz, well-being consists in the successful pursuit of personal goals (broadly understood to include projects, plans, relationships, ambitions, commitments, and so on[76]), in particular, comprehensive goals, those that have ramifications which pervade important dimensions of one's life.[77] Someone may have amongst his goals success in a particular career, a successful marriage, and achieving social status. His well-being, according to Raz, turns upon his success in achieving those aims. Raz also claims that comprehensive goals are based on existing social forms, that is forms of behaviour which are widely practised in society.[78] Marriage, for example, is possible only because society contains the institution of marriage that makes particular instances of it possible. One cannot pursue a career in law or medicine except in societies that contain the practices of law and medicine.[79] The nature and range of goals available to people are determined to a large extent by the social forms in society. Social forms bring options into existence and give them significance in the lives of those who take them up.[80] Putting these ideas together, that well-being depends upon goals and that the

---

[75] Raz, *The Morality of Freedom*, p. 390.
[76] Raz, *The Morality of Freedom*, p. 291.
[77] Raz, *The Morality of Freedom*, p. 308.
[78] Raz, *The Morality of Freedom*, p. 308.
[79] Raz, *The Morality of Freedom*, p. 310.
[80] Wall, *Liberalism, Perfectionism, and Restraint*, p. 165.

availability of goals depends upon social forms, we get the result that well-being is to a large extent dependent on the nature of the social forms existing in society.[81]

The next step of the argument focuses on the nature of the social forms prevalent in most modern western societies. Such societies are dominated by social forms that require individual choice. The prevailing attitude to marriage, for example, emphasises that the partners should choose each other, whereas in the past there was more emphasis on pre-arranged marriages. Because our institution of marriage now rests on individual choice, those who do not exercise that choice will find it difficult to take part in that institution. Similarly, attitudes to work now emphasise individual choice of occupation instead of following in one's parents' footsteps.[82] Since most projects and relationships in an autonomy-supporting environment will be outside the reach of a non-autonomous person, that person will fail to achieve any adequate level of well-being. Such considerations lead to Raz's conclusion:

> The value of personal autonomy is a fact of life. Since we live in a society whose social forms are to a considerable extent based on individual choice, and since our options are limited by what is available in our society, we can prosper in it only if we can be successfully autonomous.[83]

Wall argues along similar lines. Modern western societies are marked by geographic, technological and social mobilities. These mean that people are forced to make choices about where to live, what job to take, and who to live with. They must be capable of revising their goals and plans when circumstances require and be able to assess what options best suit their nature and talents.[84] In short, they need to be able to lead their own lives. Wall concludes that in order to flourish in modern western societies, 'people need to realize the ideal of personal autonomy at least to some substantial degree.'[85]

---

[81] Raz, *The Morality of Freedom*, p. 309; Wall, *Liberalism, Perfectionism, and Restraint*, p. 165-6.
[82] Raz, *The Morality of Freedom*, pp. 392-3
[83] Raz, *The Morality of Freedom*, p. 394.
[84] Wall, *Liberalism, Perfectionism, and Restraint*, p. 168.
[85] Wall, *Liberalism, Perfectionism, and Restraint*, p. 168, italics omitted.

One concern with this argument is that it only succeeds in establishing the value of autonomy in certain societies - those that are autonomy-supporting. It provides no reason for non-autonomy supporting societies to place any value on autonomy. When Raz claims that autonomy is 'an essential ingredient'[86] and 'constituent element'[87] of the good life, those claims are only with respect to people living in an autonomy-supporting society. He admits this when he says that autonomy is 'of value to people living in certain societies only.'[88] Wall similarly admits that his argument only attempts to establish the value of autonomy for people living in modern western societies that contain the conditions that make autonomy essential for well-being.[89]

Jeremy Waldron claims that if autonomy is valuable only because of the sort of environment that exists in autonomy-supporting societies, then the value of autonomy provides no reason to sustain an autonomy-supporting environment, just as the value of justice provides no reason not to eliminate scarcity and the value of courage provides no reason not to eliminate danger.[90] Raz's argument may support duties to protect autonomy, just as there is a duty for the government to act justly in the face of scarcity. But just as the government could at the same time attempt to eliminate scarcity, it could attempt to eliminate the social conditions which create the autonomy-based duties. It could try to eliminate the autonomy-supporting environment and replace it with one dominated by non-autonomous social forms. Raz's explanation of the value of autonomy provides no argument against such an action.

---

[86] Raz, *The Morality of Freedom*, p. 390.
[87] Raz, *The Morality of Freedom*, p. 408.
[88] Raz, *The Morality of Freedom*, p. 189, n.1.
[89] Wall, *Liberalism, Perfectionism, and Restraint*, p. 172.
[90] Jeremy Waldron 'Autonomy and Perfectionism in Raz's *Morality of Freedom*', *Southern California Law Review*, vol. 62, nos. 3 and 4, 1989, p. 1122.

Raz admits that, on his argument, those who advocate a return to a community where autonomous choice was minimal cannot be rebutted by claiming that in such a society, people's lack of autonomy makes them worse off. However, he argues that such a policy requires a large scale social design that would be a leap into the dark with no certainty that the results wouldn't be totally unlike what was hoped for. Normal politics, he says, rules out such radical changes, instead conducting affairs within existing though ever-changing social structures. This means that in normal politics, promoting the conditions of autonomy is essential for individual well-being.[91] However, it is possible that a non-autonomous society could be pursued gradually, with the encouragement of social forms that reduce individual choice and the discouragement of those that promote choice. Raz's arguments provide no objection against such a policy.

Raz suggests another possible response to this difficulty. Whilst not denying the possibility of living a worthwhile life in a non-autonomy-supporting society, he claims that it is false that any such life is better than any life with autonomy. Such alternatives are incommensurable.[92] However, I suggest that though it is true that the values of many goals and activities may not be comparable, incommensurability does not exist with regard to many of the comparisons that a perfectionist paternalist would have to make. Those comparisons involve the judgment that certain activities are more worthwhile than others and that people who engage in the former in societies where choice is limited have more well-being than those who choose the worthless in more free societies. Someone who has a challenging career and a loving family, though pushed into these by tradition is better off, for example, than a lonely couch potato in an autonomy-supporting society. Such comparative judgements seem possible. Moreover, in the absence of a sound argument for the value of autonomy

---

[91] Raz, 'Facing Up: A Reply', *Southern California Law Review*, vol. 62, nos. 3 and 4, 1989, p. 1228.
[92] Raz, 'Facing Up', p. 1228.

that would demonstrate otherwise, such judgments also seem correct. The social forms argument does not provide the needed argument.[93]

Wall has a different reply to Waldron's objection. He accepts that the social forms argument provides no reason for the state not to change the nature of the social forms that make autonomy necessary. And he seems to accept that comparisons between the quality of lives in different societies is possible. But he questions whether there is any reason to change the character of the social forms of autonomy-supporting societies. He claims that there is no reason to think that the autonomous lives of people in modern western societies are less good than the non-autonomous lives they could be living in alternative societies. Moreover, he says that there is good reason not to change the nature of society since change is disruptive and it is better to live in an environment that is familiar.[94]

There may, however, be a case for changing the character of an autonomy-supporting society. If many people are exercising their autonomy to pursue worthless goals, then they shall have lives of low value. If there is an alternative, non-autonomy-supporting, society in which people could live relatively good lives, then there is reason to change the nature of the social forms in society. There is evidence that people in modern societies make some important choices badly. The emphasis on free, rather than arranged, marriages according to one writer has meant that people often end up in worse marriages, ones that cause psychological harm, than they would under the arranged institution. Whilst arranged marriages did result in much harm and unhappiness, there was an even chance that one could marry a person whose personality would be compatible with one's own. Under free marriages, on the other hand, people's compulsive tendency to 'reproduce their childhood experiences in their

---

[93] There is not space here to consider more fully the complex issue of incommensurability. Raz discusses the issue in chapter 13 of *The Morality of Freedom*. Nothing he says there, I believe, casts doubt on my claims in the text.

[94] Wall, *Liberalism, Perfectionism, and Restraint*, p. 173.

marriages is jarred ... by the reality of the other person.'[95] If so, a perfectionist state would have reason to try to alter the nature of marriage, reducing the significance of individual choice in order to make it more feasible to limit people's choices (say by requiring the approval of marriage partners by state psychologists) without undermining well-being.

Autonomous decisions of lifestyle choice may also be less good than non-autonomous ones. Many worthwhile activities involve higher startup costs at their beginning than less worthwhile activities. The pleasures of passive consumption are high at first, whilst the psychic costs of say learning a musical instrument or becoming a painter are great. This means that even though the payoff later on of the latter type of activities is so high that they are overall the more worthwhile choices, people may fail to choose them, due to irrational time discounting or weakness of the will. Given a choice between options, people may 'not bring themselves to undergo the painful learning process that is required'[96] for the more worthwhile options. The social forms surrounding lifestyle and occupational activities in modern western societies may emphasise individual choice, but if the phenomenon just described is prevalent then there are perfectionist reasons for the state to try to change those social forms. Hence there are some reasons for a paternalistic state to change the nature of society to one that is not autonomy-supporting.[97]

---

[95] Gerald Dworkin, 'Is more choice better than less?', in his *The Theory and Practice of Autonomy*, Cambridge, Cambridge University Press, 1988 p. 74, quoting Philip Slater, *Earthwalk*, New York, Harper Row, 1974, p. 17.

[96] Elster, 'Self-realization in Work and Politics', pp. 103-7.

[97] Elster admits that state action to counter the phenomenon he describes could be objectionably paternalistic (p. 98). His solution is to advocate some form of collective self-paternalism to change people's preferences (ibid.). But this solution would succeed, I believe, only if the paternalism were unanimously self-imposed, something that is impossible to achieve in modern societies. Simple majority-based collective self-paternalism would remain objectionable, since it would still involve the imposition of conceptions of the good on those who do not accept them without their consent, unless there were unanimous support for majority rule, something else that does not exist in modern societies.

Leaving aside this objection, there is a more fundamental difficulty with the social forms argument. Even accepting that well-being is dependent upon social forms,[98] we need not agree with Raz's and Wall's insistence that social forms in modern western societies are of such a nature that they require people to be autonomous in order to participate in them. It is possible for a non-autonomous person to live a successful life in a society dominated by autonomy-supporting social forms.[99] Someone who marries out of his parents' wishes rather than his own choice could still find fulfillment in that marriage. A person forced by social pressures into a certain occupation could nevertheless successfully pursue a worthwhile occupation. A non-autonomous person in an autonomy-supporting society could still live a worthwhile life. If autonomy is not necessary for well-being, then this is important for the perfectionist case against paternalism. Perfectionism would be less able to resist its objectionably paternalistic implications. The state could force people into more worthwhile conceptions of the good, at the expense of autonomy, without damaging their well-being.

Raz considers the above objection to the claim that autonomy is a necessary condition for well-being. He asks whether there are any possibilities of a successful non-autonomous life in a society committed to autonomy.[100] It is a mistake to think there are such possibilities, he says, because it assumes that an autonomy-supporting society 'does not affect the nature of the opportunities it provides.'[101] It is not just that an autonomy-supporting environment provides more options. Rather, the nature of the options themselves are changed. The fact that they are goals that are freely chosen

---

[98] For arguments against this claim, see Roger Crisp, 'Raz on Well-Being', *Oxford Journal of Legal Studies*, vol. 17, no. 3, 1997, p. 509 and Donald Regan, 'Authority and Value: Reflections on Raz's *Morality of Freedom*', *Southern California Law Review*, vol. 62, numbers 3 and 4, 1989, pp. 1051-2.

[99] W. J. Norman makes this point in 'The Autonomy-Based Liberalism of Joseph Raz', *Canadian Journal of Law and Jurisprudence*, vol. II, no. 2, 1989, p. 155, as does Simon Caney in *Liberal perfectionism*, D.Phil. thesis, University of Oxford, 1994, pp. 254-6.

[100] Raz, *The Morality of Freedom*, p. 391.

[101] Raz, *The Morality of Freedom*, p. 392.

is part of what constitutes them.[102] An autonomy-supporting society, says Raz, 'offers its members opportunities which cannot be had in a non-autonomous environment, and lacks most of the opportunities available in the latter.'[103] Because the nature of the opportunities will be such that they require individual autonomous choice, there is no possibility that non-autonomous people can participate in them. However, this characterisation of the nature of opportunities in an autonomy-supporting society is too strong. Whilst it is true that the prevailing attitude toward marriage emphasises individual choice, the institution itself is compatible with a non-autonomous selection of partners. The social form of marriage was transformed into something new in the move from pre-arranged to self-chosen partnership. But though the new social form increased personal autonomy, it still allows for non-autonomous marriage. The same point holds for occupations. Although social practice emphasises individual choice of occupation, it is still possible for someone to be placed into an occupation without autonomously choosing it. Moreover, there is no reason to think that an autonomy-supporting society requires institutions of marriage and occupational choice that prohibit all but autonomously chosen marriages and jobs. Despite the choice-emphasising nature of social forms in an autonomy-supporting society, it is still possible for people to be forced into worthwhile forms without thereby closing access to those very forms.

Wall also considers the objection and has a different reply to it. He claims that the fact that people can live fully good, though non-autonomous, lives does not conflict with the conclusion of the social forms argument. To refute that argument, he says,

> one needs to present a way of life that is both fully good and incompatible with autonomy. This is true because if the way of life is compatible with autonomy, then it is plausible to claim ... that autonomy would make that life even better.[104]

---

[102] Raz, 'Facing Up', p. 1228.
[103] Raz, *The Morality of Freedom*, p. 392.
[104] Wall, *Liberalism, Perfectionism, and Restraint*, p. 170.

Wall argues that this cannot be shown since almost all goals are compatible with autonomy. Marriages, careers, and living life as a nun or in the military may all be engaged in non-autonomously, but they could also be chosen by an autonomous person. If goals are compatible with autonomy in this way, then non-autonomous lives that contain them do not refute the social forms argument since those lives would be even better if the goals were chosen autonomously.

Wall claims that autonomy would make those lives even better because of the value autonomy has as a result of his arguments that were examined in section 2.5. I argued that those arguments did not establish the value of autonomy. The values Wall referred to, such as giving meaning to life, having character, and responsibility, were compatible with paternalism that restricted autonomy. If so, then his argument here does not succeed. But even if his earlier arguments had succeeded, so that autonomy has value that improves the lives of those who have it, this is irrelevant for the issue at hand. Wall confuses the idea that autonomy is a necessary condition of well-being, which is what the social forms argument is supposed to show, with the notion that autonomy simply adds to the value of a person's life. Even if lives that lack autonomy could be improved if autonomy were added to them in the way he suggests, as long as it is possible to lead a good life without autonomy, even in an autonomy-supporting society, then autonomy is not necessary for well-being and the social forms argument is refuted. Wall's argument cannot therefore show that the social forms argument rules out paternalism. The state could paternalistically force people out of worthless ways of living and consequently they may, contrary to the social forms argument, live good lives even though lacking in autonomy. Their lives may have been even better, according to Wall's suggestion, had they autonomously chosen the goals that they are paternalistically guided into. But this was never a possibility; presumably what motivates the paternalism is that the paternalised do not and would not freely choose more worthwhile goals.

Hence, the nature of social forms in modern western societies do not make autonomy a necessary condition of well-being, thus leaving open the possibility of living a worthwhile, non-autonomous life. However, this does not show that the argument fails to rule out objectionable paternalism. That failure turns upon the possibility of there being successful (that is, the target is actually better off) autonomy-invading paternalism given the nature of social forms in modern western societies. For even if autonomy is not necessary for well-being, the social forms argument might still show that autonomy is so important for well-being that there is little likelihood of paternalism succeeding. The more autonomy-supporting an environment is, the less likely it is that a non-autonomous person will prosper. Although autonomy may not be necessary for well-being in such an environment, it may be extremely important. This revised position may be enough to resist most objectionable cases of paternalism.

The counter-examples to autonomy being necessary for well-being include someone inadvertently falling into a worthwhile activity, or being forced by social pressure into an occupation, or slavishly following the norms of his parents or community, thereby lacking autonomy but nevertheless leading a worthwhile life. What seems to make the resulting life still worthwhile is that its goodness rests on what there is in it rather than how it came to be. This is so even if, as Raz asserts, what there is in it is partly determined by how it came to be, in the sense that an autonomously chosen activity has a different character than the same activity not autonomously chosen. Even accepting this, it remains true that to some extent what there is in a life is independent of how it came to be. In fact, this follows from Raz's claim. If activities are *partly* constituted by being freely chosen, then presumably the rest of what constitutes them is their content, which is independent of how the activity becomes part of the life of the person who engages in it. Hence the counter-examples are worthwhile lives due to the content of those lives. But if so, then people can live a worthwhile life even if they are guided to valuable activities by the state for their own good. So long as the lives contain good activities then they can be worthwhile regardless of how they came

about. Thus, if, as seems true, it is possible for people to lead worthwhile non-paternalised though non-autonomous lives, then it is possible for people to lead worthwhile lives, even though subjected to autonomy-invading paternalism. The social forms argument, therefore, even if taken to establish that autonomy is important though not necessary for well-being, is still compatible with paternalism that forces people into worthwhile activities.

If, as I have argued, the social forms argument fails to show that autonomy is essential for well-being or even important enough to rule out much paternalism, why is it thought to be so important by Raz and Wall? It is the central argument for the value of autonomy of both. Surely, it may be thought, it must show something of importance. What it shows, I believe, is that in modern western societies autonomy is *possible* and *facilitated* to a degree never witnessed before. But being possible and facilitated is not the same as being near inescapable, which is the conclusion Raz and Wall attempt to draw. Geographic, technological, and social mobilities make autonomy along these dimensions available in a way people in other societies have never heretofore experienced. But it does not mean that these possibilities have to be taken up by people. They may stay (without choosing to do so in the sense required by autonomy) in the same place, pursue their family's traditional occupation, and associate only with members of the same social class without sacrificing their well-being. Social practices do strongly emphasise individual choice in the spheres of marriage and occupation. The result is that people can exercise choice in these spheres without pressure from others to conform to conventional norms. But choosing without such pressures is not necessary to have access to the options provided by social institutions. People can still live worthwhile, though non-autonomous lives in modern societies even though such lives are less widespread and more difficult to engage in than in previous times.

The difficulties with the social forms argument, with respect to paternalism, are that it is compatible with people living worthwhile non-autonomous lives, thereby permitting paternalism, and that it allows the state to change the nature of social forms. Donald Regan draws an argument for the value of autonomy from Raz's work that may be able to avoid these difficulties. Raz holds that a commitment constitutes a reason for action that did not exist prior to the commitment.[105] Suppose someone decides to become a writer for reasons such as that she enjoys the writing process, she has a talent for it, and is able to earn a reasonable income from what she produces. Committing herself to a writing career creates a new reason for her to pursue her writing. The other reasons still exist, but according to Raz she has a new reason for being a writer and that is that she has made a commitment to that project. Regan presents the following argument. Being able to make such commitments creates new possibilities for success or failure in one's life. Because commitments create new reasons for action and new opportunities for success and failure, it is valuable to be able to make commitments. It is not possible to be able to make commitments in the relevant sense unless one is autonomous. Therefore autonomy is valuable.[106] This argument makes autonomy valuable regardless of the nature of the social forms in society. And it may make autonomy a condition of goods, valuable because a necessary capacity for well-being for everyone.

Regan rejects this argument because of its reliance on the claim that commitments themselves are reasons for action. This, he says, is not the case. Rather, a commitment changes one's already existing reasons for action without the commitment itself being a new reason for action. He believes that it deepens one's appreciation of the project that has been committed to. The commitment creates new reasons but is not itself one of those new reasons.[107] But even if Regan is correct in

---

[105] Raz, *The Morality of Freedom*, p. 386.
[106] Regan, 'Authority and Value', p. 1078.
[107] Regan, 'Authority and Value', p. 1078.

claiming that a commitment changes one's reasons for action without itself being a new reason, the argument being considered could still succeed. Making commitments may change one's already existing reasons in a valuable way, by (as Regan claims) deepening one's already existing appreciation of the activity. It could be valuable to be able to make commitments, because it creates a new reason whether that is because the commitment itself or something else is the new reason. But the next step of the argument, that it is not possible to make commitments in the relevant sense unless one is autonomous, is false. One can be coerced or manipulated into a commitment, or make one without having had an adequate range of options to choose from, and it could still be the case that the commitment creates new possibilities for success or failure. Someone who becomes a writer because that is the best of a bad bunch of opportunities open to her nevertheless has new opportunities for success and failure. If however it is said that commitments made in these circumstances are not commitments 'in the relevant sense' then the argument simply becomes question-begging, if by 'the relevant sense' is meant the autonomous making of commitments.

2.7 Autonomy and Worthless Options

The arguments so far have failed to establish that autonomy has sufficient value to enable perfectionism to resist the charge of permitting too much paternalism. I have not tried to show that autonomy has no value at all. Rather, the value that it does have is compatible with much, though not all, paternalism. There is another difficulty with the perfectionist appeal to the value of autonomy. Even if autonomy has value, the autonomous life, as some perfectionists have realised, is valuable only to the extent that it is engaged in worthwhile goals. As Raz claims, an '[a]utonomous life is valuable only if it is spent in the pursuit of acceptable and valuable projects and relationships.'[108] If true, this would mean that the value of autonomy gives even less protection against paternalism. Paternalism that denies autonomous choice of worthless goals would not make the lives of those interfered with any worse, at least

[108] Raz, *The Morality of Freedom*, p. 417.

not worse because less autonomous. Hence perfectionism would provide little protection against paternalistic intervention with worthless activities.

Before examining Raz's argument for this proposition, an objection of Regan's warrants consideration. Regan claims that there is a contradiction between saying that autonomy is valuable on the one hand, and on the other saying that autonomy is valuable only if exercised in the pursuit of the good. The first claim seems to imply that autonomy is valuable even when exercised in the pursuit of worthless options, which is what the second claim denies.[109] Raz points out however that there is no inconsistency. Autonomy is intrinsically good as a constituent of well-being in modern societies, and well-being is good in itself. Remove other elements of this good, says Raz, such as the worthwhile activities that determine well-being, and it may turn worthless (even though autonomy is still present).[110] Thus whilst autonomy is valuable because one is unlikely to prosper without it, an autonomous life does not contribute to well-being unless it is spent in worthwhile pursuits. Autonomy is valuable because of its importance for well-being. But it is not sufficient for well-being.

Whilst Raz's position is consistent, it may be doubted that an autonomous life is valuable only if exercised in pursuit of worthwhile activities. Perhaps autonomous choice always contributes to personal well-being. Raz presents the following argument:

> has autonomy any value *qua* autonomy when it is abused? Is the autonomous wrongdoer a morally better person than the non-autonomous wrongdoer? Our intuitions rebel against such a view. It is surely the other way round. The wrongdoing casts a darker shadow on its perpetrator if it is autonomously done

---

[109] Regan, 'Authority and Value', p. 1084. The same objection is made by Robert George in *Making Men Moral*, Clarendon Press, Oxford, 1996, p. 175 and by Margaret Moore in 'Liberalism and the Ideal of the Good Life', *Review of Politics*, vol. 53, no. 4, 1991, p. 687.
[110] Raz, 'Facing Up', p. 1229, n. 162.

> by him. ... Demeaning, or narrow-minded, or ungenerous, or insensitive behaviour is worse when autonomously chosen and indulged in.[111]

Waldron criticises this argument. He claims that it makes no difference to our assessment of whether an act is right or wrong that it was performed autonomously or not, just as our assessment of the wrongness of murder is not affected by whether it was done courageously, or our assessment of the wrongness of fraud is not affected by whether it is done with prudence and temperance or not. Virtues such as courage, prudence, and temperance, says Waldron, have evaluative significance without making the acts they qualify better or worse. Autonomy, he says, is a concept that works in a similar way. Waldron claims that Raz needs to show that autonomy is not like other virtue concepts in this regard. Without doing so, Raz cannot make the comparison contained in the above argument.[112]

But this is precisely what Raz does, at least partly, show in his appeal to intuitions. Waldron may be correct when he says that courageousness makes no difference to the wrongness of murder, or that prudence makes no difference to the wrongness of fraud. But as Raz says,

> A murderer who was led to his deed by the foreseen inner logic of his autonomously chosen career is morally worse than one who murders because he momentarily succumbs to the prospect of an easy gain.[113]

And '[n]eglecting one's child freely and autonomously is worse than neglecting one's child out of necessity and without real choice.'[114] Autonomy does seem to make the bad acts it qualifies worse. Waldron fails to provide any reason for doubting these intuitions. The burden of the argument is as much on Waldron to explain why autonomy plays a similar function to the other virtues he mentions, as it is on Raz to give an explanation for the intuitions. In his reply to Waldron, Raz argues that since

---

[111] Raz, *The Morality of Freedom*, p. 380.
[112] Waldron, 'Autonomy and Perfectionism', pp. 1127-8.
[113] Raz, *The Morality of Freedom*, p. 380.
[114] Raz, 'Facing Up', pp. 1228-9, n. 160.

autonomy is a property of life rather than a virtue, the analogy that Waldron relies upon does not hold up.[115] This I think is an inadequate reply since Waldron could admit that autonomy is a property of life but insist that it is similar to the virtues he points to in that it has evaluative significance without making the acts it qualifies better or worse. Raz's examples, given above, are a better reply to Waldron's objection.

However, even though the intuitions Raz appeals to seem plausible, they merely show that autonomously harming others is morally worse than non-autonomously doing so. But that does not show that autonomously choosing a worthless way of life, putting aside its effects on others, is not better[116] for one's well-being than living an equivalent non-autonomous life. If it is better, then autonomy makes some contribution to well-being even when exercised to choose worthless goals and activities. Our intuitions with regard to autonomous harm to self are not as strong as they are with regard to autonomous harm to others. It is not so clear that someone who autonomously decides to lie on the sofa all day every day drinking beer and watching soap operas is leading as bad a life as someone who non-autonomously does the same thing, as it is that the autonomous murderer is worse than the non-autonomous one.

Perhaps an explanation of the latter type of case can provide guidance in the former. What makes the autonomous murderer so immoral may be the lack of respect he shows toward his victim by carefully considering the available options and deliberately choosing an immoral one. A murderer who momentarily succumbs to the prospect of an easy gain, while immoral and lacking respect for his victim, does not show the same or as much lack of respect. Analogously, perhaps someone who autonomously chooses a worthless way of life is showing disrespect towards himself

---

[115] Raz, 'Facing Up', p. 1228, note 160.

[116] 'Not better' rather than 'worse' since the former is all that is needed in order to disprove the claim that autonomy has value even when it is exercised in the pursuit of worthless activities.

and in virtue of that leads a worse life than someone who leads a similarly worthless life in a non-autonomous way. This argument does not work however since it need not be the case that someone who autonomously chooses worthless goals or activities is showing disrespect towards himself. It could be said of somebody who recognises the worthlessness of a goal but nevertheless autonomously chooses it that he lacks respect for himself. But someone could make a genuine mistake and believe that a goal that has no value is worthy. She may contemplate the decision carefully, comprehending the alternative options open to her, and decide to pursue what she mistakenly believes is a valuable activity. She would be making a mistake about what is valuable, and her well-being is less because of that mistake, but no more so than the non-autonomous person who pursues a similarly worthless option. Her life does not go worse due to any disrespect she shows to herself, since she has made an honest mistake.

Waldron presents an alternative argument for believing that an autonomous life contributes to one's well-being only if spent in the pursuit of the good, since he thinks that Raz's argument fails. Waldron draws on the idea, explained by Raz at length[117], that people choose projects and activities for reasons pertaining to the value they see in those projects and activities. People value autonomy, says Waldron, because it allows them to choose goals for the reasons that they apprehend, and shape their lives themselves in accordance with those reasons. He concludes that

> [o]ne does violence to the self-understanding of people aspiring to autonomy if one advises them that the moral quality of their choices is unimportant so far as the value of their autonomy is concerned.[118]

This argument as it is stated, if it establishes anything, establishes that autonomy is valuable only if exercised in the pursuit of what a person *believes* is worthwhile. If autonomy is valuable to the extent that people value it, and they value it because it

---

[117] Raz, *The Morality of Freedom*, pp. 300-5.

[118] Waldron, 'Autonomy and Perfectionism', p. 1129.

allows them to conduct their lives according to what they believe is worthwhile, then autonomy is valuable only to the extent that it is spent in the pursuit of what is believed worthwhile. But that differs from the proposition that autonomy is valuable only if it is exercised in the pursuit of what is *truly* valuable since people may mistakenly believe something is worthwhile when it is not and mistakenly believe something is worthless when it is valuable.

Of course, the argument may be adjusted by saying that people value autonomy because it allows them to choose worthwhile goals, rather than simply what they believe is worthwhile. Moreover, this is a more accurate explanation of the situation, since people would be less likely to value autonomy if it always resulted in them choosing the worthless in the mistaken belief that it is worthy. But this adjusted argument is question-begging. People do value their autonomy because it allows them to pursue the good. Does this mean, however, that they see *no* value in the exercise of their autonomy when they mistakenly aim at the bad? If Waldron's answer is in the affirmative then he will have to explain why this is so, and this explanation will be precisely what he thought he'd shown - an argument to demonstrate that autonomy has no value when exercised in the pursuit of the worthless.

Despite the failure of all of the arguments so far to establish that autonomy contributes to well-being only when spent in the pursuit of the worthwhile, I believe a successful argument can draw upon Raz's explanation of the value of autonomy in general. That argument was the following. Given that many social forms in our society are based on individual choice, it is difficult for a person who lacks autonomy to participate in those social forms, and such participation determines one's well-being. Against this, I argued that it is possible for someone to have a worthwhile life, and even be paternalistically guided into one, without autonomy. Even if so, however, the point remains that for many people autonomy remains important because they often need it in order to participate in those social forms that do require individual choice. But,

according to Raz, participating in social forms contributes to one's well-being only if the forms participated in are worthwhile ones.[119] Participating in worthless ones does not improve the goodness of one's life. This seems obviously true, perhaps even a tautology, but is the important point for the argument I am making. According to the general argument, autonomy has value because without it one is unlikely to be able to engage in worthwhile activities. If one accepts this argument, along with the claim that well-being is furthered by participation in worthwhile though not worthless goals, then one could believe that autonomy is important for engaging in social forms based on individual choice, but one could not think that doing so furthers one's well-being if the goal associated with the social form is a worthless one. That in turn means that autonomy exercised in the pursuit of the bad does not contribute to well-being. In short, if autonomy is valuable solely because it is needed in order to engage in worthwhile activities, then it is difficult to see how it can have value when used to engage in worthless activities. The autonomous life, then, is valuable only if it is exercised in the pursuit of worthwhile activities.

One implication of this argument is the following. The autonomous life spent in the pursuit of worthless activities is no worse than the non-autonomous one. They are equally bad lives. Raz however seems to suggest that the autonomous worthless life is worse than the non-autonomous one. He claims that 'autonomously choosing the bad makes one's life worse than a comparable non-autonomous life is.'[120] This was his conclusion from the intuitive argument discussed above, regarding the autonomous murderer being worse than the non-autonomous one. But the conclusion must be abandoned in light of the argument I am making.[121] If autonomy gets its value from

---

[119] Raz, *The Morality of Freedom*, pp. 298-9.

[120] Raz, *The Morality of Freedom*, p. 412. The same claim is also made on p. 380 and in 'Facing Up', p. 1228, note 160.

[121] A difficulty arises in that I relied upon the plausibility of the intuitions in order to reply to Waldron's objection to the intuitive argument. If the intuitions must be abandoned then may that reply still be made? It can since the reply depends upon the plausibility of the intuition that autonomously harming others is worse than non-autonomously doing so. The argument being made in the text requires us to abandon the conclusion that an autonomous, though worthless live is worse for the person that lives it

providing access to valuable social forms, then it loses this value with respect to its providing access to valueless social forms. But it does not, or I see no reason for it to, have negative value in this latter type of case.

This more moderate conclusion helps the perfectionist in one respect. The more extreme view that Raz seems to hold provides more scope for paternalism. If autonomously choosing the bad has negative value, then a greater benefit could be brought about by successful paternalism that guides the target into a more worthwhile option, compared with the benefit that could be produced if autonomously choosing the bad is worthless, with neither negative nor positive value. If autonomous worthless choice does not have negative value, then there is less of a benefit that could be brought about by paternalistic intervention. So here my argument helps perfectionism avoid paternalism in one way. But it is still the case that even on the moderate conclusion there would be enough of a benefit at stake to motivate paternalism.

2.8 Conclusion

The value of autonomy fails to rule out much paternalism that would ordinarily be thought objectionable. Some arguments for the value of autonomy make it incompatible with paternalism, but in any plausible perfectionist theory that value is only one amongst others. Furthering well-being requires weighing up all perfectionist values and overall perfection would in many cases be furthered with the use of paternalism. Arguments that attempt to give autonomy more value to avoid these trade-offs are either unconvincing or compatible with paternalism. The argument that tries to show that autonomy is a condition of goods in modern western societies is also unsuccessful. It exaggerates the difficulty of living a worthwhile, non-autonomous, paternalistically guided life in such societies. Moreover, there seems reason to think

---

than a similarly worthless non-autonomous one. We can abandon this conclusion while retaining the other intuitions. I have no explanation for this asymmetry here.

that autonomy has no value if spent in the pursuit of worthless goals, hence making the value of autonomy compatible with paternalistic interference with choices for such goals.

The failure of the arguments for autonomy to rule out paternalism may come as a surprise, but it should not do so. The reasons that often lead people to think that valuing autonomy is incompatible with advocating paternalism are mistaken. Being forced into, say, a monotonous factory job for one's entire life is both a bad life and a non-autonomous one. But it should not be concluded from this that a good life is impossible without autonomy or that autonomy always leads to a good life. Another widely held belief is the thought that paternalism always invades autonomy. Being forced into a way of life for one's own good is a paternalistic invasion of autonomy. But it would be a mistake to conclude from this that all cases of paternalism invade autonomy. Even if most of them do, it still does not follow (as I have tried to argue) that autonomy always has sufficient value to rule out cases of paternalism that conflict with it. Autonomy may be valuable but it does not follow from this that its value trumps all others. Another motivation for belief in the incompatibility is the thought that paternalism stunts the capacities required for autonomy. But again it need not. Paternalistic intervention, even if coercive, can block certain options from the paternalised without damaging her abilities for rational self-direction. Finally, there is the thought that paternalism, perhaps because it assumes that another person knows my own good better than I do, must necessarily express a lack of respect for me as an autonomous agent. This, however, invokes a conception of autonomy different to the perfectionist understanding of autonomy. As was seen in section one, to conceive of autonomy as a restraint on how someone may be treated, regardless of whether she may be made better or worse off by being so treated, differs from conceiving of autonomy as a value that improves the life of those who have it. It is the latter understanding that perfectionists rely upon in their arguments. In this perfectionist sense there is no incompatibility between autonomy and paternalism. Once these points are grasped, and given the arguments in this chapter, it is clear that perfectionism cannot rely upon the value of autonomy to rule out its objectionably paternalistic implications.

Chapter Three

Raz on Coercion and Paternalism

3.1 Introduction

In the previous chapter, I tried to show that several arguments, including Raz's, for the value of autonomy fail to rule out many cases of objectionable paternalism. Raz has other arguments relevant for paternalism, claiming that the value of autonomy is 'no more than part of the reason for our concern for liberty.'[1] Several of these arguments all attempt to establish that coercive paternalism is unjustified. While the value of autonomy, according to Raz, provides one explanation of the wrongness of coercive paternalism, there are others. Raz claims that independence, which requires the absence of coercion and manipulation, has an importance that 'transcends its role as a component or precondition of personal autonomy.'[2] Well-being, according to Raz, consists in the willing and whole-hearted engagement in worthwhile relationships and pursuits. He claims that coercion is incompatible with the whole-heartedness aspect of this conception of well-being.[3] If so, then this component of well-being may provide a means by which perfectionism can resist strongly paternalistic implications. This strategy, I will try to argue, is not successful. Raz admits that there are exceptions to the prohibition of coercion. As I hope to show below, once these exceptions are allowed then there is no convincing argument for why coercive paternalism should not also be acceptable.

---

[1] Joseph Raz, 'Liberty and Trust', in Robert P. George (ed.), *Natural Law, Liberalism, and Morality.* Oxford, Clarendon Press, 1996, p. 113.

[2] Raz, 'Liberty and Trust', p. 120.

[3] Raz, 'Liberty and Trust', p. 121.

Section two explain's Raz's position on coercive paternalism, partly by way of an objection made against it by Donald Regan. Raz argues for a version of the harm principle, which is usually taken to prohibit paternalism. In the third section, I try to show that two arguments for the harm principle and against coercive paternalism that Raz makes in *The Morality of Freedom* are unsuccessful. Section four examines an argument Raz makes in a subsequent article. I conclude that this argument too is unsuccessful.

3.2 Coercive Paternalism

Mill's harm principle states that the only purpose for which coercion may justly be used, against a person's will, is to prevent harm to others.[4] This rules out all coercive paternalism that is against a person's will. It would therefore rule out many of the objectionable cases of paternalism that perfectionism seems to provide a motivation for. But Raz endorses a different version of the harm principle. Raz's version of the principle holds that coercive interference is justified only if it is used to prevent harm, whether to others or the person coerced.[5] This provides more scope for paternalistic practices. It permits the coercion of people to prevent them from harming themselves.

Regan argues that this interpretation of the harm principle permits a wide range of paternalism. He emphasises that, as Raz recognises, one can harm a person not only by making him worse off than he is, but also by failing to make him as well off as he ought to be. If you owe me five pounds, for example, then you harm me by delaying its repayment even though I am not worse off than before.[6] According to Regan, this means that if the government has a duty to promote the pursuit of worthwhile goals,

---

[4] John Stuart Mill, *On Liberty*, Norton Critical Edition (David Spitz, ed.), New York, Norton, 1975, ch. 1, par. 9: 'the only purpose for which power can be rightfully exercised over any member of a civilised community, against his will, is to prevent harm to others.'

[5] Joseph Raz, *The Morality of Freedom*, Oxford, Clarendon Press, 1986, p. 413.

[6] Raz, *The Morality of Freedom*, p. 416.

then that government harms a person who is pursuing worthless projects by not making him as well off as he ought to be.[7] According to Raz's version of the harm principle, there would be no prohibition on coercively preventing people from pursuing worthless projects, despite Raz's claim to the contrary.[8]

However, this objection interprets harm more broadly than it is meant to be interpreted in Raz's theory. While Raz does say that people are harmed by not being made as well off as they ought to be, as well as being made worse off, the notions of harm and better- and worse off here are to be interpreted in terms of the conditions of autonomy.[9] Harm, on Raz's interpretation, consists only in depriving a person of the conditions of autonomy or failing to secure those conditions insofar as there is a duty to. One harms another by depriving her of, or failing to secure for her, an adequate range of options and the capacities needed for autonomy. Certain injuries and pains, so long as they are not incapacitating, are not to be counted as harms, and neither is offence if it is not serious and persistent (but incapacitating pain and serious and persistent offence are harms since they reduce a person's opportunities and capacities required for autonomy). Nor is choosing worthless activities to be counted as harming oneself. The harm principle, under the current interpretation, supports coercion to prevent harm to a person's autonomy, whether caused by others or that person herself. But, as Raz says, the harm principle excludes 'the use of coercion to discourage non-harmful opportunities.'[10] Thus people who pursue worthless goals are not harming themselves in a way that may justify coercive interference. Paternalistic coercion may be justified, according to Raz, only if it is to protect or promote a person's autonomy.

---

[7] Donald Regan, 'Authority and Value: Reflections on Raz's *Morality of Freedom*', *Southern California Law Review*, vol. 62, 1989, p. 1085, n. 207.

[8] Regan makes a stronger claim. He states that the harm in this case would justify coercive intervention upon Raz's version of the harm principle ('Authority and Value', p. 1085, note 207). But it is important to note, however, that the principle need not sanction such measures. It is a necessary but not sufficient condition for coercive interference. As Raz notes (*The Morality of Freedom*, p. 420) it does not justify all use of coercion to prevent harm, but just proscribes the use of coercion for other purposes.

[9] For this interpretation see pp. 413-4 of *The Morality of Freedom*.

[10] Raz, *The Morality of Freedom*, p. 418.

While this interpretation of the harm principle rules out many cases of unwarranted paternalism, there are others that it does not. Consider, for example, the case of lifestyle paternalism discussed in section 2.4 which threatened couch potatoes with legal sanctions in order to encourage them to take up more worthwhile activities. Such action could meet the condition set by the harm principle since being a couch potato may set back one's ability to live an autonomous life. The mental capacities needed for autonomy would become dulled by lifestyles that do not involve their exercise. And because of this many options would become closed to people living those lifestyles, perhaps to such a degree that they fail to have an adequate range of options with which they can exercise autonomy. One needs particular abilities in order to have the option of engaging in certain activities. Coercing people out of autonomy-dulling activities may help them become and remain autonomous, and thereby meet the condition for justified coercive interference set by the harm principle.

The harm principle, even on Raz's interpretation, does not therefore rule out all cases of objectionable paternalism. But it does rule out many. However, we need to know the justification for restricting coercion to only the protection and promotion of autonomy. That 'harm' can be given an interpretation is not enough. We need a justification for the principle under that interpretation. Raz suggests that the harm principle is supported by his doctrine of autonomy, elements of which were examined in the last chapter. He does not explicitly set out the derivation, but a plausible case can be made along the following lines. The value of autonomy, Raz argues, results in three types of duties to others: those to help create the personal capacities for autonomy, those to ensure others face an adequate range of options, and those to refrain from coercing or manipulating others. The principle of autonomy, says Raz, is the principle requiring that people secure these conditions of autonomy for others.[11]

[11] Raz, *The Morality of Freedom*, p. 408.

The duties may sometimes conflict, as when the only way to ensure an adequate range of options for some is to coerce others, say to make them pay taxes in order to provide opportunities for all. In the event of such conflicts, it is plausible to suppose that the duties may be overridden for the sake of ensuring compliance with other autonomy-derived duties. So independence may be invaded in order to secure for everyone the personal capacities needed for autonomy, an adequate range of options, and independence. Hence we end up with something like the harm principle: coercion may be justified in order to secure the conditions of autonomy.

One way of stating the conclusion of the last chapter is to say that Raz's argument for the value of autonomy fails to establish the third category of autonomy-based duties forbidding coercion and manipulation, at least not in a way that rules out paternalism. The social forms argument does establish that forcing someone into a way of life may result in them being unable to have a worthwhile life. But that argument for the value of autonomy does not support a duty against coercing people away from worthless options. If so, then this makes invalid the derivation of the harm principle from the principle of autonomy. There would be no duty against all coercion that has to be weighed against the other duties of autonomy. Without such a duty, there is no prohibition on using coercion to prevent non-harmful but worthless activities.

However, even if we grant the premises and accept that there is a duty not to coerce or manipulate, the argument for the claim that the harm principle follows from the autonomy principle is still not yet successful. The argument is that there is a duty not to coerce because coercion invades autonomy which is essential for well-being. But, the argument continues, sometimes coercion is justified in order to secure the other conditions of autonomy which are similarly essential for well-being. The problem with this argument is that it needs to be explained why coercion cannot also be justified in other cases. An explanation is needed because if coercion is justified in order to secure the conditions of autonomy because autonomy is essential for well-

being then it seems also to be justified to provide other essentials to well-being. And one of the essentials for well-being on a perfectionist (and any plausible) account is that a person engage in worthwhile rather than worthless goals, projects, and activities. So we need an explanation for why coercion is not justified to guide people into valuable activities. Why then is such coercion ruled out?

3.3. Two Arguments

Raz has two replies to this question. Firstly, coercive interference 'violates the condition of independence and expresses a relation of domination and an attitude of disrespect for the coerced individual.'[12] Elsewhere, Raz writes that the coerced person 'is being treated as a non-autonomous agent, an animal, a baby, or an imbecile.'[13] Secondly, coercion is a global and indiscriminate invasion of independence. Raz claims that 'there is no practical way of ensuring that the coercion will restrict the victims' choice of repugnant options but will not interfere with their other choices.'[14]

With regard to the second argument, as several commentators have pointed out, coercive measures need not be as indiscriminate as Raz maintains. Even if imprisonment does prevent a person from engaging in almost all pursuits, the threat of imprisonment may successfully prevent people from engaging in worthless activities without interfering with their other choices. Moreover, the state could coercively threaten only the providers of opportunities for repugnant options and not the choosers.[15] Paternalism could confiscate items such as drugs and pornography rather than imprison people.[16] There are 'narrowly tailored, precise coercive measures which focus sharply on prohibitions of 'morally repugnant' options, and where the spill-over

---

[12] Raz, *The Morality of Freedom*, p. 418.
[13] Raz, *The Morality of Freedom*, p. 156.
[14] Raz, *The Morality of Freedom*, p. 419.
[15] Regan, 'Authority and Value', p. 1082.
[16] See Simon Caney, *Liberal perfectionism*, D.Phil. thesis, University of Oxford, 1994, pp. 190-2 for this and other objections to Raz's argument about clumsy coercion..

effect will be minimal."[17] Examples include cases of paternalism such as coercing couch potatoes out of their lifestyles, censorship laws which restrict aesthetically bad music, and requirements to listen to state-provided relationship advisors (although the latter especially would have to be carefully tailored to avoid invading worthwhile options).

With regard to the first argument, it is difficult to see why respect for a person requires respecting his choices for worthless activities. On the contrary, paternalistically coercing a person away from repugnant goals may be motivated by respect for that person.[18] Raz himself admits that treating people with respect requires taking into account the moral quality of their chosen activities. He claims that '[o]ne would be showing disrespect to another if one ignored moral considerations in treating him."[19] On the other hand, perhaps coercive interference ignores certain moral considerations. But it is not clear what these considerations could be. It cannot be that autonomously choosing worthless options endows those choices with value since Raz denies this. Paternalism that coerces people away from worthless options may invade autonomy but does not do so in any important sense, since autonomy exercised in the pursuit of worthless activities has no value.[20]

Perhaps the moral consideration that is ignored by coercive interference is a duty not to coerce, a duty that holds regardless of the quality of autonomous choices. Raz seems to hold this position, but it is not clear that there is such a duty. As was seen in section 2.6, Raz's argument for the value of autonomy does not establish the third category of autonomy-based duties forbidding coercion and manipulation, at least not in a way that rules out paternalism. The social forms argument does not provide any

---

[17] Wojciech Sadurski, 'Joseph Raz on Liberal Neutrality and the Harm Principle', *Oxford Journal of Legal Studies*, vol. 10, no. 1, 1990, p. 132. Leslie Green argues along similar lines in 'Un-American Liberalism: Raz's Morality of Freedom', *University of Toronto Law Journal*, 38, 1988.

[18] Sadurski, 'Joseph Raz on Liberal Neutrality and the Harm Principle', p. 132.

[19] Raz, *The Morality of Freedom*, p. 157.

[20] See section 2.7 for an argument for this claim.

support for a duty not to force people away from worthless options. It supports a duty to refrain from using coercion when doing so would make it impossible for the coerced person to participate in social forms. But not all coercion, so I argued in 2.6, makes such participation impossible. The value of autonomy upon the social forms argument does not support a duty against all coercion.

Perhaps there is support for such a duty from another source. Raz suggests that coercion undermines the whole-heartedness aspect of personal well-being. But even if so, since the good life is one of whole-hearted commitment to *valuable* activities, there would be no loss of well-being when the state coerces people who are whole-heartedly pursuing valueless activities. There may be a gain in well-being if the person chooses something worthwhile instead and does so with some degree whole-heartedly.[21] Hence the argument does not provide a duty against all coercion.

Even if this argument (or any other) does support a duty against all coercion, we may ask if the duty can ever be overridden. As we have seen, Raz maintains that it may be but only to prevent harm, that is to protect and promote autonomy. But then we can ask why it may not also be overridden in other cases in order to provide other essentials for well-being. We are then back to our starting point of asking why coercion is unacceptable in these other cases.

Nevertheless Raz maintains that coercive interference violates autonomy and thus displays an attitude of disrespect towards the coerced. At several points in *The Morality of Freedom*, Raz says that the usual prohibition against coercion because of its expression of domination and disrespect does not always hold. He claims that there 'is nothing wrong with coercion used to stop one from stepping into the road and

---

[21] This argument is made by T. M. Wilkinson in his review of *Natural Law, Liberalism, and Morality* in *Utilitas*, vol. 11, no. 1, 1999, p. 136.

under a car'[22] since this is an exception to the conventional symbolic meaning associated with coercion. More generally, he says that many cases of private coercion may be justified when 'the good intentions of the coercer are not in doubt, and therefore no insult or indignity is involved.'[23] Raz says that in an ideal state where people are guaranteed adequate rights of political participation, and where the state

> is guided by a public morality expressing concern for individual autonomy, its coercive measures do not express an insult to the autonomy of individuals. It is common knowledge that they are motivated not by lack of respect for individual autonomy but by concern for it.[24]

If coercion may be justified under these conditions, when the good intentions of the coercer are not in doubt and when it is localised and doesn't impinge on the adequacy of options, then this could allow a wide range of paternalism, certainly more than is justified by Raz's version of the harm principle. If the good intentions of the perfectionist state are made clear, then it may use coercive paternalistic measures in excess of those allowed by the autonomy-interpretation of the harm principle to further people's well-being.

Raz thinks that the condition requiring good intentions is one that governments will struggle to meet. He claims that the expressive meaning associated with coercion can seldom be overcome, particularly by governments. He says that 'conditions justify suspecting governments of lacking respect for individuals.'[25] Raz admits that his argument depends upon the contingent facts that governments lack respect for people and that there are no means of coercion that are not global and indiscriminate.[26] He considers the possible case where these contingencies do not hold. A state using

---

[22] Raz, *The Morality of Freedom*, p. 378.

[23] Raz, *The Morality of Freedom*, p. 421, note 1. The other conditions that need to be satisfied for private coercion to be justified are that it be localised so that denying the coerced one option does not interfere with his ability to choose others and that those others remain adequate so that there is no loss of autonomy.

[24] Raz, *The Morality of Freedom*, p. 157.

[25] Raz, *The Morality of Freedom*, p. 419.

[26] Raz, *The Morality of Freedom*, p. 419.

coercive measures designed to prevent immoral but harmless conduct may meet his conditions. It could not be mistaken about what is worthwhile, not abuse its power, its coercive measures could be localised so that they do not restrict other options, the condition of adequacy of options could still hold, and it could have good intentions and thus be free from suspicion that it does not respect its subjects.[27] Raz claims that such a situation is too abstract to yield any concrete recommendations. The imagined circumstances 'diverge from anything we have experience of sufficiently to make it impossible for us to say how the change would affect the merits of the issue.'[28] The change in circumstances is substantial enough to bring about not only a change in the application of our values, but a change in the values themselves. These changes are, says Raz, 'as a matter of principle, unpredictable.'[29]

However, Raz overstates the abstractness of the situation and the difficulty of making a judgment of the issue. There are easily imaginable (indeed in some countries actually existing) cases such as prohibitions on enervated lifestyles and aesthetically bad artistic expressions where the coercion is localised so that no other options are interfered with and people still have an adequate range of options available. Moreover, in those cases there could be no doubt about the sincerity of the state's intentions. Those who advocated, voted for, and implemented the policies may have truly believed (and believed truly) that they are for people's own good. Coercive perfectionist paternalism would not be rule out by Raz's conditions. Moreover, the circumstances are sufficiently connected with our experience to allow us to make such a judgement. There is little reason to think our values would be changed in these circumstances.

---

[27] Raz, *The Morality of Freedom*, p. 419.
[28] Raz, *The Morality of Freedom*, p. 419.
[29] Raz, *The Morality of Freedom*, p. 420.

These two arguments of Raz's, therefore, fail to provide a compelling case for restricting coercive paternalism to only those cases permitted by his version of the harm principle.

3.4 The Trust Argument

In an article written since *The Morality of Freedom*, Raz attempts to develop the above argument to show that coercive paternalism ought to be prohibited. As in the earlier work, he claims that coercive paternalism is justified only if the good intentions of the coercer are not in doubt.[30] This, he argues, supports the second of two conditions for justified paternalistic coercion:

> first, it is undertaken for a good reason, one sufficient to make reasonable a partial loss of independence. Secondly, barring emergencies, it comes from the hands of someone reasonably trusted by the coerced.[31]

Saying that the coercer must be *reasonably* trusted by the coerced, according to Raz, makes the condition semi-objective. Some people may be deluded into believing that the government is taking their interests seriously when in fact it is not. Coercive paternalism is not justified in these cases.[32] As Raz presents these conditions, they together are sufficient for justified coercive paternalism. But he also seems to suggest that the second is a necessary condition. He says that trust is 'a precondition for justified paternalistic coercion.'[33] Moreover, making it a necessary condition is essential for Raz's aim, which is to show how perfectionism need not involve great infringement of individual liberty.[34] If the conditions are merely sufficient then that leaves open the possibility that coercive paternalism may be justified without satisfying those conditions, thus leaving perfectionism vulnerable to the initial objection. Finally, as will be seen shortly, the conclusion that Raz believes follows

---

[30] Raz, 'Liberty and Trust', p. 122.
[31] Raz, 'Liberty and Trust', p. 122. Footnote omitted.
[32] Raz, 'Liberty and Trust', p. 127.
[33] Raz, 'Liberty and Trust', p. 123.
[34] For Raz's statement that this is his aim, see 'Liberty and Trust', p. 115.

from these two conditions follows only if the trust condition is taken as necessary for justified coercive paternalism.

Raz argues that most coercive state paternalism fails to satisfy the trust condition. He claims that before a person may reasonably trust the government, she must enjoy full citizenship. She does so when she feels (and is not unreasonable to so feel) that the government recognises her as a person who matters in her own right, that her fate is a matter of intrinsic value in the eyes of the state.[35] She reasonably believes that her interests are given due consideration by the government when it makes its decisions. The condition of citizenship contains a subjective element. It depends not just on how the government treats a person but on how she feels she is being treated, so long as that feeling is not unreasonable. According to Raz, citizenship is a precondition of trust. He writes that,

> [t]hose who find it impossible to regard the government as their government, and to regard themselves as full citizens alongside everyone else, are unable to trust the government.[36]

In these circumstances, where there is full citizenship and a resulting trust in the government, the second justifying condition of paternalism is met.

But according to Raz, most paternalism undermines the existence of full citizenship. He says that if the state

> pursues coercive *moral* paternalism against me it will, by definition, be preventing me from following my way of life, and it denies, in a purported exercise of its authority, the validity of propositions I hold true and which underpin my way of life. If it does so, however, it denies me full citizenship.[37]

---

[35] Raz, 'Liberty and Trust', p. 124.

[36] Raz, 'Liberty and Trust', p. 127.

[37] Raz, 'Liberty and Trust', p. 127. Emphasis in original. I assume from this passage that *moral* paternalism is that which is perfectionist, that is, guided by a conception of the good rather than by neutral ends. Moral paternalism prevents its target from following his way of life and tries to guide him into a more worthwhile one.

In doing so it 'undercuts my trust that I and my interests are seriously being taken into account in deciding public action.'[38] Someone whose fundamental values and beliefs are rejected as worthless by the government will not trust the government as acting in her own interests. Hence, according to the argument, coercive moral paternalism is unjustified.

Raz's argument can be formally restated in the following way:

1. Coercive paternalism is justified only if the paternalist is someone reasonably trusted by the coerced.

2. The state may be reasonably trusted by a person only if she enjoys full citizenship, that is, she feels that the government recognises her as someone who matters intrinsically.

Therefore,

3. Coercive paternalism is justified only if the coerced enjoys full citizenship.

4. A person cannot feel that the government recognises her as someone who matters intrinsically if it coercively prevents her from following her way of life.

5. Coercive moral paternalism prevents a person from following her way of life.

Therefore (from 4 and 5),

6. A person cannot feel that the government recognises her as someone who matters intrinsically, that is, she cannot enjoy full citizenship, if subjected to coercive moral paternalism.

Therefore (from 3 and 6),

7. Coercive moral paternalism is never justified.

Contrary to the claim of one commentator[39], Raz does provide a rationale for the first premise. The rationale is an implication of the explanation of what is wrong with

---

[38] Raz, 'Liberty and Trust', p. 127.

[39] Christopher Wolfe, 'Being Worthy of Trust: a Response to Joseph Raz', in Robert P. George (ed.), *Natural Law, Liberalism, and Morality*. Oxford, Clarendon Press, 1996, pp. 138-9.

coercion. As was seen in chapter two, according to Raz, coercion, along with manipulation, invades independence. Independence is a condition of autonomy but its importance transcends this role. Independence is also connected to the whole-heartedness aspect of well-being, conceived as the whole-hearted pursuit of worthwhile activities. Whole-hearted pursuit of goals is incompatible with being coerced or manipulated by others.[40] Coercive paternalism therefore harms personal well-being. But coercion, says Raz, is sometimes justified. In paternalistic cases, 'coercion is normally justified only if used by friends or others whose good intentions are not in doubt.'[41] When this condition holds, the whole-heartedness aspect of well-being is not undermined. If a person does not doubt the good intentions of her coercer, then she may remain whole-heartedly committed to her goals even though coerced into them.[42] This, says Raz, reflects the nature of trust. When trust occurs, the normal prohibition on coercion may be relaxed.[43] A person who trusts his coercer is more likely to engage in the activity willingly and whole-heartedly. Thus the trust requirement lessens the bad effect on well-being that coercive paternalism has. I shall accept this and the first premise for the sake of argument.

The second premise of the argument, that in order to reasonably trust the state one needs full citizenship, is true. Full citizenship, by definition, is enjoyed by those who feel, not unreasonably, that the state regards them as intrinsically valuable when it makes its decisions. The premise, then, is false only if it is possible to reasonably trust the state even when one thinks that it is not treating one as intrinsically valuable, that in its eyes one does not matter in one's own right. But it is difficult to conceive how anyone could trust the state if that were true. Some people may be deluded into believing that the government takes their interests into account when in fact it doesn't,

---

[40] Raz, 'Liberty and Trust', p. 121. I questioned this claim above in text at note 21.
[41] Raz, 'Liberty and Trust', p. 122.
[42] Raz does not explicitly state this step of the argument, but it seems to me the most plausible way to explicate it.
[43] Raz, 'Liberty and Trust', p. 122.

and may thereby claim to trust the government. But, as Raz says, such trust would not be reasonable, because due to their delusion they do not have reasonable grounds for trusting the government.[44]

Even though the first two premises, and therefore the initial conclusion (3) are plausible, the next major premise (4, and by extension, 6) may be questioned. That premise is the claim that coercive moral paternalism denies its targets full citizenship. The argument here is that by preventing me from following my way of life, the state denies the validity of my fundamental beliefs concerning how to live my life which thus makes it impossible for me to think that the government is seriously taking my interests into account when making decisions. Of course, as Wolfe points out, coercive moral paternalism may be motivated by concern for the well-being of the subject of the coercion. The government may well view them as of intrinsic worth in their own right.[45] But the crucial question, on Raz's argument, is whether the subject of paternalism will see it this way, and would not be unreasonable to do so. Wolfe fails to discuss this and Raz provides grounds for thinking the subject would not feel the same way as the paternalist. He argues that 'it is reasonable not to trust someone who condemns the fundamentals of one's life as base and worthless.'[46] Even if the government is acting out of the best intentions,

> it is bound to seem to them as failing through moral blindness to take them and their well-being seriously, lacking the ability to give their interests the weight they deserve. The people who feel like that may be wrong. Their beliefs and way of life may indeed be worthless. But it is reasonable for them to deny that and to feel that the government has forfeited their trust.[47]

It should be noted that this stage of Raz's overall argument contains a subsidiary argument against coercive moral paternalism. The proposition being considered is

---

[44] Raz, 'Liberty and Trust', pp. 124-5.
[45] Wolfe, 'Being Worthy of Trust', pp.144, 146.
[46] Raz, 'Liberty and Trust', p. 127.
[47] Raz, 'Liberty and Trust', p. 127.

that coercive moral paternalism denies full citizenship. This proposition is part of the larger argument which says that full citizenship is needed for there to be reasonable trust and that paternalism is unjustified unless there is reasonable trust. But the proposition that coercive moral paternalism denies full citizenship is itself an argument against that type of paternalism, apart from its connection with the wider trust-based argument. Raz argues that everyone in a political community should be full citizens. If some people are not, if there are estranged groups who do not identify with the state, that society will be unstable.[48] Moreover, since humans are political animals, says Raz, 'feeling part of a larger community, and being able to identify oneself as a member of such communities is an essential ingredient in people's well-being.'[49] If the government has a duty to protect and promote the well-being of people then it should 'make sure that they all enjoy full citizenship.'[50] The government must refrain from coercive moral paternalism since such paternalism denies people full citizenship, thereby undermining their well-being.

An objection to premise 4 can be dismissed. It may be presented in the form of a dilemma. The premise claims that coercive moral paternalism denies full citizenship. If this is because all coercion, whether paternalistic or not, denies full citizenship, then since some coercion is justified, then why cannot denying full citizenship on paternalistic grounds also be justified? If on the other hand, not all coercion denies full citizenship, then what is special about coercive moral paternalism that means that it does deny full citizenship? With regard to the first part of the dilemma, Wolfe claims that coercion for non-paternalistic ends also denies full citizenship and this may sometimes be justified. Denying full citizenship is justified, for instance, where people are coercively prevented from harming others or violating their rights. Raz, argues Wolfe, has to show why it is appropriate for the government to deny full

---

[48] Raz, 'Liberty and Trust', p. 125.
[49] Raz, 'Liberty and Trust', p. 125.
[50] Raz, 'Liberty and Trust', p. 125.

citizenship to those who have harmed others but not appropriate to do so in cases only involving self-regarding actions.[51] Raz does admit that there can be considerations which override the government's duty to grant everyone full citizenship.[52] But paternalistic considerations seem ill-suited for this role. It seems at least possible for the government to rightly behave in a way that would undermine a person's belief that the government recognises him as someone who matters in his own right, in order to protect others. But to do so for that person's own good seems less plausible.[53] It would mean that the government is behaving in a way that makes a person believe that it does not consider him intrinsically important, but that that behaviour is nevertheless for his own good.

The second part of the dilemma assumes that coercion need not deny full citizenship. When the state punishes a person for harming others, it may still nevertheless view that person's interests as intrinsically valuable. If this is so, Wolfe asks, then why do coercive threats of punishment against certain self-regarding actions deny full citizenship?[54] The reply is that they do so because those coercive threats deny the validity of the person's fundamental beliefs. Coercion in order to protect others does not do the same. The government's intention in protecting others is not to deny the validity of the beliefs of those who wish to harm others. The intention is simply to protect others. Denial of the harmer's beliefs may be the effect of the intervention, in the sense that he may believe that such harm is acceptable. But this need not be true (someone harming others may agree that it is wrong) and even if it is, the denial would be the effect though not the intention of the intervention. And if it is not the government's intention then the person can less reasonably believe that the government, in its actions, is not taking his interests seriously. Paternalistic coercion, on the other hand, has as its intention the denial of the beliefs of the subject and their

---

[51] Wolfe, 'Being Worthy of Trust', p. 144.
[52] Raz, 'Liberty and Trust', p. 126.
[53] Raz suggests the same in 'Liberty and Trust', p. 126.
[54] Wolfe, 'Being Worthy of Trust', pp. 144, 146.

replacement with the beliefs of the paternalist. Since this is its intention, it is less likely to seem that the government is treating the people involved as intrinsically valuable.

So we may reject Wolfe's objection. But we still need to examine the thesis that coercive moral paternalism denies full citizenship. This claim is central to both the subsidiary and the wider argument.

It is not true that full citizenship is always denied by a government that pursues coercive moral paternalistic policies. Such policies will deny the validity of the propositions which underpin my way of life (if I am the target of the paternalism), and may condemn these fundamental beliefs as base and worthless. But under certain conditions it would nevertheless be unreasonable of me not to think that the government is treating me as someone who matters in my own right. If I know that the intentions of the government are beneficent, that it sincerely believes that my way of life is misguided and mistaken, that it believes the mistake is of such magnitude that it warrants coercion where no other measures are possible, if I know that the government has thought long and deeply about the issue when arriving at its decision, considering all the arguments or all that are available on either side of the issue, if it takes adequate steps to explain the rationale behind the paternalism, to explain how it thinks I am mistaken, and if it does its utmost to ensure that the coercion is as localised as possible so that it does not interfere with my other options, then it does not seem unreasonable for me to believe that the government is seriously taking my interests into account in its decision making. When such steps are taken by the government it would be difficult for me to think that it is not acting in order to further my interests. Of course, the state and myself will still disagree over what is in my interests and it could be the case that I am right and the state is wrong about what is in my interests. Nevertheless, it would be reasonable for me to think that the state is

acting in what it sees as my interests. Thus coercive moral paternalism could be compatible with full citizenship.

This means the subsidiary argument does not succeed. Moreover since, in the circumstances described, it would be reasonable of me to think that the government is taking my interests seriously then it would also be unreasonable not to trust the government. Therefore coercive moral paternalism could satisfy the second condition for justified coercive paternalism. This means that the wider argument does not succeed. Premise 4 is false and hence so is premise 6. Therefore the conclusion (7), that coercive moral paternalism is unjustified, is not established by the argument.

One suggestion Raz makes may be able to make the argument avoid the objection. A further requirement for full citizenship, he claims, is that 'one must be able to profess one's basic beliefs, and conduct one's life in accordance with them and with one's deepest feelings without fear of criminal sanctions.'[55] With this additional requirement, coercive moral paternalism will always undermine full citizenship, since such paternalism prevents people from following their way of life.

However, while adding this requirement for full citizenship makes the argument sound, it is not clear what the justification for the requirement is. It does not follow from Raz's characterisation of full citizenship as believing (not unreasonably) that the state recognises one's interests as intrinsically important. As was just seen above, it is possible for the state to prevent people from following their way of life while at the same time making clear to them that it regards their interests seriously, so long as it takes certain steps to make this clear. Nor, unlike the initial characterisation of full citizenship, does the additional requirement follow from the trust requirement. The connection between trust and the initial characterisation of full citizenship is straightforward, since it would not be reasonable for one to trust the state if one feels

[55] Raz, 'Liberty and Trust', p. 126.

that it does not recognise one as intrinsically important. But it is not so obvious what the connection is between trust and the added requirement that full citizenship needs a person to be able to follow his way of life. In the situation described previously, where the state makes its intentions clear, sincerely believes that people are making a serious mistake, that there are no other means available, etc., it is not unreasonable to trust the government even if it is preventing one from being able to profess one's basic beliefs and conduct one's life in accordance with them. If I know that the state has honest intentions and believes that no other course is possible, and if it attempts to explain the rationale for its behaviour, etc., then it is reasonable for me to trust the government. In the absence of a justification for the added requirement, it should be rejected, and the conclusion of the argument therefore remains unsupported.

3.5 Conclusion

Raz's argument may therefore be rejected. Even if coercive paternalism is justified only if the paternalist is reasonably trusted by the coerced, it is not true, given certain circumstances, that it is always unreasonable to trust a state that is acting paternalistically. The trust requirement will rule out some paternalism. But it does not provide the absolute prohibition that Raz intends it to. Thus we are still left with no defence against the conclusion that perfectionism permits and even requires that the state coercively prevent people from pursuing worthless activities.

Chapter Four

The Endorsement Constraint

4.1 Introduction

In previous chapters, a number of perfectionist strategies for resisting the paternalistic implications of perfectionist political morality have been examined and criticised. The value of autonomy is unable to rule out all serious paternalism, and whilst there may be principled limits to coercion, these limits do not prohibit all such paternalism. Another strategy of perfectionism appeals to a particular thesis about well-being. This is that in order for an aspect of a person's life to advance her well-being, it is necessary that she believe in the value of that aspect. A person's life cannot be improved, so the claim goes, against her conviction that it is not. This proposition may be termed the *endorsement constraint* since it claims that a person's endorsement is needed for any activity to improve her well-being.

The endorsement constraint is given most attention by Ronald Dworkin, and much of the chapter will focus on his arguments. Another adherent of the endorsement constraint is Will Kymlicka, who maintains that 'no life goes better by being led from the outside according to values the person doesn't endorse.'[1] Joseph Raz also suggests something like the endorsement constraint when he states that we cannot benefit people by forcing them into goals they do not accept because goals, to be goals, must (with some exceptions) be adopted or endorsed.[2] Neither Dworkin nor Kymlicka are

---

[1] Will Kymlicka, *Liberalism, Community, and Culture*, Clarendon Press, Oxford, 1989, p. 12. See also his *Contemporary Political Philosophy*, p. 203.

[2] Joseph Raz, *The Morality of Freedom*, Oxford, Clarendon Press, 1986, pp. 291-2. Raz does not develop the idea of endorsement at any length.

normally considered perfectionists, since they advocate some form of state neutrality.[3] However, they do not qualify as neutralists (at least in some of their works) in the sense in which it is being defined in this thesis, which is *justificatory* neutrality. As stated in chapter one, neutrality in this sense refers to abstinence in using conceptions of the good in the justification of state decisions. This is to be distinguished from neutrality of aim, which requires that the state not aim to encourage or discourage any conception of the good. Both Dworkin and Kymlicka support neutrality of aim but neither are justificatory neutralists. They believe that the state should not intentionally help any conception of the good, but their justification for this position is that it is supported by (what they believe is) the correct understanding of the nature and requirements of the good life. The endorsement constraint is one part of this supposedly correct understanding, and is thought by Dworkin and Kymlicka to provide support for neutrality of aim.

That the theories of both Dworkin and Kymlicka have perfectionist foundations is clear from their texts. Dworkin's aim in the lengthy article in which he argues for the endorsement constraint is to show how liberal politics can be seen to be supported by a plausible account of the good life. Responding to the objection that liberalism pays too little attention to the good, he tries 'to construct what we might call a *liberal ethics* - instincts and convictions about the character and ends of human life that seem particularly congenial to liberal political principles'.[4] Dworkin rejects justificatory neutrality, or what he calls 'the discontinuity strategy' (so named because it attempts to

---

[3] Dworkin's 'Liberalism' in his *A Matter of Principle*, Cambridge, Massachusetts, Harvard University Press, 1985 is one of the most quoted statements of neutrality. He also defends neutrality of a sort in 'Foundations of Liberal Equality', in G. B. Peterson (ed), *The Tanner Lectures on Human Values*. Vol. 11, Salt Lake City, University of Utah Press, 1990, pp. 110-8. The 'Foundations' article is reprinted in a shorter version in Dworkin's *Sovereign Virtue: The Theory and Practice of Equality,* Cambridge, Massachusetts, Harvard University Press, 2000, chapter 6. The arguments I examine remain unchanged in the recent book. Kymlicka defends neutrality in *Liberalism, Community, and Culture*, pp. 76-85 and in *Contemporary Political Philosophy: An Introduction*, pp. 199-205.

[4] Dworkin, 'Foundations of Liberal Equality', p. 20.

disconnect and set aside conceptions of the good as foundations for political morality), as lacking theoretical justification.[5] Instead, writes Dworkin,

> We need an account of what people's critical interests are that will show why people who accept that account and care about their own and other people's critical well-being will be led naturally toward some form of liberal polity and practice.[6]

While not as explicit as Dworkin, Kymlicka follows a similar strategy. He claims that people's 'essential interest is in leading a good life, in having those things that a good life contains'[7] and that a full account of this essential interest 'forms the basis of liberal political theory.'[8] It does so by containing a number of theses about the good life, such as that our interest is in leading what is *truly* a good life rather than simply what we *believe* is a good life. This proposition, Kymlicka believes, helps support certain civil and personal liberties such as freedom of expression.[9] Another proposition about the good life is the endorsement constraint, which according to Kymlicka, 'will justify the traditional liberal prohibitions on coercive paternalism.'[10] Hence Kymlicka, like Dworkin, rather than setting aside conceptions of the good in the foundation of political morality as justificatory neutrality requires, relies upon a conception of the good (albeit an abstract one) to justify political principles.

But once these perfectionist foundations are admitted, there is motivation for a wide range of paternalistic practices. If the state has a role in helping people to lead good lives and if people can be mistaken about which lives are good and which are not, which Dworkin and Kymlicka along with other perfectionists seem to admit, then

---

[5] See Dworkin, 'Foundations of Liberal Equality', pp. 22-35 for Dworkin's rejection of arguments for justificatory neutrality.

[6] Dworkin, 'Foundations of Liberal Equality', p. 46. What Dworkin means by 'critical' well-being shall be explained below.

[7] Kymlicka, *Liberalism*, p. 10.

[8] Kymlicka, *Liberalism*, p. 13.

[9] Kymlicka, *Liberalism*, p. 13.

[10] Kymlicka, *Liberalism*, p. 18.

there is a rationale for forcing people into good lives and away from bad. Dworkin tries to deflect this difficulty from the start. It assumes, he says, 'that we have some standard of what a good life is that transcends the question of what circumstances are appropriate for people deciding how to live.'[11] If we did, he argues, then we could shape those circumstances to ensure people achieve that standard. But on the correct understanding of the nature of the good life, according to Dworkin, living well is simply responding appropriately to circumstances rightly judged. This notion, Dworkin seems to claim, rules out any standards by which we can judge circumstances as being more or less likely to produce good lives. I shall call this the absence of transcendent standards argument, or ATS for short.

If ATS is correct then it makes any appeal to the endorsement constraint unnecessary. What motivates that constraint is the thought that the state may be able to create circumstances in which people could be forced into better lives The endorsement constraint is supposed to show that even though the state could do this, it ought not to since, in the absence of endorsement, well-being will not be advanced. Since Dworkin relies heavily upon the endorsement constraint, it seems that he does not put much weight on ATS. Moreover, Dworkin seems to accept that there are standards of the good life compatible with the model of living he advocates. He claims that someone can accept that model and think that religious devotion is part of living well or that homosexuality blights one's life.[12] This implies a rejection of ATS. Dworkin argues that the good life consists of a skillful performance in the face of the challenge of living and that our interests lie in meeting that challenge well.[13] That implies that it is possible to meet the challenge badly.[14] If so, then it is possible that the state could block options that are more likely to result in meeting the challenge of life badly, such

---

[11] Dworkin, 'Foundations of Liberal Equality', p. 84.
[12] Dworkin, 'Foundations of Liberal Equality', p. 78.
[13] Dworkin, 'Foundations of Liberal Equality', p. 57.
[14] Dworkin seems to admit the possibility on p. 58 of 'Foundations of Liberal Equality'.

as being a couch potato. So Dworkin's own understanding of his model of living rejects the central claim of ATS and that argument should be rejected.

The endorsement constraint, Dworkin and Kymlicka hope, provides the explanation for how their theories may avoid paternalistic implications. Neutrality (in the justificatory sense) does not require such a constraint since the motivation does not exist in the first place. Perfectionist considerations are set aside in the justification of state action. However, as I try to show in this chapter, the endorsement constraint is unable to do the work that Dworkin and Kymlicka hope. After clarifying the notion of endorsement, I will defend it from some objections. The endorsement constraint has some intuitive support, but in section 4.4 I try to show that a weaker alternative view is more plausible. I examine and reject some arguments for the stronger endorsement constraint, thus concluding that the weaker view should be accepted. This weaker view fails to rule out as much paternalism as the endorsement constraint.

4.2 Endorsement

It will be helpful first to explain the notion of endorsement in more detail. Dworkin distinguishes two questions that may be asked when evaluating someone's well-being:

> We can ask, first, how far his life includes whatever experiences or relationships or events or achievements we count as components of a good or decent life. We can ask, second, how far he recognizes whatever components of the good life his own life contains, whether he sought them, regarded them as valuable, in short endorsed them as serving his critical interests.[15]

Dworkin then asks what view we should take of the relationship between the components of a person's life and his attitude towards those components. He claims that there are two possible views:

> The *additive* view holds that we can judge his life a good or bad one without consulting his opinions of its value. If his life has the components of a good

[15] Dworkin, 'Foundations of Liberal Equality', p. 50.

> life, then it is good for that reason. If he endorses those components, then this increases the goodness of his life; it is frosting on the cake. But if he does not, the ethical value of the components remains. He may have a very good life in virtue of experiences and achievements he does not endorse, though not so good a life, perhaps, as if he had endorsed them.
>
> The *constitutive* view, on the other hand, argues that no component may even so much as contribute to the value of a person's life without his endorsement. So if a misanthrope is much loved but disdains the love of others as worthless, his life is not more valuable for their affection.[16]

The two views have different implications with regard to paternalism. Consider a public policy, motivated on paternalistic grounds that makes it impossible to lead a life entirely devoted to surfing.[17] Suppose someone refrains from being a full-time surfer out of fear of punishment and fails to endorse the life he leads as a result. Upon both the additive and constitutive views, his life has gone worse in one respect due to the lack of endorsement. However, on the additive view, his life may be better if the critical value of his life is improved by not surfing and if this improvement is enough to outweigh the loss of endorsement. According to the constitutive view however, his life has not been improved at all by the paternalistic policy since endorsement is a necessary condition for a component to contribute to the critical value of someone's life. The constitutive view will make such a paternalistic policy pointless.

It is important to note what the endorsement constraint applies to. Dworkin claims that it applies to a person's critical interests. These are distinguished from volitional interests. One's critical well-being is improved when one has or achieves what one should want. People may fail, therefore, to see what is in their critical interests. By contrast, people also have, according to Dworkin, volitional interests which are furthered when people have or achieve what in fact they do want.[18] The endorsement constraint applies to critical interests only. For a person's critical well-being to be improved, according to EC, she must endorse the critical value of the goal or activity.

---

[16] Dworkin, 'Foundations of Liberal Equality', p. 50. Italics in original.

[17] Workfare schemes, where people receive state welfare payments only on the condition that they do community work, may be examples of this.

[18] Dworkin, 'Foundations of Liberal Equality', p. 43.

The distinction is important for paternalism. The two types of interests correspond to two forms of paternalism. Critical paternalism restricts somebody's opportunity to choose in order give him what he ought to have, even if he does not want it. It is justified upon the constitutive view only if the person endorses the value of the goal or activity. Volitional paternalism, on the other hand, denies his choice in order to help him achieve what he wants to achieve and may be justified without endorsement. Volitional paternalism is means-ends paternalism; it forces a person into the means required for an end they acknowledge. Seatbelt laws are an example. As Dworkin construes such laws, 'the state makes people wear seat belts in order to keep them from harm it assumes they already want to avoid.'[19] Paternalism here may be justified even if the paternalised does not endorse the value of the means. Critical paternalism, by contrast, is paternalism of ultimate ends.[20] It tries to force a person into more worthwhile ends than he currently holds. The constitutive view says that those ends must be endorsed for them to improve the person's well-being.

Another exemption to the endorsement constraint should be noted in order to ward off some simple objections. If I knock a glass of poisoned wine from the hand of a person about to drink it, I benefit the person even if he does not endorse the value of being denied a drink. Similarly I benefit a person by moving her out of the way of an oncoming vehicle. People are benefited when their biological needs are met, such as when they are fed, kept warm, and relieved from pain. These things improve people's well-being regardless of their endorsement. It may be possible to justify some actions that satisfy people's biological needs on grounds of their volitional interests since people usually want those needs satisfied and they may be instrumental to other desires. But this would not be a fully adequate explanation. When biological needs are met, people are better off in at least one respect even if they do not want the needs

---

[19] Dworkin, 'Foundations of Liberal Equality', p. 77. Such laws are not always means-end paternalism since someone may not have the end the example assumes or he may be willing to weigh that end against others such as unrestricted comfort while driving.

[20] I borrow the phrase from Wilkinson, 'Dworkin on Paternalism and Well-being', p. 433.

met or they are not connected to any other desires. Hence the endorsement constraint, if it applies at all, applies only to a person's ultimate ends other than her biological needs.

4.3 Three Objections

The endorsement constraint is able to withstand three objections that have been made against it. Thomas Hurka claims that it is implausible to think that if I do not believe my activities are good, then they are valueless. Imagine, he says, an artist who creates brilliant works but thinks they are worthless, perhaps due to the extremely high standards she sets herself. According to the endorsement constraint, her artistic activity has no value. This, says Hurka, is unacceptable. We may think her activity is more valuable if she believes it has value, but we cannot think that her activity in creating brilliant works has no value at all.[21]

This objection may be overcome by drawing a distinction between value as such and value for a person. Someone who is kidnapped and whose various body parts are used in order to save lives clearly has value in some sense. But it is value to others. His life does not go better due to the compulsory operation.[22] Similarly with Hurka's artist. Her works may have value to others because of the pleasure they get from viewing them. Or they may have value in themselves regardless of how they affect people. But they may not be valuable for the artist, that is, contribute to her well-being, if she thinks them worthless. Since the endorsement constraint is a proposition about value for a person, about well-being, then that proposition may still be true even if we agree with Hurka that the artist's creative activity does have value; it may be

---

[21] Thomas Hurka, 'Indirect Perfectionism: Kymlicka on Liberal Neutrality', *Journal of Political Philosophy*, vol. 3, no. 1, 1995, p. 42.

[22] This example is from Wilkinson, 'Dworkin on Paternalism and Well-being', p. 438.

valuable independently of whether it contributes to her well-being. Hurka's objection therefore may be put aside.

Another objection is pressed by Robert George. He also denies that a component of the good life can have no value if it is not endorsed. Basic human goods, he claims, are intrinsic aspects of well-being. They provide reasons for choice and action, reasons that are not reducible to satisfactions. These reasons remain, according to George, even when people believe a basic good to be valueless. Something can be valuable for a person even if he does not value it. Hence, George concludes, endorsement is not necessary for something to contribute to well-being.[23] But the constitutive view can account for basic human goods being of intrinsic value for a person. They are valuable for a person even if he does not value them because his well-being is advanced in their presence, so long as they are endorsed, but undermined in their absence. On the constitutive view, basic goods still provide reasons for choice and action. They are good for a person even if he does not want them. Basic goods would contribute to his well-being if he would recognise and engage in them. The endorsement constraint says that these goods cannot be fully realised, that there will be no advancement of a person's well-being, unless he comes to realise their goodness and thereby endorse them, but they remain of value even if he does not recognise them.

An objection is developed along similar lines by Steven Wall. The endorsement constraint implies that if one does not believe that some worthwhile activity makes one's life better, then it does not. This is implausible, says Wall, because it rules out the possibility of the person being mistaken about the issue. The endorsement constraint means that if a person believes the activity does not advance his well-being, then he cannot be wrong; it is not possible that the activity advances his well-being even though he thinks it does not. But, argues Wall, we believe that it is possible for a

---

[23] Robert George, *Making Men Moral*, Clarendon Press, Oxford, 1993, pp. 105-6.

person to be mistaken about that question. Asking "Did this pursuit add value to my life?" is an intelligible question. The endorsement constraint makes it unintelligible since answering it in the negative results in a self-fulfilling negative answer.[24]

The constitutive view is able to reply to this objection. Upon that view, mistakes are still possible. There are two types of mistakes that could be made. Firstly, one can mistakenly think that *x* is valuable when in fact it is not. The endorsement constraint expresses what is supposed to be a necessary but not sufficient condition for well-being. It does not claim that believing an activity to be valuable makes it so. Instead, it claims that of those activities that are truly valuable, and not those that are worthless, only those that are endorsed contribute to well-being. Endorsing worthless activities does not contribute to well-being. Hence someone that does so has made a mistake in believing that the activity is valuable when it is not.

The second type of mistake, when a person thinks that *x* does not advance her well-being when it does, requires a less straightforward explanation. The endorsement constraint does deny that a person's life can be improved against her conviction that it is not. It is this denial which gives the constraint its anti-paternalistic implications. But that does not mean that a person cannot be mistaken in thinking that some experience or activity would not advance her well-being. The endorsement constraint does not deny that she may mistakenly think that *x* is not valuable when in fact it is. If *x* is valuable then it remains valuable even when a person thinks it is not. But making it part of her life will not advance her well-being without her endorsement. The mistake she makes is in not realising that something is in fact valuable. If she had, then her well-being would have been advanced. By not realising *x*'s value, she loses the chance of engaging in something that would advance her well-being. But *x* nevertheless remains of value and she is mistaken in thinking it not. As Dworkin

---

[24] Steven Wall, *Liberalism, Perfectionism, and Restraint*, Cambridge University Press, Cambridge, 1998, p. 194.

says, a person 'might be wrong in not recognizing and endorsing some feature of his life that, had he recognized it, would have made his life better.'[25]

The mistake is more significant if the endorsement constraint is true. If endorsement is not required, if it just adds to the value of an activity without being necessary for it, then a person who mistakenly withholds her endorsement may still have her well-being advanced by the presence of the valuable activity. But under the endorsement constraint, the consequences of so withholding are more severe. Her well-being is more severely undermined. Hence Wall's objection may be turned against him in favour of the endorsement constraint. It seems plausible to think that when people are mistaken about the value of their lives, this mistake must have significant consequences for them. Without the endorsement constraint, it could be the case the someone who makes the mistake of thinking x is not of value when it is could nevertheless live just as good a life or only a slightly worse one. The endorsement constraint explains why making the mistake can have drastic effects; by not realising what is of value, a person loses out entirely on the contribution that that value could have made to her well-being.

The endorsement constraint, therefore, is able to resist a number of objections. But this does not make it true. In the next section I examine a number of considerations which seem to provide support for the constitutive view. I then argue that there is an alternative view which is more plausible and able to account for the considerations that also provide support for the constitutive view. But this view provides less protection against paternalism. I then examine Dworkin's two arguments for the stronger, constitutive view.

4.4 The Intuitive Argument

---

[25] Dworkin, 'Foundations of Liberal Equality', p. 50.

A number of straightforward intuitions seem to support the endorsement constraint. For example, as Dworkin says, a misanthrope who is much loved by others but who disdains that love does not have a more valuable life due to their affection.[26] It also seems implausible to think that someone's life is improved if she is forced into religious observance that she thinks worthless, even if it is true that religion is part of the good life.[27] Even if (as some falsely believe) a homosexual blights his life by a failure to understand the point of sexual love, it is unlikely that he overcomes that defect if he abstains, against his convictions and only out of fear of punishment.[28] A life in politics without genuine satisfaction or self-approval can hardly be thought better than a religious one endorsed as valuable, even if we think religion is worthless.[29] The constitutive view captures and explains these intuitions. In all of them, valuable projects and activities seem not to contribute to the well-being of the person involved. And in all of them, there is an absence of endorsement.

There are other types of judgements about well-being that are similarly captured by the constitutive view. The above examples concern people being made to lead lives contrary to their profound ethical convictions. Consider instead cases where people are made to live against unreflective opinions and powerful passions. Despite one commentator's claim, it seems likely that, at least in some cases, a person can lead a worse life when made to live in ways contrary to his unreflective opinions or powerful passions.[30] Someone who never reflects upon the components of her life but

---

[26] Dworkin, 'Foundations of Liberal Equality', p. 50. See also pp. 51, 77.

[27] Dworkin, 'Foundations of Liberal Equality', pp. 51, 78. Kymlicka also relies heavily upon the case of coerced religious worship to support the endorsement constraint. See *Contemporary Political Philosophy*, p. 204 and *Liberalism, Community, and Culture*, p. 12

[28] Dworkin, 'Foundations of Liberal Equality', p. 78.

[29] Dworkin, 'Foundations of Liberal Equality', pp. 79-80, 81-2. Dworkin makes a stronger claim: that unendorsed politics is *worse* than endorsed religion, even if religion is worthless. As Wilkinson shows, such a ranking is inconsistent with the constitutive view. That view says that since religion is worthless, endorsing it adds nothing and though politics is valuable, without endorsement there is no value. Hence both lives are equally worthless. See Wilkinson, 'Dworkin on Paternalism and Well-being', pp. 436-7.

[30] Christopher Wolfe, 'Liberalism and Paternalism: A Critique of Ronald Dworkin', *The Review of Politics*, vol. 56, no. 4, 1994, pp. 627-8.

nevertheless very much enjoys it would surely be worse off if made to do something she doesn't enjoy. Consider also a person brought up in a religious background and who never seriously reflects upon his religious beliefs. He may still be committed to them in such a way that without them his life would be significantly worse. A similar point may be made concerning powerful passions. A homosexual's beliefs in the goodness of his life may have been adopted due to the passions that impel him. A Marxist may also be impelled by powerful passions concerning the injustice of capitalism that lead him to adopt his Marxist opinions. Both the homosexual's and the Marxist's lives in these cases would go worse if made to act contrary to their passions, in a way similar to the worse life led by someone living against the grain of his most profound ethical convictions. The constitutive view embraces all of these judgements. They all involve a withholding of endorsement. Endorsement involves a positive attitude towards one's life or aspects of it and such an attitude need not be negated solely by being made to live against consciously held ethical beliefs.

Consider a more general thesis about well-being. Dworkin emphasises the importance of leading a life without a sense of resentment towards it. He claims that 'it is implausible to think that someone can lead a better life against the grain of his most profound ethical convictions than at peace with them.'[31] Similarly, he writes '[i]t seems preposterous that it could be in someone's interests, even in the critical sense, to lead a life he despises and thinks unworthy.'[32] These thoughts also support the endorsement constraint.

More support for the endorsement constraint may be drawn from Dworkin's text. The constitutive view solves the problem of satisfying two conditions concerning well-being which seem contradictory but both of which seem plausible. One condition is objectivity, which says that well-being depends on having features and components in

---

[31] Dworkin, 'Liberal Community', p. 486.
[32] Dworkin, 'Foundations of Liberal Equality', p. 76.

one's life that are truly valuable. The other condition is subjectivity, which says that well-being is a function of getting and achieving whatever one wants. Both conditions seem plausible. We often think of well-being as objective; 'a particular life cannot be good for me just because I think it is, and I can make a mistake in thinking a particular life good'.[33] But at the same time it seems implausible that people can be made better off by giving them things they do not want; '[i]t seems preposterous that it could be in someone's interests ... to lead a life he despises and thinks unworthy'.[34] Well-being seems subjective after all, and having a good life must be a matter of thinking one's life good. But again, it seems impossible to think one's life good unless its goodness does not depend on one thinking it so.[35]

The constitutive view solves this dilemma. The endorsement constraint captures the plausibility of the notion of subjectivity. It rules out the possibility that someone's life can be improved if it contains features they despise or think worthless. The value of the good life, according to the constitutive view, does depend to some extent on thinking it valuable. But the constitutive view is also suitably objective. According to it, one's life isn't good just because one thinks it is, and it is possible to be mistaken about the value of one's life.

Some other views of well-being are unable to solve this dilemma. An account of well-being that takes it to consist purely of preference satisfaction fails the test of objectivity since it says certain objects are of value because one wants them rather than being wanted because objectively valuable. A view that holds well-being to consist of the presence or absence of a list of goods such as self-reliance, close family ties, and fulfillment in one's employment fails the test of subjectivity since it holds

---

[33] Dworkin, 'Foundations of Liberal Equality', pp. 75-6.
[34]Dworkin, 'Foundations of Liberal Equality', p. 76.
[35] Dworkin, 'Foundations of Liberal Equality', p. 76.

that these goods further the well-being of those who think them worthless.[36] The additive view also fails the subjectivity test, because even though it sees endorsement as valuable and may give great weighting to it, it nevertheless holds that some components of a person's life may contribute to well-being without endorsement and even if the person thinks them worthless. The constitutive view, therefore, has an important advantage over these alternative views.

The above points may be summarised in the following way. The constitutive view is able to satisfy three related criteria for a plausible view of well-being. It can account for judgements we make in a range of particular cases, it captures the importance of the absence of resentment in a good life, and it solves the subjective-objective dilemma.

But there is an alternative view which also satisfies these three criteria. The important aspect of this view is that it contains what may be termed a *non-resentment constraint*. Rather than a component of one's life having to be endorsed for it to have value, it may be thought necessary merely that it not be despised and thought worthless. That is, a necessary condition for a component to contribute to well-being is that it not be resented, rather than having to be positively endorsed. Such a non-resentment constraint (henceforth NRC) is different from the endorsement constraint (EC). The former allows the possibility that a person's well-being may be improved by the presence of components that she does not endorse so long as she does not despise them. EC, on the other hand, rules out the possibility of someone's life being improved even where there is no resentment, just a neutral attitude towards the component, since such a neutral attitude is not endorsement.

---

[36] For more discussion of these views, see Derek Parfit, *Reasons and Persons*. Oxford, Oxford University Press, 1984, Appendix I; and James Griffin, *Well-Being: Its Meaning, Measurement, and Moral Importance*. Oxford, Oxford University Press, 1986.

NRC can account for the range of judgements that seem to support EC. All of them involved people disapproving of the lives they live. The badness of these lives could be explained by the element of resentment in them rather than the lack of endorsement. NRC also of course satisfies the second criterion which is just that a good life ought not to contain resentment. NRC is in fact *all* that is implied by the second criterion. EC makes a stronger claim than that suggested by the criterion. The criterion is based on the suggestion that a component of a person's life cannot be valuable if that person dislikes that component. The constitutive view, however, makes the stronger claim that to be of value, a component must be endorsed as valuable by the person. Even if it is true that in order to have a good life one must not despise its components, it may not be true that the good life requires positive endorsement. A worthwhile life may contain components that are not disliked nor thought of as particularly valuable. Not having components that one despises does not mean that one believes the components of one's life are valuable. Finally, NRC also satisfies the third criterion. It is objective because it accepts that one's life cannot be good just because one thinks it is, and that it is possible to be mistaken about the value of one's life. But it also retains an important element of subjectivity. It rules out the possibility that someone's life can be improved if it contains features they despise or think worthless. The value of the good life, on this view, does depend to some extent on one's attitude towards it.

Both the constitutive view and a similar view containing NRC (rather than EC) satisfy the three criteria. Another consideration weighs in favour of NRC. The constitutive view implies that it is impossible for there to be an improvement in one's life due to the presence of components that one neither endorses nor despises. But this is implausible. Suppose a person has an occupation that she enjoys and that stimulates her mind and exercises her body. Nevertheless, when she thinks about it she doesn't regard that occupation as particularly valuable. But nor does she think it valueless or dislike it in any way. It seems plausible that she is better off through having that

occupation than if she had one she dislikes, other things being equal. But the constitutive view, because of EC, views both occupations as valueless. It is consistent with NRC, on the other hand, that her life is better in virtue of her occupation.[37]

A possible reply to this argument is to hold that the constitutive view be combined with a claim about the badness of living a life one despises. That is, there is nothing inconsistent in holding that endorsement is necessary for a valuable component to contribute to well-being *and* that someone's life is worse if she despises it than if she is neutral towards it, neither despising nor endorsing it. This would be so if the despised life would involve negative amounts of well-being whilst the neutral life, since it lacks endorsement, involves no amount of well-being. Adding the claim about the badness of living a life one dislikes allows us to say, consistent with the constitutive view, that a person's life where she is neutral towards its value, neither endorsing nor disliking it, is better than one she dislikes since the latter has negative value. This reply, however, does not succeed since it insists on what is false, namely that the neutral life contains no well-being. It is not just that the despised life is *worse than* one about which a person is neutral, although that is true. The example suggests that a life containing a worthwhile occupation is one of positive value, even when a person does not particularly value it. If she despises it, it loses value, perhaps all value and may even become of negative value. But with a neutral attitude there can still be positive value. EC, however, even if combined with a claim about the badness of living a life one dislikes, denies this.

So there is some reason to prefer NRC over EC. The former but not the latter allows us to say that well-being may be advanced by the presence of valuable activities just so long as these are not despised by the person. The two views have different implications for paternalism. NRC provides weaker protection against paternalism

---

[37] A similar objection is made by Richard Arneson in 'Human Flourishing Versus Desire Satisfaction', *Social Philosophy and Policy*, vol. 16, no. 1, 1999, section XIV.

than EC. EC makes paternalism self-defeating even when the activity that the paternalised is guided into is viewed indifferently by that person. She may believe it to be neither valuable nor worthless or she may hold no belief regarding its value. In all such circumstances there is no endorsement and hence the person's life is not improved. In such circumstances the condition set by NRC, on the other hand, is met and paternalism may be justified.

This difference may be seen more clearly by considering two types of cases, set out by Thomas Hurka, that show how paternalism is compatible with the constitutive view.[38] The first concerns efforts to force people into a single best activity. Such intervention may be justified, on the constitutive view, if the paternalised person endorses the value of the activity. For example, 'a child who is forced to practice music is very likely later to endorse the coercion by agreeing that it did, in fact, make his life better'.[39] People initially resistant to the activities may, once forced into them, come to believe in their value. Paternalism here may be justified.[40] The second type of paternalism that is compatible with the constitutive view involves forbidding a single worst activity while leaving a range of others open to a person. Even if the prohibited option would have been endorsed by the paternalised, she instead will choose another option and she may endorse it. If a range of other activities are left open then it is likely that one of them could be chosen and endorsed by the paternalised person. It seems unlikely that of a range of options open to a person, only one of them could be endorsed.

---

[38] See the arguments in Thomas Hurka, 'Indirect Perfectionism', pp. 44-8.

[39] Dworkin, 'Foundations of Liberal Equality', p. 79. Kymlicka also acknowledges that endorsement may justify paternalism. See *Contemporary Political Philosophy*, pp. 232-3.

[40] It may be but it may not be. The endorsement constraint sets a necessary but not sufficient condition for justified paternalism. Other conditions may have to be met before the paternalism is justified. Both Dworkin and Kymlicka state that the intervention must also be short-term (See Dworkin, 'Foundations of Liberal Equality', p. 78 and Kymlicka, *Contemporary Political Philosophy*, pp. 232-3.). Critics have claimed that such a condition is unfounded since if endorsement is important, it need not matter whether the intervention is short- or long-term (See Hurka, 'Indirect Perfectionism', p. 45 and Wolfe, 'Liberalism and Paternalism', p. 629). But this criticism amounts to forcing upon Dworkin and Kymlicka a claim that neither of them make; namely that endorsement is sufficient for justified paternalism.

Dworkin tries to show that the possibility of endorsed paternalism is a narrow one. He argues that for intervention to be justified, the endorsement must be genuine and that it is not genuine if the mechanisms used to secure it lessen a person's ability to consider the critical merits of the change in a reflective way. Most state paternalism fails the test of genuine endorsement since 'threats of criminal punishment corrupt rather than enhance critical judgement.'[41] Dworkin gives no argument for this assertion, and it creates problems for his argument that a child forced to take music lessons can later endorse the intervention. Since threats are most probably used to make the child take the lessons, the endorsement, upon Dworkin's argument, would not be genuine. Moreover there is reason to think the assertion is false. People can use their capacity for critical judgement to consider whether certain options ought to be prohibited, and they can do this whether or not those options actually are prohibited. Threats of punishment need not undermine this capacity, but might instead simply add a consideration, with the intention that it be a strong or overriding one, for or against a certain course of action. Threats assume the capacity for critical judgement, otherwise the threats would not be effective. For threats to succeed, the targets of those threats have to use their capacity for critical judgement to decide to act the way the threateners intend them to act.

However, we may remain sceptical about the possibility of endorsed paternalism. Hurka himself admits that the first form of paternalism is unlikely to be accompanied by endorsement.[42] But both types are more likely to be permitted by NRC than by EC. Consider the second type of case, involving the prohibition of a single activity while leaving other options open. Imagine that the state makes it difficult or impossible to live a couch potato lifestyle. A number of options for living a more active life remain open. Some people prevented from being couch potatoes may engage in one of these

---

[41] Dworkin, 'Liberal Community', p. 486.
[42] Hurka, 'Indirect Perfectionism', p. 46.

other options but not view them as valuable. They do not dislike the activities, but have a neutral attitude towards them. According to EC, they are no better off and paternalism is self-defeating. According to NRC, since there is no resentment,[43] their lives may be improved and paternalism could be justified. Hence this type of paternalism, involving prohibition of bad options, is more likely to be permitted by NRC than by EC. The same argument, that interventions will be permitted by NRC but not EC, extends to other cases of paternalism. As well as cases of lifestyle paternalism like that described above involving couch potatoes, it would apply to cases of aesthetic paternalism (prohibiting bad artistic expressions). Prohibiting some genre of worthless art may not result in people who would have engaged in that genre endorsing anything else instead, but they may not resent and view as worthless that which is left open to them.

NRC is also more likely to permit the stronger type of paternalism that forces a person into a single activity. The test that needs to be met for this type of paternalism to be justified is weaker than that set by EC. Rather than any positive endorsement, all that is required is that the person not disapprove of the forced activity. Kymlicka outlines the process by which endorsed paternalism may occur: 'One way to get people to pursue something for the right reasons is to get them to pursue it for the wrong reasons, and hope they will then see its true value.'[44] This may be a vain hope since people would probably resent the intervention and refuse to appreciate the value of the activity. But they may remain neutral towards its value. While disliking the intervention, they may not deny the value of the activity itself even if they do not endorse its value. Consider, for example, the case of a child forced to take music lessons. She may dislike being forced to do so and refuse to believe that taking music

---

[43] There may be resentment towards the act of paternalism itself. But NRC concerns people's attitudes towards the activities people are paternalised into, rather than the entire intervention. It is possible to endorse the value of the activities while resenting the paternalism, say because one believes it to be unjustified due to a belief in state neutrality.

[44] Kymlicka, *Contemporary Political Philosophy*, pp. 233.

lessons is a valuable activity. But she may not think that it is a bad activity, of negative worth. Whilst EC says that this intervention cannot improve her life, such improvement is possible under NRC. For this type of paternalism to be justified according to NRC, the state does not have to ensure that people come to believe in the worth of an activity. It just has to ensure that they do not come to believe that the activity is of negative worth.

Hence NRC is likely to allow a wider range of paternalistic practices than EC. But there seems little reason to think that NRC is further from the truth as a thesis about well-being than EC. NRC in fact seems more plausible than EC. We need to see therefore whether there is an argument strong enough to make us accept EC. Without such an argument, perfectionism can only rely upon NRC rather than the stronger EC to block paternalism and it (NRC) is likely to permit a wide range of paternalistic policies.

4.5 Performance and Endorsement

Dworkin has two major arguments for the constitutive view. Both attempt to show how that view is supported by the conception of the good life that he explains and defends throughout 'Foundations of Liberal Equality' (elements of which have already been presented in the discussion of ATS in 4.1 above). This conception, the challenge model of well-being, holds that 'a good life has the inherent value of a skillful performance.'[45] According to this model, living a life is a performance that demands skill. It is the most important challenge we face, and our critical interests consist in the achievements, events, and experiences that mean that we have met the challenge well.[46]

---

[45] Dworkin, 'Foundations of Liberal Equality', p. 57.
[46] Dworkin, 'Foundations of Liberal Equality', p. 57.

Dworkin's first argument claims that there is a connection between his conception of the good life as a skillful performance in meeting the challenge of life and endorsement. He argues that the challenge model presses towards the constitutive view in the following way:

> intention is part of performance: we do not give credit to a performer for some feature of his performance he was struggling to avoid, or would not recognize, even in retrospect, as good or desirable.[47]

According to Dworkin, the connection between the challenge view and the endorsement constraint means that someone who accepts the challenge model might think that certain activities are an essential part of how human beings should respond to their place in the universe, and therefore be part of living well, but he cannot think that those activities, undertaken involuntarily, have any ethical value. On the challenge model, 'it is performance that counts, not mere external result, and the right motive or sense is necessary to the right performance.'[48] Dworkin's argument in short is that the good life is one that is a skillful performance, such a performance requires intention, intention is the same as or requires endorsement, and therefore a good life requires endorsement.[49]

However, as T. M. Wilkinson points out, intention and endorsement are not the same and nor does the former require the latter. We may agree with Dworkin that a skillful performance requires intention. As he says, we do not give credit to a novice painter if his hand is moved by a master for every stroke (even if the novice endorses each move).[50] But there can be intention without endorsement. A person can intentionally perform an activity, and do so skillfully, without regarding that activity as valuable.

---

[47] Dworkin, 'Foundations of Liberal Equality', p. 77.
[48] Dworkin, 'Foundations of Liberal Equality', p. 78.
[49] This summary of Dworkin's argument is from Wilkinson, 'Dworkin on Paternalism and Well-being', p. 439.
[50] Dworkin, 'Foundations of Liberal Equality', p. 77.

An author can write stories he regards as potboilers produced simply to earn money.[51] An artist may intentionally produce highly accomplished works and yet, due to her high standards, she may think of those works as worthless. Clearly, there can be intentional skillful performance without any endorsement of the value of that performance or its products.

Although Dworkin's argument from the challenge view to endorsement via intention fails, there could be a more direct connection between the challenge view and endorsement. That is, we may agree with Dworkin that a skillful performance is undermined when the person does not regard it as valuable. The above examples go against this claim; they show that a performance can be skillful even if the performer does not value his activities. But Wilkinson argues that whilst it may not be generally so, endorsement is needed for some activities to have value. Forced religious worship, for example, would be pointless, since a religious life only has value (if it has value at all) if it contains sincere devotion rather than merely the external actions of worship, brought about by force. In such cases, says Wilkinson, endorsement may be necessary and the constitutive view is plausible.[52] Similarly, Kymlicka argues that the endorsement constraint is true 'for a broad range of human excellences, from reading poetry and writing novels to friendship, conversation, and social ritual',[53] even if it cannot be extended to all activities.

Even in these cases, however, endorsement is not necessary for value. The right motive may be necessary for performance, but from this it does not follow that endorsement is needed. Sincere religious devotion, for example, does not require the belief that such devotion is worthwhile. It requires that devotion is motivated for its

---

[51] The example is from Wilkinson, 'Dworkin on Paternalism and Well-being', pp. 439-40 who uses Conan Doyle's attitude to his Sherlock Holmes stories as an illustration.

[52] Wilkinson, 'Dworkin on Paternalism and Well-being', p. 440.

[53] Kymlicka, *Liberalism, Community, and Culture*, p. 19. It has already been seen that endorsement is not necessary for the value of writing novels.

own sake and not just because one is threatened. This much is true and may be supported by the challenge model. As Dworkin says, according to that model, 'involuntary religious observance, prayer in the shadow of the rack'[54] will have no value. But a person may be devoted to a religion for its own sake and not due to threats but at the same time not think such devotion is a valuable component of her life. She may think that it does not contribute to her well-being, but that that does not mean she should not do what she believes she must. Intention to do something for its own sake, rather than endorsement, is also what is needed for the other activities Kymlicka lists to have value.

Although endorsement is not needed for many activities to contribute to the value of a person's life, it is clear that intentions are important. Many activities need to be done with the right motives to have value. This is obviously relevant for the perfectionist case against paternalism. This argument will be examined in the next chapter.

4.6 Integrity and Endorsement

Dworkin's next argument for the constitutive view appeals to the value of ethical integrity as part of the challenge model of the good life. Ethical integrity, according to Dworkin,

> is the condition someone achieves who is able to live out of the conviction that his life, in its central features, is an appropriate one for him, that no other life he might live would be a plainly better response to the parameters of his ethical situation rightly judged.[55]

Ethical integrity is reduced when people live mechanically without conviction, when they set their convictions aside sensing that they are not living as they should, or when they believe that they have less resources than justice requires. Most importantly for the issue of paternalism, ethical integrity also 'fails conspicuously when people are

---

[54] Dworkin, 'Foundations of Liberal Equality', p. 78.
[55] Dworkin, 'Foundations of Liberal Equality', p. 80.

made to live in a way they regret, and never endorse, by the fiat of other people.'[56] Dworkin argues that the challenge model of living supports giving priority to ethical integrity in judgements about people's well-being. Such priority would mean that ethical integrity is not just one value that can be weighed against other features of the good life. Balancing of that sort would result in the additive rather than the constitutive view. The priority of ethical integrity means that a life than never achieves integrity can never be better than one that does.[57] Forcing someone into a way of life he does not endorse undermines his sense of living the best life he can, and such a loss can never be made up by a gain in other values.

It may be thought that the condition of ethical integrity is too strong. Very few people would believe that there is no way they could have responded better to their lifelong situation. Everyone has nagging doubts that things could have been better, that mistakes have been made, and opportunities were not taken up. If so, few if any would ever achieve the condition of ethical integrity. However, this is to interpret the condition as an all-or-nothing requirement. According to Dworkin, ethical integrity is a matter of degree.[58] Thus people can have successful lives with ethical integrity even if those lives could be further improved if they could live up to their convictions to an even greater degree.

However, there are a number of problems with Dworkin's argument. It may be formally restated in the following way:

1. A good life is responding in the right way to the challenge of living.

2. Responding in the right way to the challenge of living requires giving ethical integrity priority over other values.

---

[56] Dworkin, 'Foundations of Liberal Equality', p. 80.

[57] Dworkin, 'Foundations of Liberal Equality', p. 80.

[58] Dworkin, 'Foundations of Liberal Equality', p. 83. Dworkin says that integrity is a matter of degree 'over some range'. Thus there may be some threshold below which people can not achieve the condition at all. Above it, integrity is a matter of degree.

3. The priority of ethical integrity implies the endorsement constraint.

Therefore:

4. A good life requires the endorsement constraint.

Whilst there is some connection between the notions of ethical integrity and endorsement, the third premise is false. Living out of the conviction that no other life would be a better response to one's situation implies that one thinks that the components of one's life are valuable, since responding well to one's situation as one believes it ought to be involves choosing activities, projects, and relationships that one regards as worthwhile.[59] If one does not think those components are valuable, one would choose differently and thereby respond in a different way. However despite this connection, the move from the priority of ethical integrity (henceforth PEI) to the endorsement constraint is not valid. PEI means that living up to one's convictions, whatever they are, has value and that this value has priority over all others. But this means that, ethical integrity apart, there could be other values that can contribute to a person's well-being. Of two lives containing the same measure of ethical integrity, one containing more of some other value is better than one containing less. Hence, on this view well-being can be advanced independently of the endorsement of the person whose life it is. Endorsement therefore would not be a necessary condition of value, which is what the constitutive view claims.[60] PEI results in the priority of the value of endorsement rather than its necessity. Priority and necessity differ since according to the former, a life may still have value despite the absence of endorsement (but never as much as any life containing endorsement). This is what the latter denies. So EC cannot follow from PEI since they make a contradictory claim about the value of unendorsed components of a person's life.

---

[59] The stipulation that the response be to what one believes one's situation ought to be is important since responding to one's situation whatever it is would not amount to an endorsement of one's life if that situation is an unjust or impaired one. Rather, such a response could simply be making the best of a bad situation.

[60] This point is from Wilkinson, 'Dworkin on Paternalism and Well-being', pp. 437-8.

But this is a small objection. Whilst giving ethical integrity priority results in a logically different position than that implied by the constitutive view, the difference in the implications for paternalism of each are not great. Both make paternalism self-defeating when it tries to force people into ways of life they do not accept. Those people will be worse off, either because though getting some improvement they are always worse off overall because they miss out on a value that is prior to all others or because they get no improvement at all.

So whilst the link between PEI and EC is not sound, the weakness may not trouble us much. But the second premise, concerning the connection between the challenge view of the good life and the priority of ethical integrity, may also be questioned. Dworkin gives no clear argument for how the former leads to the latter. PEI is difficult to defend. Consider two possible lives. The first contains a certain amount of integrity and components with a small amount of value. The second has only a slightly lesser amount of integrity and components that are of vastly greater value. If ethical integrity has priority then the first is the better life but it seems implausible that this could always be true regardless of how much other value the second life has.[61]

Is there an argument for giving priority to ethical integrity? Some of Dworkin's comments suggest one. He imagines that someone falsely believes that a life in religious orders is of greater value than one in politics. Assume that a life in politics is worthwhile and a religious life is not, but the person believes the reverse is the case. There are two lives he might lead. He might adhere to his convictions and live a life of religious devotion with full satisfaction and confidence in his choice. Or he could bow to the advice of his friends (who think his beliefs are mistaken) and enter politics, against his convictions. Dworkin claims he could be successful there, 'but will find no genuine satisfaction or self-approval, and will therefore never cease regretting his

[61] Again, this point is from Wilkinson, 'Dworkin on Paternalism and Well-being', p. 442.

choice.'[62] Dworkin claims the first life is better than the second and that his ranking reflects his commitment to the priority of integrity. Even if it is true, says Dworkin, that a life in politics would be more worthwhile than one in religion, given the person's unshaken convictions (even if they are misguided), the religious life is the only one he can lead at peace with himself and it is therefore the best he can do in meeting the challenge of his situation.[63]

This example spells out the implications of the priority of ethical integrity and may be thought to assume the truth of that view. But perhaps it is an argument for that view. Say we accept the ranking Dworkin sets out. If a view provides a plausible explanation for the ranking, so the argument could go, then that counts in favour of the view. The priority of ethical integrity provides such an explanation. But there is another plausible explanation of the case: that the ranking reflects NRC, which states that nothing can contribute to a person's well-being if he resents it. The life in politics would be led against the person's 'instincts and convictions', with 'no genuine satisfaction or self-approval' and unceasing regret.[64] He leads, says Dworkin 'a false, distorted life, at war with his own ethical sense.'[65] Such a life would therefore violate NRC. Hence, even if we agree with Dworkin's ranking of the two possible lives, that is not necessarily because we share his commitment to the priority of integrity. Such a ranking can be explained by NRC.

Moreover, the explanation provided by NRC (which as has been seen is also plausible for other reasons) is more plausible than that provided by giving ethical integrity priority. One difference between the two views is that according to PEI, unendorsed politics is worse than endorsed religion[66], and this is always the case no matter how

---

[62] Dworkin, 'Foundations of Liberal Equality', p. 82.
[63] Dworkin, 'Foundations of Liberal Equality', p. 82.
[64] Dworkin, 'Foundations of Liberal Equality', p. 82.
[65] Dworkin, 'Foundations of Liberal Equality', p. 80.
[66] Here again it can be seen how PEI and EC differ. The latter says that both lives are equally worthless. Religion has no value and endorsement adds nothing to it, while politics is valuable but

much more value, endorsement aside, politics has than religion. Living in accordance with one's beliefs, PEI implies, has priority over all other values insofar as they contribute to well-being. NRC on the other hand does not have the same implication. On that view, unendorsed politics that is also resented (as it seems to be in the case Dworkin describes) has no value. Endorsed religion has none because religion (by hypothesis) is worthless. Hence the two lives are equally worthless; unendorsed politics is not worse than endorsed religion. But NRC can be combined with a claim that if the resentment is severe enough, then this may produce negative amounts of well-being. Resentment undermines well-being; an activity does not improve a person's life if she resents it and may make her worse off if she resents it to a significant degree. This is why we could accept Dworkin's ranking that unendorsed politics is worse than endorsed religion. The way he describes the former makes the resentment sound severe. The explanation I have given makes the ranking possible and sometimes true but not necessarily true, as PEI implies. But this is an advantage of the explanation rather than a weakness. We should not think that a life of unendorsed good activities is *always* worse than a life of endorsed bad ones. If there is severe resentment then it may be. But if the good activities have great value and only a small amount or no resentment then it need not be worse.

Moreover there is reason to think that ethical integrity cannot have priority upon Dworkin's challenge model of the good life. One implication of that model is that distributive justice is a parameter of the good life. That means that a person's life cannot be better with an unjust share of resources. If so, however, ethical integrity cannot have priority.

---

does not contribute to well-being due to the absence of endorsement. PEI says that unendorsed politics has some value due to the value of politics, but endorsed religion (always) has more value since while religion has no value, endorsement does and this value always has priority over others. For this to be true, I assume that PEI implies that endorsement adds value to a person's life even if the other components are worthless. A different version, one which seems at odds with Dworkin's view, would say that ethical integrity is prior to other values but only has value so long as there are some other values in a person's life. This view would make endorsed religion worthless (assuming religion has no value) and therefore less valuable for a person that unendorsed politics.

To see this, some brief exegesis is needed. According to Dworkin, the parameters of one's ethical situation are the conditions that help define what a good life would be. Some circumstances in life can be classed as limitations on how well one can live. A person may count his lack of talent, or wealth, or the range of options open to him as limitations on how good his life can be; without the limitations, he would be able to lead a better life. But not all circumstances, claims Dworkin, should be classed as limitations. Just as an artist working within the tradition of a certain genre will count the requirements of that genre as parameters which help define his success in working in that genre rather than as limitations on how good his work can be, so certain conditions in life can count as parameters for living a good life rather than as limitations. Being a member of a certain political community rather than another, for example, could be such a parameter. It states a condition of a good life as one that is 'appropriate to someone whose situation includes that connection [to the political community].'[67] Dworkin argues that justice in the distribution of wealth is one correct parameter. A person's life goes worse if he is denied the resources that justice requires that he should have (and also if he has more than justice allows).[68]

Ethical integrity, for Dworkin, occurs when a person best responds to the parameters of his ethical situation, where she determines which of her circumstances are parameters. The subjectivity of the classification of parameters gives the notion of ethical integrity its anti-paternalistic nature. If people count paternalistic restrictions as limitations then those restrictions will undermine ethical integrity since people will believe that their lives could have been even better without those restrictions. Ethical integrity fails 'when people believe, *rightly or wrongly*, that the correct normative

---

[67] Dworkin, 'Foundations of Liberal Equality', p. 68.
[68] Dworkin, 'Foundations of Liberal Equality', pp. 71-5.

parameters have not been met for them ... [and] ... when made to live in a way they regret, and never endorse, by the fiat of other people.'[69]

But the problem is this: making ethical integrity turn upon how a person judges her parameters means that justice is not necessarily a parameter of the good life. Say a person does not believe that justice is a parameter of the good life. Unconvinced by Dworkin's argument to the contrary, he believes that justice is a limitation on how good his life could be. He believes that if he had more resources than justice allows, then his life would be better. If there is a just resource distribution, then he will believe that his life is not as good as it could have been and therefore not achieve ethical integrity. By giving him more resources, more than justice permits, he will achieve ethical integrity. And according to Dworkin, ethical integrity has priority over other values. But this makes it difficult to see how justice can be a parameter of the good life. It certainly cannot be a *hard* parameter, one that if not met means life has no value at all. If that were the case then in the absence of a just distribution of resources, a person's life has no value no matter how much ethical integrity he has. But in the case described, the person's life must have much value due to his high ethical integrity. But justice also cannot be a *soft* parameter, one that if not met makes a life worse but still able to have some value, at least not the sort of soft parameter that Dworkin imagines justice to be.[70] He claims that whilst someone with more resources than justice permits could live a better life than he could have led in a just society, most people would be worse off; they 'will do nothing so brilliant or amazing with the surplus over justice that it will compensate for their inability to lead a life good for a just community.'[71] However, so long as they regard justice as a limit rather than a parameter then the priority of ethical integrity will mean that, with more resources than justice permits, these people *always* lead better lives than they would if

---

[69] Dworkin, 'Foundations of Liberal Equality', p. 80. Emphasis added.

[70] For the distinction between hard and soft parameters, see Dworkin, 'Foundations of Liberal Equality', p. 70. For his argument that justice is a soft parameter see p. 75.

[71] Dworkin, 'Foundations of Liberal Equality', p. 75.

they had no more than permitted by justice. With PEI, living in accordance with one's convictions, including convictions about parameters, has priority over other values, including the value of living in accordance with actual parameters. Accepting PEI means that justice can only be an *extra soft* parameter. A person's life does go better in one way if society is just, but never better all things considered if a person believes he would be better off with more than justice gives him. Injustice never makes a person's life worse overall so long as it produces ethical integrity. In a move toward justice, well-being is improved in one respect but always overall worse. A person's life can never be overall better in a move to justice. A move away from justice is always an improvement. A person's life could never be better with only the resources justice permits than it could if he had more, since without more he lacks ethical integrity and such integrity is prior to other values.

Hence it does not seem possible to hold both the priority of ethical integrity and justice as a parameter of the good life, at least in the sense that people's lives can be worse overall in an unjust society when they believe they are living the best lives they could. PEI says that their lives are always better than if there was a just distribution. One way out of this dilemma would be to drop the claim that justice is a parameter of the good life. But as Dworkin argues, such a parameter seems implied by the challenge model of living. Since according to that model living well means responding skillfully to the right challenge, 'then a life goes worse when the right challenge cannot be faced.'[72] And justice is part of the right challenge. So abandoning justice as a parameter would require abandoning the challenge model of living, and then we would be without premise 1 of the argument for EC from the challenge model. We would need an alternative argument for the priority of ethical integrity since, as we have seen, it is counter-intuitive.

---

[72] Dworkin, 'Foundations of Liberal Equality', p. 73. He also states that people's lives go badly 'because we have the wrong circumstances.' (p. 69.)

challenge model of living. That model supports holding justice as a parameter of the good life and such a parameter conflicts (except in the weak sense of justice being an extra soft parameter) with giving priority to ethical integrity. Without support from the challenge model, PEI should be rejected since it is counter-intuitive and not supported by Dworkin's other argument for it (the argument from the ranking of unendorsed politics being worse than endorsed religion). Dworkin's argument for EC therefore fails since the second premise is false. It fails in another way too. Establishing PEI would be an argument against paternalism in itself regardless of its support for EC, since PEI alone clearly has anti-paternalistic implications. But as we have seen, Dworkin's arguments fail to establish the truth of PEI.

4.7 Conclusion

Dworkin's arguments for the endorsement constraint therefore fail. We are left with the conclusion that the non-resentment constraint is more plausible than the endorsement constraint. NRC can account for the considerations that seemed to support EC and NRC implies (whilst EC denies) what seems to be true, namely that a person's life can go better through having valuable components when she has a neutral attitude towards those components, neither endorsing them nor resenting them. Finally, NRC can provide an explanation of the ranking of lives (unendorsed politics versus endorsed religion) that Dworkin believes supports the priority of ethical integrity and even explains how there can be exceptions to that ranking, which such priority denies. NRC, however, as I attempted to show in 4.4, is likely to allow significantly more paternalism than EC. Hence perfectionism cannot rely upon this argument to show how it can avoid its strongly paternalistic implications.

Chapter Five

Hurka on Activity and Intention

5.1 Introduction

As was seen in chapter two, one of the arguments that Thomas Hurka, along with other perfectionists, makes in order resist the paternalistic implications of perfectionism appeals to the value of autonomy. This chapter examines several other arguments against paternalism made by Hurka. As was noted in chapter two, Hurka's theory begins with an account of the good life, characterising this as realising what is central to human nature. When extended to politics, perfectionism holds that 'the best political act, institution, or government is that which most promotes the perfection of all humans.'[1] Because perfectionism holds that some ways of life are better than others regardless of whether people want or would choose them, Hurka admits, it may be thought to favour state coercion to force people into excellence. Hurka's arguments aim to resist this charge. He tries to show how paternalism would threaten certain perfectionist goods. I will try to show that the arguments do not succeed in their aim. Hurka's perfectionism remains committed to a wide range of paternalistic policies.

The strategy I take is the same as the one I took in section 2.4 towards Hurka's arguments for autonomy. I describe three hypothetical cases of what I take to be unacceptable paternalism and attempt to show how a state committed to Hurka's theory would support paternalism in each of these cases. Section 5.2 summarises Hurka's arguments. Sections 5.3 through 5.5 each focus on a case of unacceptable paternalism and try to show that Hurka's theory not only fails to rule out paternalism

[1] Thomas Hurka, *Perfectionism*, New York and Oxford, Oxford University Press, 1993, p. 147.

in each case but would actually support it. The final section attempts to determine where my argument, if it is correct, leaves Hurka's theory. I suggest that it should lead us to reject perfectionism and accept state neutrality, at least with regard to paternalism.

5.2 Hurka's Arguments

5.2.1 Hurka's main argument rests on the notion that perfection is not something that individuals can directly produce in others. Other people, including governments, can supply necessary conditions that make people's perfection more likely to occur but beyond those conditions, 'each person's achievement of perfection must be his own.'[2] There is an asymmetry between what a person can do to bring about his own perfection and what he can do to bring about others'. This is because perfection is *active* and *inner*. Perfection is active in the sense that it

> involves doing things, forming goals and realizing them in the world. And each person's doing must be largely her own, reflecting her energy and commitment.[3]

A similar point is made by Raz. An important aspect of well-being, according to him, is that it is active. This means that 'no one can make a success of another person's life.'[4] Bringing about others' well-being is logically impossible. We can avoid interfering and provide encouragement and resources. But we cannot produce others' perfection ourselves. The achievement of active goods must be a person's own. My life can only go well, upon this argument, if I achieve my goals myself rather than have others attempt to bring about their achievement.

---

[2] Hurka, *Perfectionism*, p. 152.

[3] Hurka, *Perfectionism*, p. 64.

[4] Joseph Raz, 'Duties of Well-Being', in his *Ethics in the Public Domain*, revised edition, Oxford, Clarendon Press, 1996, p. 8.

Perfection is *inner* in the sense that it 'involves acting from complex, structured intentions.'[5] An activity has more value to the extent that it is performed with the right intentions and less when those intentions are absent. This is similar to the idea examined in chapter four, namely that for any activity to contribute to well-being, it's value must be endorsed by the person involved. The two ideas are, however, not the same. There can be complex, structured intentions without endorsement, as the example of an author who writes complex works without believing them valuable illustrates. And there can be endorsement without complex, structured intentions, as in the case of an apprentice artist who has his hand moved across the canvas by his master. If complex, structured intentions are needed for an activity to have perfectionist value, paternalist laws requiring perfection cannot be enforced since people's inner states cannot be enforced. Requiring or prohibiting certain acts will not necessarily further perfection since the required inner states may not be present.[6]

5.2.2 Hurka claims that the asymmetry shows that many paternalistic interventions would do no good. But he adds that this would only make interference morally indifferent and 'a serious defence of liberty must show that it is wrong.'[7] He gives three reasons for thinking it harmful.[8] First, paternalist laws can limit people's routes to excellence. Any activity, if approached in the right spirit, can have value. To forbid some options is to deny some possibilities of genuine perfectionist worth. Those who miss out on such possibilities without discovering anything else as good are harmed.

Second, paternalist laws can make people's intentions worse. According to Hurka, choices for or against a course of action are best when part of a plan for one's whole

---

[5] Hurka, *Perfectionism*, p. 65. These intentions may be unconscious. See *Perfectionism*, pp. 125-6.
[6] Hurka, *Perfectionism*, p. 153. This has an obvious historical affinity with the Lockean argument that religious conversion cannot succeed with the use of coercion since it would not result in genuine belief.
[7] Hurka, *Perfectionism*, p. 153.
[8] The three reasons are on pp. 154-5 of *Perfectionism*.

life or for one's community. Legal threats cause people to act through fear of punishment, thus supplanting these more worthwhile motives. When people act because of this fear and would have acted for better reasons, there has been a perfectionist loss. Paternalism makes people's intentions worse in another way. According to Hurka, 'the best acts are chosen partly for themselves rather than just as a means to something beyond them.' Legal threats undermine this by making intrinsic choosing less likely. Hence, there are two supposed harms involved, both of which worsen people's intentions. The first emphasises the goodness of choosing activities with the intention that they be part of a wider plan of life, whilst the second emphasises the goodness of choosing activities because of intrinsic rather than instrumental qualities they have. The threat of punishment, argues Hurka, undermines both these types of intentions. The two harms differ because a wider plan of life may or may not be dominated by goals that are chosen for intrinsic reasons.

Thirdly, paternalistic intervention undermines the general character needed for excellence by diminishing energy and self-direction. People become indisposed towards exercising their capacities for human perfection.

According to Hurka, these arguments (along with the appeal to autonomy) 'comprise a composite perfectionist case for liberty.'[9] State interference with self-regarding actions rarely succeeds in promoting other perfections, and can diminish them by removing routes to excellence, inducing less valuable motives, and weakening self-direction. Hurka illustrates his case for liberty by pointing out that his theory would be opposed to paternalist laws seeking to ban certain sexual activities. He claims that his 'Aristotelian sexual ideal differs importantly from the views normally behind a call for the enforcement of sexual morals.'[10] Holders of such views usually object to the physical aspects of non-procreative sex, but Hurka's perfectionism has no such

---

[9] Hurka, *Perfectionism*, p. 155.
[10] Hurka, *Perfectionism*, p. 156.

objections if the activities are approached with the right intentions. Prohibitions on sexual activities would also invade autonomy and the exercise of practical rationality that it involves, may make valuable options unavailable for many people, and may supplant worthwhile intentions by encouraging people to act through fear of punishment.[11] On the other hand, Hurka's theory will allow paternalism in other areas. He says that perfectionism may approve seatbelt legislation, compulsory medical insurance and perhaps laws discouraging smoking but that it has serious reservations about these laws. They may only be implemented after a careful weighing of the costs.[12] Whilst these examples may seem to illustrate the plausibility of Hurka's theory, other cases, as I hope to show, reveal that it is unacceptable.

5.3 Lifestyle Paternalism

5.3.1 One type of paternalism mentioned in chapter two was *lifestyle paternalism*, which aims to guide people into more worthwhile lifestyles. An illustration of this is a government policy of discouraging couch potato lifestyles. Such a policy could be possible and enforceable.[13] We are rightly suspicious that lifestyle paternalism would be a legitimate state policy. Couch potatoes should be free to pursue their chosen lifestyle unhindered by the state. It would be unacceptable for the state to prosecute people for watching television and drinking beer too much. But it is not clear that Hurka's theory can explain why. On the contrary, his theory would seem to support lifestyle paternalism. This may be shown by examining each of Hurka's arguments as they apply to the case of couch potatoes.

5.3.2 Hurka's argument from the active character of perfection does not rule out lifestyle paternalism. That argument claimed that perfection must be one's own doing; others must not choose or achieve one's own goals for one. The argument rules out

---

[11] Hurka, *Perfectionism*, pp. 156-7.
[12] Hurka, *Perfectionism*, p. 158.
[13] See section 2.4.2 for some discussion of its enforceability.

forcing someone into specific projects and bringing about the achievement of the goals associated with those projects since those choices and achievements would not be the person's own doing. But it does not rule out prohibiting some options while leaving many others available. Whilst forcing couch potatoes out of their beer-drinking, soap/gameshow-watching lifestyles does prevent them from pursuing their chosen projects (if they count as projects), it allows them to choose from amongst many more. Since it is inevitable that they will have to choose something else instead (they cannot do nothing; that is what they are being prevented from doing), then there is likely to be no loss of activity overall.

Moreover, there is a likely gain. It is doubtful that a couch potato's lifestyle counts as a perfectionist activity at all. It is an example of forming goals and realising them in the world only in the weakest sense possible and hardly reflects her 'energy and commitment.'[14] By being prevented from living that lifestyle she will be forced to choose something else instead, something more likely to count as activity in the perfectionist sense. There will thus be a perfectionist improvement in her life.

Even if a couch potato's lifestyle does count as an activity in the perfectionist sense, if she is prevented from choosing that lifestyle then she will choose another which is likely to be more valuable and there will again be a perfectionist improvement in her life. This is because her more active lifestyle will be better than her beer-drinking, television-watching one if, firstly, it affects more people in more significant ways and, secondly, if it involves a hierarchy of goals consisting of means and ends.[15] Since a couch potato's chosen lifestyle is so low on both these counts then anything she chooses instead as a result of lifestyle paternalism is likely to be an improvement.

---

[14] Hurka, *Perfectionism*, p. 64.

[15] These are the two criteria according to which Hurka claims the value of activities should be judged. See Hurka, *Perfectionism*, pp. 115-6.

Hurka's argument that perfection requires certain inner states cannot help couch potatoes either. The argument tries to show that unless they are approached with the right intentions, certain activities will not contribute to well-being. Like the above argument, this one also makes forcing a person into one specific option probably pointless since doing so will not ensure that he will have the right intentions. But it has less force against prohibiting one option while leaving open many others, which is what is involved in lifestyle paternalism. That would allow significant scope for the person to choose something else. With this wider scope, there is the possibility that he will choose something with the right intentions.

Although it is true that forcing him out of his bad lifestyle does not ensure that he will come to have the intentions that are needed for a more worthwhile life, a couch potato does not have the appropriate intentions for a good life anyway, so lifestyle paternalism would not result in any perfectionist loss. A person's activities have value, according to Hurka, when they are performed with the intention to affect many significant aspects of one's own and other people's lives and with the intention that the activity form part of a complex hierarchy of goals. Activities intended for trivial reasons or that have no relation to broader ends have little value.[16] Hence a couch potato's life of drinking beer and watching TV soaps and gameshows would have little value since, by hypothesis, his activities are not intended to form part of any organised structure of life.

There could in fact be some perfectionist gain as a result of lifestyle paternalism. The state cannot directly ensure that couch potatoes have the intentions that help make a good life, but it can provide conditions that make them more likely to occur. Any coercive lifestyle paternalism could be accompanied by public education programmes using leaflets and advertising, to convince people of the value of having an organised structure of life where goals are pursued as part of a hierarchy of aims that have

[16] Hurka, *Perfectionism*, pp. 153, 65, 121.

pervasive effects on one's life. These non-coercive measures in themselves may be unobjectionable policies. But my point is that they would help make perfection, on Hurka's account of it, more likely to occur through the use of coercive measures, which are objectionable.

5.3.3 Hurka's argument that intervention can cause harm by limiting people's routes to excellence is not convincing in a couch potato's case. As already noted, such a lifestyle can only lead to perfectionist worth if it somehow involves critical deliberation, or a fully conscious choice of the lifestyle and its qualities and a rejection of available alternatives, or is intended to form a structured hierarchy of pervasive goals. But couch potatoes fail to satisfy any of these criteria. Hence even if it is true that, as Hurka says, 'almost any act can, if approached in the right spirit, embody some perfection'[17], a couch potato's lifestyle does not because he does not approach it in the right spirit.

It could be replied here that it is still possible that a couch potato may somehow get some perfectionist value from his lifestyle and this possibility is enough to rule out interference. Perhaps his intentions may change in such a way that he approaches his activities in a way more conducive to the achievement of perfection. But if so, then it begins to be difficult to see what work Hurka's theory is doing. If the mere possibility of perfectionist value is sufficient to protect any activity from intervention, there is very little that a state can do to promote perfectionism. Not only may it not prohibit any options but even promoting some activities through public education programmes, without any actual interference at all, may, if successful, turn people away from activities with possible perfectionist worth. This may not be a problem. Maybe on Hurka's account the state should not target certain activities, but just promote intentions and critical attitudes. However, this conflicts with what Hurka

[17] Hurka, *Perfectionism*, p. 154.

explicitly says. He admits that his arguments do allow forbidding some bad activity while allowing others. Such a policy

> reduces autonomy only a little, removes just one route to excellence, and leaves considerable scope in the choice among remaining options for the operation of good motives. If it steers enough people into better activities, it can be on balance justified and can be recognized by a government as justified. So although perfectionism has a strong general commitment to liberty, it may sometimes favour legal restrictions on especially bad activities.[18]

Hurka cites drug use as an example of these bad activities. But idle lifestyles such as those of couch potatoes would also seem to be covered.

Another way that paternalistic intervention can cause harm, according to Hurka, is by making people's intentions worse. Legal threats may substitute fear of punishment for more worthwhile motivations. This presupposes that a couch potato would choose, in the absence of lifestyle paternalism, to reject for perfectionist reasons a life of drinking beer and watching television. Then lifestyle paternalism might result in him so choosing only to avoid punishment. However, it may be assumed that a couch potato would not choose to reject his current lifestyle and that a state considering lifestyle paternalism can be sure of this. That is presumably what motivates the paternalistic measure; if there was reason to think that couch potatoes will eventually reject their lifestyles then paternalism is unnecessary. They are content in their lifestyles and will continue watching soaps and game shows while drinking beer for the remainder of their lives. They have no worthwhile motives to be supplanted by bad ones.

There is another way in which it may be thought that interference with couch potatoes may make their intentions worse. Even if they will never do anything other than drink beer and watch television, perhaps they will come to do this for more worthwhile motives. Perhaps a couch potato will form 'an organizing plan for one's whole life'

---

[18] Hurka, *Perfectionism*, p. 156.

which is 'extended in times or persons or often part of a complex hierarchy.'[19] Lifestyle paternalism would then be harming him by bringing about a perfectionist loss when it forces him out of his lifestyle. However, the rebuttal in the previous paragraph applies here too. Many couch potatoes, it may be assumed, do not and will not ever have more worthwhile motives for their lifestyles. They are unreflective and completely unconcerned with having an organised structure or overall plan of life. Lifestyle paternalism's interference does not supplant any worthwhile motives because there are none to supplant.[20]

Hurka's assumption in the above argument, that fear of punishment from legal sources may supplant better motives for activities, is usually false. Someone who would choose against the couch potato lifestyle in the absence of lifestyle paternalism, say because she views it as a worthless way of living, would not cease to think the same once lifestyle paternalism began. The threat of legal punishment would simply be an additional motivation for choosing as she does. She would still choose against the couch potato lifestyle because of the worthlessness she sees in it. There are legal sanctions against murder and theft. Yet their presence does not supplant more worthwhile motivations. Many people, perhaps most, refrain from such acts out of a belief that the acts would be immoral, as well as or even regardless of the threat of punishment.[21]

---

[19] Hurka, *Perfectionism*, p. 154.

[20] Moreover, the distinction in this paragraph between couch potatoes and those who live similar lifestyles but with better motives may be recognised in law. People prosecuted by a policy of lifestyle paternalism could defend themselves by providing evidence that their lifestyle has perfectionist worth. This evidence could include, for example, demonstrating a knowledge of the subtle flavours of different beers or of the semiotics of soap operas.

[21] This analogy was suggested to me by Paul Markwick. In making this argument, I do not wish to deny that the presence of sanctions may undermine some of the reasons that may otherwise be present in the situation. In particular, Meir Dan-Cohen (see 'In Defense of Defiance', *Philosophy and Public Affairs*, vol. 23, no. 1, 1994) claims that the threat of coercion undercuts the authority of the state. He argues that where there is authority and therefore a reason to x simply because the authority has commanded x, once the command is backed by the threat of coercion, the reason to x simply because it is commanded is undermined and sometimes obliterated. However, Dan-Cohen does not argue that, and nor do any of his arguments seem to support the proposition that, the threat of coercion also undermines the non-authoritative reasons to do x, reasons that depend upon the merits or substance of doing x. He admits this on p. 49.

The same point may be made against Hurka's next argument. This is that paternalistic intervention would undermine intrinsic choosing, that is, choosing activities for themselves rather than as means to some other end. The best activities are chosen partly for themselves, according to Hurka, and not out of fear of punishment. Again, there is no reason to believe that people would cease to choose against bad lifestyles like the couch potato's because of their intrinsic worthlessness when lifestyle paternalism starts. The legal threats of lifestyle paternalism would simply provide additional motivation not to choose those lifestyles.

In fact, on Hurka's account their lives would be improved once lifestyle paternalism starts. His claim is that the best acts are chosen partly for themselves rather than just as means to something else. He admits that the best situation is one where an activity is chosen both for itself and as a means to something beyond it. This would be a more complex hierarchy of goals.[22] Since the threat of punishment gives them an additional reason for acting as they do, lifestyle paternalism improves the lives of those who would choose against a couch potato lifestyle anyway, on Hurka's account, by improving the complexity of their hierarchies of goals.

Even if Hurka is correct and lifestyle paternalism would undermine intrinsic choosing, the argument of course does nothing to help couch potatoes. It only helps those who would choose some other lifestyle instead for intrinsic reasons and end up choosing only out of fear of punishment. They suffer some perfectionist loss. But couch potatoes do not and would not choose something else.

Several of these arguments of Hurka's stress the bad effects that paternalism would have on those who would freely choose against the bad anyway. But perhaps there is some way that lifestyle paternalism can target couch potatoes while not harming the

---

[22] Hurka, *Perfectionism*, pp. 126-7.

other class of people who would otherwise choose against couch potatoism for worthwhile reasons. If so, then Hurka's account would stand accused of supporting objectionable paternalism since it would support coercion against couch potatoes. If not, that is, if targeting couch potatoes also supplants the worthwhile motives of those who reject couch potatoism, then Hurka still faces the difficulty of showing that the perfectionist loss that would accompany lifestyle paternalism outweighs the perfectionist gain that would accompany it. Why do the gains from eliminating the couch potato lifestyle not outweigh the costs of undermining intrinsic choosing?[23] Moreover, even if a perfectionist balancing of gains and losses would come out against lifestyle paternalism, Hurka still faces another difficulty. The badness of lifestyle paternalism, on Hurka's present argument, consists in its effects on those who would otherwise choose intrinsically not to be couch potatoes. The argument does nothing to show that couch potatoes would be wronged by lifestyle paternalism, only that others (non-couch potatoes) might be harmed. A more satisfying explanation would show that *couch potatoes* themselves would be wronged by lifestyle paternalism.

A variation of Hurka's argument may help the couch potato. Instead of undermining the possibility of her intrinsically choosing some more worthwhile ways of living, perhaps there is some value in her intrinsically choosing her otherwise worthless lifestyle so that lifestyle paternalism does her harm when it prevents her from making that choice. However, the couch potato's lifestyle has no value in itself as well as not being connected to any other worthwhile goals. It is not a goal that she has committed herself to for her entire life. Rather, it is just something she does all day every day without much thought. It does not form part of, nor is it connected to, any hierarchy of goals. Nor is it a life organised around a single end. On the contrary it is

---

[23] See Richard Arneson, 'Paternalism, Utility, and Fairness', *Review of International Philosophy*, vol. 43, no. 170, 1989 for some interesting thoughts on the balancing of costs and benefits for different people of paternalism. Nothing he says there would seem to help Hurka's case. He argues that, on grounds of fairness, the gains for the potential targets of paternalism would outweigh costs to others.

characterised by an absence of ends. Nor is it in any way complex, intricate, or challenging.[24]

Even if there is some value in the couch potato's intrinsic choice of his lifestyle, it still does not follow that lifestyle paternalism is harming her by preventing her from making that choice. There is only harm if there is nothing else of value for her to choose instead. But since lifestyle paternalism leaves open many other valuable options, it is unlikely that she will be unable to find anything else that she believes is valuable in itself. If so, then she is not worse off through being denied intrinsically choosing his lifestyle of lying on the sofa, drinking beer, and watching television.

Hurka's final argument is easily rebutted. Intervention with couch potatoes would not diminish energy and self-direction, thereby undermining the general character needed for excellence. On the contrary, they have no energy and self-direction anyway and by hypothesis will not have any in the future. Lifestyle paternalism may do something to stimulate them into activity.

Hurka's arguments do not succeed in ruling out lifestyle paternalism. On the contrary, they support it. Lifestyle paternalism would protect and further couch potatoes' autonomy by making them rationally deliberate and exercise fully conscious agency. It will enable them to achieve more perfection by performing activities that affect themselves and others in significant ways and that involve a more complex hierarchy of goals. Lifestyle paternalism could give couch potatoes more intentions and motivations in life and encourage them to choose activities for worthwhile reasons. Finally, lifestyle paternalism could stimulate their energy and self-direction and thereby improve their general characters.

---

[24] Being organised around a single end and being complex, intricate, and challenging are other ways that make activities valuable according to Hurka. See *Perfectionism*, pp. 121 and 123. Hurka justifies these claims by reference to the value of having pervasive goals and being part of a complex hierarchy. See *Perfectionism*, pp. 121-2 and 123-5.

5.4 Aesthetic Paternalism

5.4.1 A second type of paternalism is *aesthetic paternalism*, which occurs when the state tries to force people into aesthetically more worthwhile activities. Imagine that the state prohibits the broadcast, sale, and distribution of recordings of a genre of music that is inferior say because the genre consists of songs with repetitive dance beats, uncomplicated melodies played by electronic synthesizers, and vacuous lyrics sung by people who cannot sing. The music has instant superficial appeal but no lasting value. Many people choose it over more worthwhile music due to its instant appeal, even though in the long run they would gain a deeper appreciation of music and more enjoyment overall if they listened to something else. The state could suppress such music in the hope that people will listen to more worthwhile music instead.

Although the above example is hypothetical, similar policies have been implemented by several governments. Most governments censor individual songs or records[25] but some target entire genres.[26] We are, I believe, rightly suspicious that aesthetic paternalism would be a legitimate state policy. Hurka's arguments however do not prevent the state from adopting aesthetic paternalism. On the contrary, his theory actually supports it.

5.4.2 The active character of perfectionism cannot rule out aesthetic paternalism. Achievements must, the argument goes, be one's own doing if they are to contribute to one's excellence. The argument rules out forcing a person into a particular activity,

---

[25] For the U.K. see Martin Cloonan, *Banned! Censorship of popular music in Britain 1967-92.* Aldershot, Arena, 1996.

[26] See 2.4.3 and Craig A. Lockard, *Dance of Life: Popular Music and Politics in Southeast Asia.* Honolulu, University of Hawai'i Press, 1998, pp. 28-9, 35-6, 77, 96-101, 248. For more information and articles on the censorship of pop music see *Index on Censorship*, issue 6 of 1998.

but allows prohibiting one option while leaving others open. Aesthetic paternalism leaves open all other available options, allowing people ample opportunity to form goals and realise them in the world themselves. It is true that aesthetic paternalism makes it impossible to achieve, without interference, the goal of avoiding a particular genre of bad music. But many people, those who would end up listening to bad music in the absence of aesthetic paternalism, would not achieve this goal anyway. And even if they could achieve some perfectionist value despite not achieving the goal so that aesthetic paternalism may involve a perfectionist loss in this respect, this loss may be outweighed by a gain in other respects. Instead of listening to bad music the person may choose some more worthwhile activity that involves more pervasive and structured goals, or that are more complex, intricate, and challenging. The active character of perfectionism therefore does not rule out aesthetic paternalism.

A possible reply here is that since aesthetic paternalism eliminates at least two possible goals (namely, listening to bad music and avoiding listening to bad music without the help of the state) and creates no new goals then there is an overall perfectionist loss since there are fewer opportunities to form and achieve a goal. However, this reply misunderstands the nature of Hurka's account of perfection. One's perfection is advanced when one actually forms a goal and achieves it oneself. The presence or absence of opportunities for perfection do not themselves contribute towards it. Of course if opportunities are slim then there will be less perfection. But it is possible for certain opportunities to be eliminated without any perfectionist loss so long as people choose some other activity. And this is almost certain to occur if there are an adequate range of other opportunities available. As long as people can choose and engage in activities themselves, the active character of perfection is preserved.

The same reply may be made to another suggestion concerning the effects of aesthetic paternalism on the possibility of an active life. According to this suggestion, the

presence of bad music will lead people to participate in activities that have perfectionist value. Some may join or start up bad music fan clubs, getting involved in their organisation, and thereby have new goals to pursue. The production and sale of bad music records may help provide employment and new opportunities for many. Others may be inspired by the music to pursue a career as popular singers. However, if bad music is prohibited, people will engage in other activities, get involved in other organisations, find employment in other areas and pursue other careers. The would-be bad music purchasers will either spend their money elsewhere thereby creating employment or save it which will lead to greater investment and thereby stimulate the economy, creating new opportunities. I am not claiming that everyone who would be employed in the bad music industry will find employment when aesthetic paternalism is implemented. But the effects on the economy would be so diverse and widespread that it would be difficult to say one way or the other whether aesthetic paternalism would result in more or less opportunities for achievement in the pursuit of goals and possibilities for an active life. There is little reason to suppose that allowing the presence of bad music would mean a greater amount of opportunities for achievement than its absence.

The proposition that perfection requires that activities be accompanied by the right intentions does not, even if true, help bad music listeners. Hurka claims that because perfection is inner, it is not open to public inspection and cannot be enforced by the state. But the aim of aesthetic paternalism is not to directly bring about desirable intentions. Rather, it restricts conduct and leaves people free to pursue other options, hoping that they do so with the right intentions. Hurka also claims that acts done with the wrong intentions have no value and that choosing simply to avoid punishment counts as a wrong intention. Whilst this does mean that perfection is not advanced when a person chooses not to listen to bad music only because of its prohibition by aesthetic paternalism, there is still plenty of opportunity for perfectionist improvement since people may pursue the remaining options left open with the right intentions.

The claim that activities must be accompanied by the right intentions to have value, if true, rules out forcing someone into a particular activity but not prohibiting one option while leaving open many others.

It may be thought that aesthetic paternalism is unacceptable on perfectionist grounds because people may listen to bad music with the right intentions, thus giving that activity some value. However, it is unlikely that listening to bad music may be accompanied with intentions that, on Hurka's account, have value. Activities must be intended to affect many significant aspects of one's own or others' lives or intended to form part of a complex hierarchy of goals.[27] It is difficult to see how buying or listening to bad music should be accompanied by either of these intentions. Most would listen to it as a time-filler, without it playing any significant part in their life at all.

5.4.3 Hurka's remaining arguments may be rejected for reasons that parallel those for rejecting the same arguments as they applied to lifestyle. It is difficult to see how prohibiting bad music could limit people's routes to excellence. Excellence is only achieved if choosing an activity involves critical deliberation, or is the result of a fully conscious choice concerning its and the alternatives' merits, or is intended to affect many aspects of one's own of others' lives, or is intended to form part of a complex hierarchy of activities. Listening to bad music would hardly ever satisfy any of these conditions. It seems to be an exception to Hurka's general rule that 'almost any activity can, if approached in the right spirit, embody some perfection'[28] because it is difficult to see how it could be approached in the right spirit. In those few occasions where it would involve some excellence, such as when a person critically reflects upon the merits and demerits of bad music, it is unlikely that there would be no alternative option which would enable him to achieve as much or more excellence.

---

[27] See footnote 16 and accompanying text .

[28] Hurka, *Perfectionism*, p. 154.

Prohibiting bad music does not deny people something of genuine perfectionist worth that cannot be had elsewhere.

The next two arguments may be dealt with simultaneously since they both make the general claim that intervention will result in the substitution of bad motives (fear of punishment) for good (choosing as part of an organising plan for one's whole life or choosing an activity for itself rather than as a means to an end). Hurka claims that if people act as the law requires just to avoid punishment and if they would otherwise have acted in the same way for better reasons, then there has been a perfectionist loss.[29] However, this argument only rules out paternalism toward those who would abstain from listening to bad music in the absence of aesthetic paternalism for worthwhile reasons and would no longer do so for those reasons once aesthetic paternalism is in operation. It does not show that paternalism toward those who would listen to bad music if permitted would be unjustified. And it does not rule it out toward those who, due to the threats imposed by aesthetic paternalism, refrain from listening to it but at the same time also do so for other more worthwhile reasons of the type that Hurka stresses. Moreover, in the latter situation, Hurka's theory would seem to support aesthetic paternalism since by providing additional reasons for avoiding bad music, it increases the complexity of their hierarchy of goals.

The next argument also fails. Prohibiting the sale, distribution, and broadcast of bad music would not diminish people's energy and self-direction and thereby undermine the general character needed for excellence. People's energy and self-direction would be the same in a society without bad music, or perhaps greater since they would avoid hearing such inane music which may have a pacifying effect.

Hurka's arguments for liberty fail to rule out aesthetic paternalism. Moreover, it is not difficult to see how aesthetic paternalism would be supported by Hurka's theory. As

---

[29] Hurka, *Perfectionism*, p. 154.

we have seen, listening to bad music cannot plausibly be thought of as a route to excellence on Hurka's grounds. Listening to other, more complex and sophisticated, forms of music would. Listening to bad music would not normally involve any critical deliberation, conscious choice, or be intended to form part of a structured hierarchy of pervasive goals. Eliminating it as an option would make these things more likely to occur, since the menu of options people face would then contain a higher proportion of routes to excellence. Moreover, Hurka claims that activities are valuable when they are organised around a single end or are complex, intricate, and challenging.[30] Listening to bad music would normally fit none of these criteria. Eliminating it would again increase the likelihood of an activity that does meet them, such as listening to more sophisticated forms of music, being chosen. Thus perfection, on Hurka's account, is likely to be advanced by aesthetic paternalism.

5.5 Relationship Paternalism

5.5.1 With *relationship paternalism*, the state aims to promote and support healthy adult relationships. It believes that such relationships are part of the good life and tries to ensure that they are maintained. To help achieve this end, the state employs relationship advisors whose tasks are to regularly call upon couples, whether married or not, and attempt to identify and offer help with relationship problems. The state makes it illegal for couples not to meet the advisors.[31]

As was noted in chapter two, Hurka's perfectionism holds that close relationships are valuable since they create common goals for a partnership and require coordinated acts to achieve them. A perfectionist therefore has reason to adopt relationship paternalism. However, I take it that such a policy would be unjustified. State officials should not try to forcibly tell people how to live their lives in this manner. Hurka's

---

[30] Hurka, *Perfectionism*, pp. 121, 123.

[31] See 2.4.4 for more on the feasibility of such a scheme, and the suggestion that the British government is considering a similar, though less objectionable, set of policies.

arguments, however, fail to provide any reason for why the state should not implement that policy.

5.5.2 The leading of an active life is the value most threatened by relationship paternalism. There is the danger that the relationship advisors would be controlling the relationship rather than the parties to it themselves. The advisors may take all the challenge out of a relationship, in the same way that an aspiring writer's goal of being accepted for publication is denied if a friend pays a publisher to accept the writer's book, or that flattening mountains to ensure success for mountaineers eliminates the challenge of the activity.[32] But this is not an inevitable result of relationship paternalism. It must be remembered that it is not required that the advice be followed. People must still make the final decisions themselves. Moreover, it could be required of the advisors that they refrain from conducting themselves in such a way that their advisees are denied any real opportunity to control their own relationship. Rather than, say, giving a full and thorough analysis of the situation so that there seems only one rational course of action that the couple can take, the advisor could simply clarify the issues, and provide some guidance by pointing them in the best general direction. By conducting themselves in this way, advisors will avoid diminishing the active life and bringing about a perfectionist loss.

The idea that perfection requires acting from certain inner states does not work against relationship paternalism. This policy does not force people into activities when they do not have the appropriate inner states. Couples are not forced into, nor forced to remain in, relationships by the state. Moreover, the advisors can help to bring about the desirable inner states. By explaining to couples how their relationship involves a complex hierarchy of pervasive goals that unites both of them in a 'single intended and accomplished state' they can help ensure that they have valuable inner states by viewing their relationship in a way that leads to greater perfectionist worth.

---

[32] The latter example is from Raz, 'Duties of Well-Being', p. 14.

5.5.3 Hurka's argument that paternalism may limit people's routes to excellence does not help resist relationship paternalism. Such paternalism does block the option of couples determining their relationship completely uninfluenced by an advisor, which is an option that could have perfectionist worth. However, this loss must be weighed against the perfectionist gain of relationship paternalism. That is, one must also keep in mind the likelihood that the relationship advisors will help to bring about longer-lasting and more valuable relationships that would not occur in the advisors' absence. Consider only those couples who would normally fail to maintain a healthy relationship. In their case the route to excellence that relationship paternalism blocks, that of conducting their relationship without the help of government advisors, is already unavailable since by hypothesis their relationships would not succeed. It is possible that advisors may help rescue and improve these relationships which would otherwise fail. Hurka's theory would therefore support the policy with regard to these couples. Even if there are perfectionist costs to others, Hurka assumes without argument that the balance of perfections would come out against relationship paternalism.

Relationship paternalism will not involve the kind of harm that Hurka claims will occur when restrictive laws make people's intentions worse by substituting fear of punishment for the more worthwhile motives of choosing as part of an organising plan over an entire life and choosing something for itself rather than only as a means to an end. This harm will occur only if couples would freely choose, in the absence of the policy, to hold sessions with relationship advisors for those worthwhile reasons and fail to have the same motivations when relationship paternalism is implemented. But there is little reason to think that many couples would seek relationship advice on their own. Most prefer to conduct their relationships privately.[33] Even amongst those

[33] Empirical evidence for this claim includes the fact that while over 150,000 divorces take place in the United Kingdom each year (Home Office, *Supporting Families: A Consultation Document*, London, HMSO 1998, p. 31) only just over 5,000 people have attended individual information meetings prior to

who would voluntarily use consultancy, it is not clear why their motivations for doing so, assuming they are worthwhile, would be supplanted by fear of punishment once relationship paternalism is introduced. Couples could approve of the scheme of relationship advice for the worthwhile reasons that Hurka emphasises, and at the same time be motivated by the threat of punishment. The less worthwhile motive may simply complement rather than supplant the worthy motives.

A possible objection is that the worsening of intentions will occur with regard to the couple's attitude to the relationship itself, rather than their attitude toward the advisor. If so then there is potentially a highly significant perfectionist loss. But being forced to spend time with a relationship advisor need not make their attitudes towards the relationship worse. It will do so if the advisors give bad, damaging advice. But we may assume that the advisors are competent. Then, as has been seen, the advisor may help improve their attitudes. Moreover the couple may be brought closer together, not as a result of the advice received, but due to a shared resentment of the government's intrusion into their lives.

Hurka's final argument also fails to block relationship paternalism. As long as the relationship advisors conduct themselves appropriately, they will not diminish couples' energy or self-direction or undermine the general character needed for excellence. Of course this could be the result if advisors attempt to take over and control the relationship completely. But providing considerate, friendly advice that is respectful of the couples' need for control over their own relationship need not have this consequence.

A state devoted to Hurka's perfectionism has a reason for implementing relationship paternalism due to the value that the theory places on adult relationships. None of

---

starting divorce or separation proceedings since 1996 (*Supporting Families*, p. 35) even though they are legally required to do so by the Family Law Act 1996.

Hurka's arguments for liberty are able to show why the state should not implement it. On the contrary, they seem to provide further support for relationship paternalism since it will facilitate more valuable intentions and motivations, giving them a more complex hierarchy of pervasive goals and uniting couples in a single intended state.

5.6 Conclusion

Hurka's arguments fail to rule out lifestyle, aesthetic, and relationship paternalism and some aspects of his theory support these policies. One possible route here is to hold that, despite Hurka's understanding of perfectionism, the activities that these types of paternalism aim to discourage are not in fact bad activities. Perhaps being a couch potato, listening to bad music, or being in relationships such as those that relationship paternalism tried to correct are not relatively worthless things to do. But this is implausible. The activities are of little worth and any plausible perfectionist theory must say that they are.

Another way that Hurka's theory could avoid supporting paternalism is for it to abandon its agent-neutral structure in favour of an agent-relative perfectionism. According to the former each person should care about the perfection of all as well as his own. According to the latter, each person must pursue just his own perfection.[34] If Hurka's perfectionism is agent-relative then the state has no reason to promote perfection. Perfectionism, on that view, simply tells each person to seek her own perfection.[35] If the state has no reason to promote perfection, then Hurka's theory does not face the objection that is supports lifestyle, aesthetic, and relationship paternalism. But as Hurka argues, viewing perfectionism as agent-relative is implausible. It is difficult to see how someone could think perfection is good for herself without also

---

[34] Hurka, *Perfectionism*, p. 62.

[35] This is not to say that an agent-relative perfectionism has no implications for state policy. If the state has a duty to help people fulfill their duties then it may have a duty to provide what is necessary for perfection. But then the state would not be promoting perfection for perfection's sake. Here I set aside such complications.

thinking it good for others.[36] Thus perfectionism must be agent-neutral. We still, then, need to know whether there is a way of avoiding the counter-intuitive implications of Hurka's theory.

Another possibility is that perfectionism is one part of a pluralist morality which includes other ideals and principles such as utility and rights.[37] If so, and if these different elements of morality are to be weighed against each other, then the unattractive consequences could be avoided if the counter-intuitive requirements of perfectionism are outweighed by the requirements of some other element of morality. The ideal of state neutrality can be understood as one possible version of this pluralist view. A neutralist could admit that morality consists of perfectionist and non-perfectionist considerations but hold that the former have no weight in political morality whilst the latter do. Neutrality gives a weighting of zero to perfectionism in politics. It says that perfectionist ideals, such as those advocated by Hurka, should be excluded as a legitimate goal of state action.

A slightly different version of this pluralist view is to hold that perfectionist considerations have zero weight not in all of political morality but only in the paternalistic sphere of it. Perfectionist considerations are to be ignored, on this view, when the state interferes with people for their own good, but those considerations may have weight both in other state policies and outside politics. According to this view perfectionist considerations are not appropriate ones for state paternalism. Hurka's perfectionist theory, rather than being extended into politics, should be bracketed or set aside when thinking about what the state ought to paternalistically do. Instead of grounding state paternalistic action on Hurka's perfectionism, the state should base such intervention, if there is to be any, on non-perfectionist considerations. This is not to deny the truth of Hurka's theory. He could be correct about the nature of

---

[36] Hurka, *Perfectionism*, p. 63.

[37] Hurka, *Perfectionism*, p. 6.

perfectionism and its (non-political) requirements. We would simply be denying his claim that the theory should be extended to paternalism. People may have duties to promote others' perfection. But they just must not do so paternalistically through the state. This seems a plausible position since we would not object to someone who tried to get couch potatoes to live more worthwhile lives, or tried to encourage others not to listen to bad music, or gave others advice on their relationships.[38] It is when the state tries to pursue these aims that we object.

This view, one that prohibits perfectionism in the motivation of paternalistic policies whilst not making perfectionism impermissible elsewhere in political morality, is the one that I suggest should be accepted. The arguments in this and the previous chapters have sought to establish the negative case for the view, that perfectionism ought not to motivate state paternalism. The positive case, that the exclusion of perfectionist considerations will result in a more plausible position on state paternalism, is the task of the next chapter.

---

[38] We may object to some methods used to achieve these objectives, such as coercion or manipulation, even when they do not involve the state. This would mean either that the exclusion of perfectionist ideals extends beyond political morality to other areas of morality or that those ideals, while not excluded from those areas, are sometimes outweighed by other considerations (that may not apply when the ideals are implemented through the state).

Chapter Six

Neutral Paternalism

6.1 Introduction

The arguments in previous chapters have attempted to show that perfectionists are unable to demonstrate convincingly how they can avoid the paternalistic implications of their theories. This difficulty does not mean that perfectionism should be rejected. But it does support rejecting it as the basis for *some* state actions, that is, paternalistic ones. The state could be neutral and set aside conceptions of the good when it acts paternalistically, while not necessarily abandoning perfectionism in its non-paternalistic actions. Perfectionism may be justified in some spheres of state action even if it is not in cases of state paternalism. Showing that perfectionist arguments are unacceptable in their implications for paternalism does not show that they are unacceptable elsewhere. In a similar fashion, some theorists hold that neutrality is a requirement with respect to constitutional essentials and matters of basic justice but that perfectionism may legitimately be resorted to outside these spheres.[1] Conceivably, we could have a view that requires neutrality for paternalistic policies, but permits perfectionism elsewhere.

This chapter tries to strengthen this conclusion by showing how there can be a plausible neutral position on state paternalism. I try to show that a neutral principle of paternalism avoids the problem faced by perfectionism of supporting too much paternalism. I consider the neutral principles of paternalism that are proposed by John

---

[1] See for example John Rawls, *Political Liberalism*, New York, Columbia University Press, 1993, pp. 214-5 and Brian Barry, *Justice as Impartiality*, Oxford, Clarendon Press, 1995, pp. 143-4.

Rawls and David Richards and try to show that whilst they rule out perfectionist cases of paternalism, they still permit too much. I then present a different principle which I argue avoids this objection.

This strategy may be thought a flawed one since my concern is to defend the claim that the neutrality constraint sets a necessary condition for justified paternalism. Accepting the neutrality constraint for state paternalism means that for paternalism to be acceptable it must not appeal to perfectionist considerations. But that does not mean that *any* instances of paternalism that are not perfectionist are acceptable. The neutrality constraint sets a test for justified paternalism. Any cases of paternalism that fail that test are unjustified. But passing the test does not mean that a case of paternalism is justified. We need a set of sufficient conditions to tell us that. Showing that some neutral principle of paternalism permits too much paternalism just shows that neutrality is not sufficient for justified paternalism, that is, that there must be some other conditions, besides neutrality, that must be met before paternalism can be said to be permissible. But that is no argument against neutrality since it is no argument against a necessary condition that it is not sufficient.

An alternative strategy would be to try to demonstrate that neutral paternalism, rather than supporting too much paternalism, supports too little. Limiting state paternalism to only cases that comply with neutrality, it could be argued, results in ruling out types of paternalism that should not be ruled out. A necessary condition cannot support too much but it can support too little. The difficulty with this strategy is that it is not obvious that there are any cases of justified paternalism which must be accepted. Some seem more reasonable than others but there are none that we are so confident about that it would seriously weaken a theory if it said that those cases of paternalism are always unacceptable. The intuition that some cases of paternalism are clearly

unacceptable is stronger than the intuition that some cases of paternalism are clearly justified.[2]

There is however reason for why finding neutral principles of paternalism that do not support too much paternalism helps to establish that neutrality is required for justified paternalism. If neutrality rules out some cases of paternalism but not others, then that raises the possibility that there is some other necessary condition that can rule out those others as well as the cases that neutrality rules out. If there is such a condition, then neutrality may not be necessary. The other condition could do all the work that I argue neutrality does. So showing that the neutrality condition can rule out as many cases of unacceptable paternalism as possible and does so in a plausible manner strengthens the claim that neutrality is a necessary condition for justified paternalism.[3]

Moreover, the more work that neutrality can do as a necessary condition for justified paternalism, the more significant a conclusion it is. We want not just a necessary condition, but one that is able to do as much work as possible in ruling out paternalism where it should be ruled out. Another necessary condition for paternalism and therefore for justified paternalism is that there be a paternalist. If there is no paternalist, there can be no justified paternalism. But that condition does not do much work in ruling out unjustified paternalism. Hence, if it can be shown that neutrality

---

[2] The strategy that neutrality rules out too much justified paternalism is pursued by Richard Kraut in 'Politics, Neutrality, and the Good', *Social Philosophy and Policy*, vol. 16, no. 1, 1999. But his argument is unconvincing for the reason just given: none of his cases strike me as obviously justified interferences. The only one that may be acceptable, perfectionist education of children, can be set aside here since the issue I focus on is paternalism towards adults. For arguments that the education of children should respect neutrality, see Matthew Clayton, *Educating Liberals; an argument about neutrality, equality of opportunity, and parental autonomy*, D.Phil. thesis, University of Oxford, 1997.

[3] The argument is not complete since even if neutrality can rule out all the cases of unacceptable paternalism, there may still exist another condition that can also do the same. If so then given one, the other would not be necessary. Instead we would have to say that a necessary condition for justified paternalism is that it be neutral or satisfy X where X is the other condition. In the next chapter I argue against several proposals for X. If it can be eliminated, then we are left with neutrality as a necessary condition.

can rule out almost all cases of wrongful paternalism, then we will have a necessary condition for justified paternalism that is a condition of some significance.

In the next section I set out a method for determining neutral principles of paternalism. Sections 6.3 and 6.4 examine the principles proposed by John Rawls and David Richards respectively. Section 6.5 presents my own proposal which I argue avoids the difficulties faced by Rawls's and Richards's principles.

6.2 Rawlsian Contractualism

Neutrality itself does not entail any particular principles of paternalism. It simply limits the range of arguments that may be appealed to to justify state paternalism, ruling out policies that are based on a conception of the good. Further arguments are needed to show what neutrality with regard to state paternalism would look like. These arguments must be compatible with neutrality, although they might not follow from the idea alone. The method that I will use is the contractualist one set out by John Rawls. According to Rawlsian contractualism, principles of justice are the outcome of a hypothetical social contract. We are to imagine that members of society, conceived as rational persons concerned to advance their interests and mutually disinterested towards others' interests, are to agree to principles for the regulation of the basic structure of society. The parties to the contract are placed behind a 'veil of ignorance'[4] which disguises from them their own particular circumstances so that they are unable to choose principles tailored to their own particular advantage. The veil hides from them their class position, social status, talents, abilities, strength, intelligence, and conception of the good.[5]

---

[4] John Rawls, *A Theory of Justice*, Oxford, Oxford University Press, 1972, p. 12 and section 24.

[5] Rawls, *A Theory of Justice*, pp. 12, 18-19, 137. Although Rawls does not explicitly say so, the veil must also presumably hide one's age, race, and sex. He suggests this on pp. 149-50.

To determine which principles will be the outcome of this procedure, we are to work out which principles everyone would accept as most likely to advance their interests, given that they are unaware of their own particular circumstances. The parties are to look at the consequences of each principle for people in various social positions and choose in light of the possibility that they may turn out to be in those social positions.[6] So for example, when choosing principles for the distribution of income and wealth, the parties in the original position are to look at the consequences of the alternative principles for those in worse off and better off positions under those principles and decide which of the principles is best for them given that they do not know which of these positions they will inhabit.

The primary concern of Rawlsian contractualism is with distributive justice and hence may be thought to be an inappropriate tool for dealing with paternalism. It is not obvious that a theory designed mainly to set out principles for the distribution of benefits and burdens in society should also be used to determine what principles should regulate state intervention in people's lives for their own good. But Rawls suggests the possibility that contractualism 'can be extended to the choice of more or less an entire ethical system."[7] If "justice as fairness" (the name Rawls gives to his theory) succeeds reasonably well, claims Rawls, the next step is to examine 'the more general view suggested by the name "rightness as fairness."'[8] Determining principles of paternalism that are the outcome of Rawlsian contractualism may be viewed as part of this project.

Moreover, Rawlsian contractualism is well-suited to finding principles of paternalism due to both the *subject* and the *role* of the theory. Its subject is the basic structure of society, made up of major social institutions, such as the political constitution and the

---

[6] The parties know that it is possible that they will end up in any social position, but they do not know the probabilities of so ending up. See Rawls, *A Theory of Justice*, p. 155.

[7] Rawls, *A Theory of Justice*, p. 17.

[8] Rawls, *A Theory of Justice*, p. 17.

major economic and social arrangements, which distribute fundamental rights and duties and determine the division of advantages from social cooperation.[9] The basic structure is the main subject of Rawlsian contractualism because of its influence on people's life-prospects, on what will happen to them and on how well they will do. The effects of the basic structure are 'profound and present from the start [of people's lives].'[10] Paternalism is part of the subject of Rawlsian contractualism. The basic structure of society includes institutions and arrangements that are potentially paternalistic. It may contain rules that limit people's options and these limits may be intended for their own good or it may, if it is extremely libertarian, contain no such rules. And paternalistic institutions and arrangements, or their absence, have profound effects on people's lives. A major reason why issues such as euthanasia, drug taking, homosexuality, gambling, prostitution, commercial surrogacy, and health and safety regulations are often controversial is their significance for the lives of the people involved.

The *role* of Rawlsian contractualism is to provide a method for resolving disagreement over questions of political morality. It provides a method for finding principles of justice needed in the face of the conflict of interests that exists over how the benefits and burdens of social cooperation are to be distributed.[11] In his more recent work, Rawls has emphasised the conflict, not only of interests, but also of different religious, philosophical, and moral doctrines.[12] In the face of possible conflict between the recommendations of different doctrines, Rawlsian contractualism provides a procedure for resolving disagreements. These disagreements could include disagreements over paternalism. Different moral doctrines will recommend possibly conflicting practices of paternalism. In response to such conflict, principles are needed that will set out what rights there are to be free from paternalistic interference

---

[9] Rawls, *A Theory of Justice*, p. 7.
[10] Rawls, *A Theory of Justice*, p. 7.
[11] Rawls, *A Theory of Justice*, p. 4.
[12] See for instance Rawls, *Political Liberalism*, pp. xvi, 3-4.

by others along with correlative duties of non-interference, and rights and duties, if there are any, to intervene paternalistically with others. Like principles of justice, principles of paternalism define 'a common point of view from which [people's] claims may be adjudicated'[13], claims regarding the rightness and wrongness of paternalistic intervention.

Rawlsian contractualism purports to be neutral and therefore offers a method for discovering neutral principles of paternalism. Some critics, however, have questioned the neutrality of the theory. Their objection can take several forms. Some claim that the description of the original position, the situation in which the contractors choose principles, smuggles in a conception of the good and that the resulting principles are thereby perfectionist rather than neutral. This objection is unconvincing since the features of the original position, such as the veil of ignorance, describe conditions needed for fairness in the agreement procedure rather than conceptions of valuable ways of life. However there is not space here to argue this in full, so I will simply set aside the objection.[14] Another form of the objection that Rawlsian contractualism fails to be neutral concedes the neutrality of the premises of the argument but questions whether the outcome of the agreement need be neutral. The contractors, it is argued, could agree to perfectionist principles. Since they are motivated to advance their own interests and since also it is in one's interest to have a true conception of the good and engage in activities that are truly worthwhile, then they will want to ensure that the basic structure of society helps them to secure this interest. Perfectionist policies would do this. This conclusion may perhaps be avoided by assuming, as Rawls does, that the parties to the contract have a desire to pursue their conceptions of the good, whatever they are.[15] But this desire is a misguided one. People do not have an interest

---

[13] Rawls, *A Theory of Justice*, p. 5.

[14] For arguments against this objection, see Will Kymlicka, 'Liberal Individualism and Liberal Neutrality', *Ethics*, vol. 99, no. 4, 1989.

[15] Rawls, *A Theory of Justice*, pp. 128, 250, 327. Even this would not rule out all perfectionism. Some perfectionist policies would be compatible with permitting people complete freedom to pursue their conception of the good.

in pursuing a misguided conception of the good in the mistaken belief that it is worthwhile. That would probably make them worse off. Their interest lies in leading truly worthwhile lives, not those that they merely believe are worthwhile and they want to pursue their conceptions of the good only insofar as they are worthwhile.

This objection is a strong one and shows that Rawlsian contractualism does not, without additional premises, result in the rejection of perfectionism.[16] But the difficulty may be avoided by simply assuming that the contractors must observe the neutrality constraint. All perfectionist principles are ruled out of consideration; the contractors must choose only amongst those principles that do not appeal to a conception of the good and that therefore pass the test of neutrality. This may seem like cheating. It is if we want an argument for neutrality. But that is not my purpose. We want to know, given the difficulties that we have seen perfectionism faces, whether neutrality fares any better. So we can take neutrality for granted and see whether a plausible position on paternalism is compatible with it. In order to do this we may use Rawlsian contractualism, with the added assumption of neutrality at the start. This will allow us to determine plausible neutral principles of paternalism (if there are any). The contractors' desires to pursue their conceptions of the good whatever they are, although a misdescription of people's motivations in the real world, is a way of modeling neutrality in the original position.

6.3 Rawls on Paternalism

It may be objected that since the outcome of the contract is the result of a neutral procedure, then that outcome is itself neutral. Whatever results from the original position complies with neutrality. But this would be a mistake. Principles based upon a conception of the good are, by definition, not neutral. They remain non-neutral even if they are the outcome of a procedure that does not itself assume that (or some other) conception of the good.

[16] The same difficulty applies to other forms of contractualism, such as the type developed by T. M. Scanlon and Brian Barry. See Simon Clarke, 'Contractarianism, Liberal Neutrality and Epistemology', *Political Studies*, vol. 47, no. 4, 1999.

6.3.1 A good place to start is with what Rawls himself says about paternalism. He argues that one of the principles that would be agreed to in the original position is a principle of equal basic liberties. According to this principle, each person is to have an equal right to the most extensive total system of equal basic liberties compatible with a similar system of liberty for all.[17] These liberties include freedom of speech and assembly, liberty of conscience and freedom of thought, and freedom of the person.[18] But the equal basic liberties principle is chosen only on the assumption that in society people are rational and able to manage their own affairs. Parties in the original position will, however, agree to exceptions; principles of paternalism that apply in the event that this assumption does not hold.[19] Rawls sets out three principles of paternalism.

The first is that there must be an 'evident failure or absence of reason and will'[20] on the part of the paternalised. There are several types of cases where this occurs. One is where 'their powers are undeveloped and they cannot rationally advance their interests, as in the case of children.'[21] Another is where 'through some misfortune or accident they are unable to make decisions for their own good, as in the case of those seriously injured or mentally disturbed.'[22] A third type of case concerns the 'irrational inclinations'[23] of normal adults. Since the parties in the original position will want protection against harming themselves in these types of cases, they will agree to a principle allowing intervention.

The second principle insists that paternalistic intervention be 'guided by the principles of justice and what is known about the subject's more permanent aims and

---

[17] Rawls, *A Theory of Justice*, p. 250.
[18] Rawls, *A Theory of Justice*, p. 61.
[19] Rawls, *A Theory of Justice*, p. 248.
[20] Rawls, *A Theory of Justice*, p. 250.
[21] Rawls, *A Theory of Justice*, p. 249.
[22] Rawls, *A Theory of Justice*, p. 249.
[23] Rawls, *A Theory of Justice*, p. 249.

preferences, or by the account of primary goods.'[24] It is puzzling what Rawls means by 'guided by the principles of justice,' since the principles of paternalism concern exceptions to his first principle of justice, the equal basic liberties principle. He does not explain this part of the principle more fully but I will assume that it places two sorts of constraints on paternalism. The first is that it places constraints on paternalism insofar as it affects those other that the target(s). So if acting paternalistically towards *x*, whose reason has failed, would also deprive *y* and *z*, who are not irrational in any way, of equal basic liberties, then the paternalism is not permitted. Secondly, the condition sets limits on how the paternalist may treat *x*; intervention may justifiably invade those basic liberties of *x*'s that must be restricted for the paternalism to be effective, but must not invade other of *x*'s liberties or anything else of *x*'s that are guaranteed by the principles of justice.

The principle also insists that intervention must be guided by the subject's permanent aims and preferences or as Rawls says elsewhere by the individual's 'own settled preferences and interests insofar as they are not irrational'.[25] I take it that at least part of what Rawls means by this is that paternalism must be guided by the subject's own conception of the good and not one that she does not accept. This seems implied by Rawls's statement that '[p]aternalistic principles ... must not be interpreted to license assaults on one's convictions and character'.[26] It also follows from his argument for the principle from the original position. Although the parties do not know their conceptions of the good, they wish to advance them whatever they turn out to be. They want to 'guarantee the integrity of their person and their final ends and beliefs whatever these are.'[27] In order to do so, they will require that paternalism be guided by their own conception of the good.

---

[24] Rawls, *A Theory of Justice*, p. 250.
[25] Rawls, *A Theory of Justice*, p. 249.
[26] Rawls, *A Theory of Justice*, p. 250.
[27] Rawls, *A Theory of Justice*, p. 250. See also pp. 128, 206, 327 for expressions of this interest of the contractors in the original position.

In the absence of knowledge of the paternalised's conception of the good, the paternalist is to be guided by the account of primary goods. Primary goods are goods that all rational people want whatever else they want. Regardless of a person's conception of the good, she would prefer more primary goods to less. They allow her to advance her interests whatever they are.[28] When discussing distributive justice, Rawls's concern is with primary *social* goods such as rights and liberties, opportunities and powers, and income and wealth.[29] This condition of paternalism, though, is concerned with primary goods in general, *natural* (such as health and vigour, intelligence and imagination[30]) as well as social. When there is no knowledge of the paternalised's preferences and aims, paternalism guided by the account of primary goods will try 'to get for [the subject] the things he presumably wants whatever else he wants.'[31]

The third principle of paternalism is that the subject of paternalism must, once her rational powers have been developed or recovered, accept that the intervention on her behalf was for her own good.[32] This condition alone, says Rawls, cannot be sufficient since we can imagine cases where people are forced against their wishes to affirm new religious beliefs and where they may come to sincerely accept the new beliefs. Such intervention would nevertheless not be permitted.[33] But whilst not sufficient, the condition is necessary.

---

[28] Rawls, *A Theory of Justice*, p. 92.
[29] Rawls, *A Theory of Justice*, p. 92.
[30] Rawls, *A Theory of Justice*, p. 62.
[31] Rawls, *A Theory of Justice*, p. 249.
[32] Rawls, *A Theory of Justice*, p. 249. Rawls's statement that she accepts that it was 'the best thing' for her is too strong. Many interventions may be reasonable and warranted even if there was some alternative that would have been even better. Perhaps there is some threshold that must be reached before the intervention is justified. I set aside this problem here.
[33] Rawls, *A Theory of Justice*, pp. 249-50.

The second principle ensures that paternalism will be neutral. The first and third cannot do this. Even if paternalism is limited to cases involving the failure or absence of reason and will, it could nevertheless be guided by a conception of the good. And subsequent assent could be forthcoming from someone who has been the subject of paternalism guided by perfectionist values. The second principle ensures the neutrality of paternalism in three ways, corresponding to its three parts.[34] First, since the principles of justice are neutral (or so Rawls argues), paternalism guided by them will be neutral. Second, requiring that paternalism be guided by the subject's own conception of the good rules out paternalism guided by a conception of the good that is not held by the subject. This may seem to allow the possibility of perfectionist paternalism so long as the subject shares the conception of the good that motivates the paternalism. However, I take Rawls's stipulation to mean that paternalism must be guided by the person's own conception of the good because it is his or her own view and not because it is true, worthwhile, etc. It would thus not be a case of perfectionism. Paternalism must be guided by the target's own view even if it is a false one, something that a perfectionist would not recommend. Third, since primary goods are neutral, being things that everyone wants regardless of their conception of the good, paternalism that is guided by them will be neutral. Since paternalism must, according to Rawls, be guided by primary goods in the absence of knowledge of a person's conception of the good, there is no space for perfectionist paternalism.

Rawls does not explicitly state the status of the principles, but I take him to mean that all three of the principles are necessary conditions for justified permissible paternalism and that together they are sufficient. That the first two are intended as necessary is indicated by Rawls's comment that two 'stipulations are necessary' immediately preceding their presentation. That the third is necessary is indicated when Rawls claims it is a 'requirement'. That they are together sufficient is implied

---

[34] To reiterate, these are that paternalism be guided by (1) the principles of justice, (2) the subject's conception of the good, and (3) the account of primary goods.

by Rawls's statement that they are 'principles stipulating when others are authorized to act in [people's] behalf' and by the claim that the third condition alone 'is not, however, by any means sufficient' immediately prior to reiterating the first two principles.[35]

Some general points should be noted even though I will not investigate them further. When the conditions are met, intervention is permissible and, claims Rawls, 'sometimes required'[36]. He does not state what determines when the obligation exists as contrasted with mere permission but I will focus on the principles as determining permissibility. Rawls also repeatedly mentions when 'others' are entitled to paternalistically intervene, which seems to imply that the principles are intended to apply not only to the state. However, I will only focus on their implications for state action. Finally, the principles set out when the liberties defined by the equal basic liberties principle may be restricted for a person's own good. They may, therefore, not be the same principles that apply to paternalistic interventions other than those that limit basic liberties. Perhaps the conditions required for justified paternalism in these other cases are weaker. However, since most serious forms of paternalism restrict at least one of the basic liberties, it seems reasonable to focus only on these cases.

6.3.2 The claim that the above principles of paternalism would be chosen by individuals in the original position is a plausible one. Agreement on the first principle is straightforward. In cases where there is no failure or absence of reason and will, people are able to rationally manage their own affairs and thus will not need any assistance from others. So the contractors will limit intervention only to cases where they cannot advance their own good themselves. The second principle will be agreed to only on the assumptions of neutrality, modelled in the original position by a

---

[35] All quotations from Rawls, *A Theory of Justice*, pp. 249-50.

[36] Rawls, *A Theory of Justice*, p. 249.

contractor's desire to pursue his conception of the good whatever it is. Without this assumption, parties in the original position trying to advance their own interests could sanction perfectionist principles of paternalism. Since, by stipulation, the parties wish to pursue their conception of the good, they will require that paternalistic intervention be guided by their conception of the good. And since they are constrained by neutrality, in the absence of knowledge about their conception of the good they will require that paternalism be guided by only neutral considerations.[37] The third principle is a further protection against perfectionist paternalism; a person forced into a way of life contrary to his own conception of the good is unlikely to accept afterwards that that was the best thing for him. The principle is not strictly necessary if the first two principles are respected and nor is it sufficient since it is possible that a person will come to accept a way of life that he has been forced into, but is does provide an additional safeguard against the type of paternalism that perfectionism recommends. Since the contractors are bound, by hypothesis, by neutrality, they will endorse the third principle.

An adjustment to Rawls's principles is necessary. Those principles, according to Joel Feinberg, warrant intervention with cases of weakness of the will. This is a disposition to act in a way one knows at the time to be unreasonable, perhaps for example when lighting another cigarette.[38] Intervention in such cases would be supported by Rawls's principles since they involve failure of the will, intervention could be guided by the person's own conception of the good, and subsequent assent could be forthcoming. State paternalism, in some such cases at least, would be unacceptable even though supported by Rawls's principles. But Feinberg fails to appreciate that Rawls's first principle requires that there be a failure of *reason and*

---

[37] Without the neutrality constraint this part of the second principle would not, for the reasons set out in 6.2, be agreed to. Since it is in people's interest to conform to what is good, parties in the original position could agree that, in the absence of knowledge about a person's conception of the good, paternalism be guided by true perfectionist considerations.

[38] Feinberg, *Harm to Self*, p. 109.

will, that is, a failure of reason as well as a failure of will and not merely the latter. Weakness of the will does not necessarily involve a failure of reason. A weak-willed person may reason correctly but be unable to conform his will to his reasoning. Where he does not reason incorrectly, intervention is not permitted by Rawls's principles.

But by making the permissibility of paternalism turn upon the failure of both reason and will, Rawls's principles allow too little paternalism. Requiring failure or absence of reason *and* will would be too strong. Young children often have a failure or absence of reason but often do not lack will. The same is true for a person trying to cross a bridge unaware of its danger. But paternalism in these cases seems nevertheless permissible. I therefore suggest that the 'and will' part of Rawls's first condition be dropped so that paternalism may proceed only when there is an evident failure or absence of reason.[39]

Rawls's principles of paternalism can account for the unacceptability of much of the paternalism that perfectionism supports. People who are led by a misguided conception of the good in their behaviour in some sense suffer from a failure of reason. Since they fail to realise what is truly good, they are fail to perceive right reason and thereby satisfy the first condition of justified permissible paternalism. But intervention to guide them to the good would violate the second condition. It would not be guided by their own conception of the good or by primary goods. Forcing someone out of his couch potato lifestyle (which he does not regret living) would be contrary to the paternalised's own conception of the good. Censoring bad music for the good of those who would otherwise listen to it involves intervention that is motivated by a conception of the good. A Jehovah's Witness's refusal of a blood

---

[39] Some people may think that the resulting principle is too narrow, that it should be amended to cover failure of reason *or* will, since intervention should sometimes be acceptable in cases of weakness of the will without any failure of reason. Unfortunately there is not space here to investigate this question, or that of whether weakness of the will is even possible.

transfusion may constitute a failure of reason but forcing one upon him would violate his own understanding of his good as conforming to his interpretation of the teachings of Jehovah. Intervention with a mountain climber who may not fully appreciate the risks involved would fail to meet the second condition, so long as embarking on risky mountain climbing even when there is no full appreciation of the risks is part of his conception of the good. Neutrality, embodied in Rawls's second principle, blocks these cases of paternalism that are driven by perfectionist considerations.

An initial difficulty with Rawls's set of principles is that they sanction interventions to prevent trivial harms. I may fail to reason correctly when I prefer to watch television in the evening rather than attend a local community meeting. The latter may be the more fulfilling thing for me to do. I have a failure of reason of sorts when about to drink a glass of wine without knowing that, although not poisoned, it is an inferior quality of wine than what I believe it to be. A similar failure of reason occurs when I unknowingly purchase bruised fruit in the supermarket.[40] Interference in these and other cases would be supported by Rawls's principles since the intervention would be guided by the paternalised's own views, and, we may assume, subsequent assent would be forthcoming. Intervention in such cases, however, seems troubling. Part of the concern is whether the interventions would be effective. But it still seems troubling even if we could be sure that a paternalist could force me away from the television, and to avoid inferior wine and bruised fruit without causing me any other harm.

The above cases all involve only seemingly trivial harms. One response here could be to add a further condition, which states that paternalism must prevent non-trivial harms (and provide non-trivial goods) only. This however would go too far.

---

[40] Perhaps this is to take the notion of failure of reason too broadly. The person does not reason badly but merely lacks knowledge. Assume then, that evidence of the mistakes people are making is available to them, but that they fail to take that evidence sufficiently into account in their deliberations. The cases thereby become ones of failure of reason.

Intervention to prevent trivial harms is sometimes permissible so long as the form of intervention is appropriate. Legislating for fruit quality standards in supermarkets seems unobjectionable even if having state officials standing by and seizing bad fruit from customers' hands does not. What is needed is a principle that sets out what sorts of interferences are appropriate in different circumstances. A plausible such principle would probably make the type of intervention turn partly on the significance of the harm threatened. Legislated standards to prevent bad fruit seem permissible whilst direct intervention is less so. But direct intervention is permissible when a person is about to eat fruit containing lethal poison. Similarly, knocking a glass of poisoned wine out of a person's hand is justified, but knocking a glass of inferior wine out is less so. Legislating labeling requirements that may indicate the qualities or otherwise of the wine seems unobjectionable. A principle stating what forms of intervention can take is needed before Rawls's principles are acceptable. Without going into the details of it, I assume some such principle is available. It could be stated in a very general form, such as that the intrusiveness of the intervention must be proportional to the significance of the harm involved.

Even if corrected in the above way, Rawls's principles of paternalism may still seem to sanction too much intervention. Feinberg argues that the first principle in particular is overly inclusive. Much of the conduct of competent adults involves a failure or absence of reason. Protecting people against their 'irrational inclinations' by imposing 'penalties that may give them a sufficient motive to avoid foolish actions' and 'undo the unfortunate consequences of their imprudent behaviour'[41] allows paternalistic intervention with many decisions that seemingly should be left unhindered. Some decisions, important and trivial, could be accompanied by a degree of irrationality, foolishness, or imprudence without warranting intervention. Feinberg does not seem to consider the possibility that Rawls's other two principles may block

[41] Rawls, *A Theory of Justice*, p. 249.

intervention sanctioned by the first. But I will argue that even considered alone, the first principle does not support as much paternalism as Feinberg thinks.

Whether it does depends upon the conception of rationality that is appealed to, since this will determine how wide the scope of conduct that may be interfered with is. According to Feinberg, on Rawls's account any decision made by an actual person that deprives her of primary goods is irrational because rational parties in the original position would not make such decisions. He claims that 'Rawls's hypothetical "rational parties" ... would not make such choices [to give up primary goods], so those choices are clearly not "rational," in the sense he favours, when made by actual persons.'[42] If so, then all decisions to forgo primary goods, because "irrational," are therefore subject to intervention on Rawls's first principle.

This objection, however, is mistaken. Feinberg is correct to say that it would be a mistake to think that decisions to forgo primary goods are necessarily irrational for actual persons, even if they would be irrational for parties in the original position. For example, it may not be irrational for someone to sign away all his income and wealth when he becomes an adherent of a particular religion. If that is a condition of adhering to the religion, and if so adhering is what he has most reason to do then the decision is rational. It would however be an irrational decision for someone in the original position to take since he does not know what his religious views are. What is irrational from the point of view of the original position need not be so for actual people. But there is no reason why a more reasonable conception of rationality cannot be accepted by people in the original position. They need not think that actual people must be taken to be acting irrationally when they act in ways that would be irrational were they behind a veil of ignorance. They can accept, from their position behind the veil of ignorance, a principle warranting intervention in the event of irrationality, but

---

[42] Joel Feinberg, *Harm to Self (volume three of The Moral Limits of the Criminal Law)* New York and Oxford, Oxford University Press, 1986, p. 185.

on the condition that what is irrational is not to be defined as what would be irrational from their position behind the veil of ignorance. If a more reasonable conception of rationality is taken, the scope of 'irrationalities' that Feinberg takes Rawls's first principle to cover is not as wide as Feinberg makes out. Hence in this respect the principle does not warrant as much intervention as Feinberg seems to think.

Not only need we not accept the conception of rationality that Feinberg maintains is part of Rawls's theory, but Rawls himself does not hold that conception. Rawls sets out several principles of rationality. These principles are formal ones such as (when other things are equal) that given a desired end we are to adopt the means that realises it in the best way, that we should prefer an activity if it achieves all the aims and more of another activity, and that of two alternatives that achieve the same ends we should prefer that which has the greater likelihood of achieving those ends.[43] Other criteria of rationality are that there are no errors of calculation or reasoning, facts are correctly assessed, and all relevant facts about the situation and consequences of a decision are known.[44] These are the criteria of rationality to use in interpreting Rawls's first principles of paternalism, and not what would be rational from the original position. Only when these conditions of rationality are not satisfied, irrationality occurs and paternalism is premitted by Rawls's first principle.

What seems to lead Feinberg into the mistake is that whilst he claims, rightly, that Rawls wants to protect people from irrational choices that are at the expense of primary goods[45], he takes this to mean that all decisions that are at the expense of primary goods are irrational. He claims that 'Rawls would ... protect people from "irrational choices" in the ... sense of choices at the expense of ... "primary goods".'[46] What is more plausible, and what Rawls presumably means, is that of the choices that

---

[43] Rawls, *A Theory of Justice*, pp. 411-2.
[44] Rawls, *A Theory of Justice*, p. 417.
[45] Feinberg, *Harm to Self*, p. 184.
[46] Feinberg, *Harm to Self*, p. 184.

lessen primary goods, some of them are irrational; ones that people should be protected against (subject to the constraints of the other two principles).

Feinberg also argues to the effect that even on this more reasonable conception of rationality, the first principle allows too much. He claims that Rawls fails to distinguish 'irrationality in the strongest sense (incompetence)' from 'irrationality in a weaker sense (chronic foolishness, unreasonableness, or imprudence).'[47] It is only the former and not the latter, says Feinberg, which warrant intervention. He provides several cases that he maintains fall in the latter category but should be free from paternalism. His main examples concern departures from rationality that 'are made by genuine preference, taste, or conviction.'[48] Illustrations include impulsive adventurers and gamblers, and those living lives of spontaneity, impulse, and risk.[49] Such ways of life may not be deliberately chosen by the people involved but may 'reflect their characters in deep and important ways.'[50]

The final sentence above indicates that the types of cases Feinberg has in mind would fail to pass the test for permissible intervention of Rawls's second principle even if they comply with the first. Combining the first two principles gives the result that irrationalities that are part of the person's conception of the good must not be interfered with. Feinberg's examples would not pass this test. They are irrationalities that are integral to people's conceptions of their own good. Paternalism is unacceptable when it uproots a person's 'irrational trait that is prominent in his self-conception or even in his ideal of himself.'[51] These concerns are accounted for by Rawls's principles.

---

[47] Feinberg, *Harm to Self*, p. 185. John Kleinig seems to endorse the objection in *Paternalism*, Manchester, Manchester University Press, 1983, p. 66.
[48] Feinberg, *Harm to Self*, p. 109.
[49] Feinberg, *Harm to Self*, p. 109.
[50] Feinberg, *Harm to Self*, p. 109.
[51] Feinberg, *Harm to Self*, p. 110, quoting Richard Arneson, 'Mill versus Paternalism', *Ethics*, vol. 90, no. 4, 1980, p. 474.

Feinberg's examples, moreover, may not necessarily involve any irrationalities, so they would fail Rawls's first principle of permissibility as well as the second. A life of spontaneity, impulse, etc. may be a good one to lead, say because it involves less psychological frustration that might accompany carefully considering various available options and may result in engaging in activities that are just as good. Or someone may find that the options engaged in are better ones when they do not rationally consider their merits than when they attempt to choose rationally, due to a lack of expertise in this matter. An impulsive gambler for example may win more when he is impulsive than when careful. Or even if he does not, the enjoyment he gets may make it worthwhile. None of Rawls's criteria of rationality are violated by such behaviour. Rawls admits as much when he claims that the extent to which rational deliberation should be engaged in is itself subject to rational decision. There is nothing irrational, he says, 'in an aversion to deliberation itself provided that one is prepared to accept the consequences.'[52]

There are, however, cases that illustrate impermissible paternalism supported by Rawls's principles. Imagine someone about to enter into marriage in the false belief that it will be the blissful union that he has desired all his life. It is clear however that the marriage is unsuitable and will end in failure. His conduct involves a failure of reason since, due to his strong desire for wedded bliss, he fails to perceive the indications that the marriage will not be successful. Rawls's first principle is therefore satisfied. Prevention of the marriage would be in line with the subject's own conception of the good, a component of which is the desire to be married happily ever after. Rawls's second criterion is therefore also satisfied. And the third is too since it could be the case that the subject's subsequent assent will be forthcoming once he realises the mistake he almost made. The fourth criterion that was proposed, concerning the proportionality of the intrusiveness of the intervention with the

[52] Rawls, *A Theory of Justice*, p. 418.

significance of the harm involved, could also be satisfied. Since the harm here is not a trivial one, the intrusiveness of the intervention need not be limited as that condition requires. Similar cases could be imagined concerning decisions such as choice of occupation, where to live, what sort of family to have and many other matters. All could involve decisions made in the presence of some failure of reason where intervention could be in keeping with the person's view of the good life and be subsequently approved of by the paternalised. The scope of paternalism that is supported by Rawls's principles is wider than he thinks.[53]

One initial reply concerning these cases is that whilst Rawls's principles may in theory support intervention with the decisions described, the circumstances will never occur in practice since others, including the state, are never in a position to know better than the person about to make the decision whether it is the right one. Even if this is true much of the time, there are surely some cases where others can know better than the person involved and paternalism could be effective. And even in those cases where others cannot know better, we may want to know why it is wrong in principle, rather than merely in practice, to intervene. As I have tried to show, Rawls's principles do not provide such an explanation.

6.4 Richards on Paternalism

Rawls's principles therefore seem to support too much paternalism. David Richards has proposed an alternative principle of paternalism which he argues would be agreed to from behind the veil of ignorance. The principle takes the following form:

---

[53] The reason why Feinberg does not raise these straightforward cases is probably because he is searching for cases of voluntary irrationality and the above examples are not voluntary according to his conception of voluntariness. Since he claims that paternalism is permitted when the conduct is not fully voluntary (Feinberg, *Harm to Self*, p. 12), the above cases also work against Feinberg's view. To take account of them, he could either change his notion of voluntariness or make the criteria of involuntariness necessary but not sufficient for permissible paternalism. In the next chapter, I argue that he should take the second course.

> If a person x either lacks the developed capacity of rational choice and deliberation or the full opportunity to exercise such capacities, which lack of either or both is likely substantially to frustrate the person's interests, where such frustration is extremely irrational, given the persons' desires, competences, and circumstances, then another person y, who is available and can stop x, at no great cost to himself, is to do so, and, if also at little cost to himself, to help him develop his capacities of rationality, if possible, and aid him in securing the full opportunity to exercise these capacities, if feasible.[54]

This principle is presented as a sufficient condition for justified paternalism but Richards understands it as also necessary since he suggests that all other instances of paternalistic interference are forbidden. The contractors will agree to this principle since they know general facts about people, in particular the possibility that they will act irrationally or be non-rational and will want to ensure that their interests are protected in these circumstances. The principle of paternalism 'if publicly known and generally acted on, will substantially advance human interests'.[55] The contractors will not permit paternalism in circumstances other than the ones described by the principle, since people's interests are better advanced by letting them pursue their interests alone and letting them learn from their mistakes.[56]

Richards stresses that this principle must be taken to include two important constraints on the permissibility of paternalism. The first is the constraint of neutrality. Paternalism must not be guided by perfectionist considerations. The notion of irrationality must not be taken to permit interference in order to guide the paternalised into a more worthwhile conception of the good. Instead, irrationality 'must be defined relative to the agent's system of ends.... Accordingly, only those acts are irrational that frustrate the agent's own system of ends, whatever those ends are.'[57] The second

---

[54] David Richards, *A Theory of Reasons for Action*, Oxford, Clarendon Press, 1971, pp. 192-3.
[55] Richards, *A Theory*, p. 192.
[56] Richards, *A Theory*, p. 193.
[57] David Richards, 'Commercial Sex and the Rights of the Person', *University of Pennsylvania Law Review*, vol. 127, no. 5, 1979, p. 1263.

constraint is that paternalism is permissible only when the irrationality is severe and systematic and a serious, permanent impairment of interests is threatened.[58]

Like Rawls's principles, Richards's principle of paternalism is able to rule out the objectionable cases of paternalism that perfectionism supports. The constraint of neutrality ensures this. Paternalism guided by a conception of the good that is not accepted by the paternalised is made impermissible by Richards's principle. The second of Richards's constraints works to rule out intervention in cases of trivial harm that I argued were permitted by Rawls's principles (but that could be avoided by adding an appropriate condition). But Richards's principle is subject to a similar objection made against Rawls's principles. Richards's principle seems to support a range of paternalistic practices that a plausible account of anti-paternalism should rule out. People sometimes make mistaken choices about their own good, choices that are irrational given their own ends, and interference could prevent the significant harm that would occur as a result of these mistakes. For example, as was suggested in the previous section, someone who longs for a perfect marriage may believe that a particular partner is the perfect one for him. However, he fails to see what is clear to everyone else, that the union is doomed from the start. In this case he makes a mistaken decision which will significantly frustrate his interests, and it is irrational from within his own conception of the good.

Perhaps, however, in this case and in other cases of choices that are irrational given the person's own conception of the good, the condition of Richards's principle that the target must lack 'the developed capacity of rational choice and deliberation or the full opportunity to exercise such capacities' is not met. In many such cases, including the one of the person contemplating marriage, people could be said to have the capacity of rationality and to not be denied the opportunity to exercise it, but simply fail to arrive

---

[58] Richards, 'Commercial Sex', p. 1264.

at the correct decision in the exercise of their capacities. Others, such as children and the insane on the other hand, clearly do meet the condition.

However, this condition seems an arbitrary one, especially from the point of view of the contractors in the original position. They seek to advance their own interests and since irrational decisions made by people even with the capacity for rational choice and full opportunity to exercise this capacity can significantly set back their interests, the contractors will want to protect themselves against these harms as well as any that may occur when they lack the capacity of rationality such as children and the insane. They do not have reason to confine paternalism to cases where the capacity or opportunity is absent. Intervention seems more acceptable where the capacity or opportunity is lacking because it is in those circumstances that people are likely to make a mistake in choosing for themselves. But if so, then why rule out intervention also where people make mistakes even though they do not lack the capacity or opportunity? The importance of being able to learn from one's own mistakes, which Richards mentions, can explain why the contractors will not permit intervention with some irrational decisions. But it is not of so much importance that it can explain why the contractors would prohibit interference with all the irrational decisions of rational adults.[59]

6.5 A Revised Principle

Richards's principle therefore faces the same problem that faced Rawls's principles. Whilst both accounts rule out a significant range of paternalistic practices (those that are supported by perfectionist considerations) they fail to rule out another type of paternalism: interventions that, rather than trying to guide her towards a worthwhile

---

[59] This may be why Richards, in his later work, does not limit paternalism to those who lack rationality or the opportunity for its exercise. In 'Commercial Sex' he claims that the contractors would 'agree on an insurance principle against certain of [the] more serious irrationalities in the event they might occur to them.' (p. 1263.) That is, paternalism is not limited to children, the insane, and others who lack rationality.

conception of the good, attempt to force her to live up to her own conception of the good. Much of this latter type of paternalism seems unacceptable. As was argued in section one, neutralists such as Rawls and Richards do not strictly need to give sufficient conditions for justified paternalism. A necessary condition, which is what neutrality purports to be, need not rule out all unjustified paternalism. But the more that it can rule out, the more useful and significant it is.

One way of blocking some paternalistic policies justified from within a person's own conception of the good is not available to the neutralist. It could be suggested that intervention in these types of cases is unacceptable on the grounds of some conception of the good, such as that it is worthwhile for people to make these decisions autonomously for themselves. But such grounds violate the neutrality constraint by appealing to a conception of the good. They may not be appealed to in a neutralist approach to paternalism. If we are to appeal to perfectionist considerations to try to rule out the implications of neutral paternalism, we would be compelled to allow all perfectionist considerations to count in political morality, including those that motivate paternalism. We cannot allow just the ones we want while ignoring the rest. But then we would face all the difficulties that were described in previous chapters. Moreover, as I have argued in previous chapters, perfectionist appeals to autonomy, endorsement, and intention do not succeed in ruling out all unacceptable paternalism.

Another possible move is also problematic. One minimal condition for justified paternalism is that it should be successful, that is, that it should advance the paternalised's good. Paternalism is not justified if it makes the person worse off or no better off.[60] But getting a person to live up to his conception of the good will not be for his good if his conception is a false one. That would probably make him worse off. So paternalism that is guided by the target's own conception of the good where

[60] There may be a few cases where paternalism fails to benefit the paternalised but was nevertheless not wrong because it was only due to circumstances that were impossible to foresee that the paternalised was not benefited.

that conception is a false one would not be supported by any set of neutral principles so long as that set includes what any plausible set should include, namely that paternalism, to be justified, must advance its target's good.

But this reply is inadequate for two reasons. Firstly it does not rule out all instances of the type of paternalism that is problematic. If a person's conception of the good is not a false one then getting him to live up to it would be for his good and therefore pass the minimal test for justified paternalism. For example, belief in the value of having a perfect marriage partner is not necessarily a false belief. It may be part of a conception of the good that is not misguided, and interfering with someone's plans to marry a bad partner would truly be for his own good. But state interference of this type seems unacceptable. It is not just interferences that make people conform to their *false* conceptions of the good that are problematically supported by Rawls's and Richards's principles, but also interferences that make people live up to their conceptions of the good even if they are *true* conceptions.

The second problem with the suggested reply is that it also violates the neutrality constraint. If the reply is accepted then paternalism that is guided by the target's own conception of the good is impermissible when that conception is a false one, but it is permissible when the conception is a true one. Hence before the state decides whether to act paternalistically, it must determine whether its target's conception of the good is true or false. If it is true, the state may proceed (so long as other necessary conditions are met). If it is false, then the state must not. The state would therefore be acting on the basis of the truth or falsity of the person's conception of the good, which is impermissible according to the neutrality constraint.

A better solution to the objection that neutral principles support too much paternalism is to abandon Rawls's and Richards's principles and instead accept the following Principle Of Neutral Paternalism (PNP):

State paternalism is justified only if (1) there is a failure or absence of reason; and (2) intervention is not guided by any conception of the good (whether that conception is held by the paternalised or not).

Clause 1 should be understood in such a way as I argued that Rawls's principles should be understood in order to avoid Feinberg's objections: the irrationalities covered must not include all choices that would be irrational from behind the veil of ignorance; and interference is not permitted with irrationalities that are part of a person's conception of the good. As well as ruling out paternalism that is guided by perfectionist considerations, the second condition rules out forcing people to conform to their own conception of the good. Forcing someone who is committed to the ideal of a romantic marriage not to marry a less than ideal partner is forbidden by the principle.

PNP sets two necessary conditions for justified paternalism. Many other conditions must be met for sufficiency. Paternalism must be successful. That is, in aiming to protect or benefit the paternalised, the paternalist must succeed in actually doing so. An incompetent paternalist, even with the best intentions, acts wrongly if he makes his target worse off. As was argued in section three, an additional principle is that the intrusiveness of the intervention must be proportional to the significance of the harm involved. Moreover, sometimes it is better to let people make mistakes and learn from them. Only when these benefits of non-interference are clearly outweighed by the harm that will occur without intervention may the paternalist proceed. Other paternalistic policies may be too administratively cumbersome to be worth implementing. Two other conditions that I claimed applied to Rawls's principles are also needed here. If acting paternalistically towards x, whose reason has failed, would also deprive others, who are not irrational in any way, of what justice requires or significantly harms them, then the paternalism may not be permitted. Secondly, intervention may justifiably invade those basic liberties of x's that must be restricted for the paternalism to be effective, but must not invade other of x's liberties or

anything else of x's that are guaranteed by the principles of justice. When these other conditions are all met, then PNP becomes one that sets sufficient conditions for justified state paternalism.

PNP does not rule out all paternalism. It permits intervention that provides neutral goods, that is, goods that are compatible with a wide range of conceptions of the good. Examples of such paternalism include health and safety legislation. Compulsory hygiene regulations for food manufacturers, for example, restrict the freedom of consumers to buy unregulated food (which would probably sell at a cheaper price). Insofar as such regulations are intended for consumers' own good, they are paternalistic. They are also neutral, since few conceptions would require not avoiding contaminated food. If some people do hold conceptions of the good that involve using unregulated food, then the neutral state would have to try to find some way of exempting them from the regulations. Similarly, other cases of paternalism that would be permitted by PNP are traffic regulations, food labelling requirements, vehicle and machinery safety regulations, and laws against misleading advertising.

It may be thought that PNP still covers too much. Since health is a general good, does it sanction forcing people into healthy activities (so long as there is also a failure or absence of reason)? If so, it supports compulsory transfusions for Jehovah's Witnesses, laws against junk food, risky leisure activities, and other unwarranted interventions. However, people's conceptions of the good include rankings of primary goods. Forcing a particular ranking of goods, such as placing health above religious beliefs or above the excitement gained from engaging in risky activities, would be to impose a conception of the good. Such intervention is ruled out by the second condition of PNP. According to the principle, the state may act paternalistically to provide general goods only when doing so does not override the paternalised's ranking of goods according to his conception of the good. When it does not, such as

preventing a person from crossing a bridge which, unknown to him, is dangerous, or preventing someone drinking unknowingly poisoned wine, interference is permitted.

PNP may seem to ignore an important insight that Rawls's and Richards's proposals both capture. According to them, paternalism ought to be guided by the paternalised's own values rather than those of the paternalist. When deciding whether to give a Jehovah's Witness a blood transfusion, we ought to respect his beliefs even though we may think them mistaken. But PNP may seem problematic in that it requires that we ignore his religious view in deciding what to do. However, whilst the statement of the principle may make no mention of the view of the paternalised, part of the motivation for it is the desire to respect a person's own conception of the good. By limiting state paternalism to the provision of neutral goods, PNP allows people to pursue their own conceptions of the good in their own way. Giving a blood transfusion to a Jehovah's Witness on the basis of the perceived falsity of his religious views is forbidden by the principle. My point here is not just that PNP gives a tolerant result - that Jehovah's Witnesses ought not to have blood transfusions forcibly imposed on them. Like Rawls's and Richards's principles, although less obviously than theirs', my proposal also gives the result for a plausible reason - that a person's decisions ought to be respected even though guided by a mistaken conception of the good.

Another difficulty of PNP is whether it would be agreed to by the parties in the original position. Since the parties are motivated to pursue their conception of the good whatever it is, it seems that they would sanction measures that would guide them into conforming with their own conception of the good. PNP rules out (and is designed to avoid) such measures. The solution, I suggest, is to drop the assumption that the parties wish to pursue their conception of the good whatever it is. That assumption misdescribes people in the real world - they wish to pursue their conception of the good only insofar as it is not misguided or false - and it leads to too much paternalism when made one of the premises of the original position. Instead it

should be assumed that the contractors are simply motivated to advance their own interests but not their own or any conception of the good. They would agree to condition 1 of PNP since when they are not suffering from a failure or absence of reason, they can advance their own interests without paternalistic interference from others. The contractors would also agree to condition 2 since they do not wish to advance any conception of the good. Agreeing to that condition nevertheless allows them to protect their interests by providing themselves with neutral goods when they are unable to do so themselves due to failure or absence of reason. But they will not endorse further paternalistic measures to force them to live up to their conceptions of the good because they do not, under my proposed adjustment to the original position, wish to live up to their conception of the good whatever it is.[61]

6.6 Conclusion

Paternalism, according to the principle I have suggested, must be confined to cases where interference provides goods that are compatible with a wide range of conceptions of the good. The principle is plausible in its implications and would be chosen by parties to a Rawlsian hypothetical social contract, so long as some plausible adjustments are made to the circumstances of the contractors' original position.

---

[61] Whether this adjustment would affect the argument for principles of justice or whether it is acceptable to have different descriptions of the original position when determining principles of paternalism than when determining principles of justice, are questions that are beyond the scope of my thesis.

Chapter Seven

Some Alternatives to Neutrality

7.1 Introduction

The argument of the first five chapters was that since perfectionism supports too much paternalism, neutrality should be accepted as a restraint on the state's paternalistic actions. That conclusion was strengthened in the previous chapter by setting out how respecting the neutrality constraint may result in a plausible position on paternalism. This chapter tries to strengthen the conclusion in a second way. I consider a number of principles which provide alternative explanations for why state paternalism should be limited. Like the neutrality constraint, these principles also provide necessary conditions that must be met before paternalism may legitimately proceed. Unlike the arguments considered in chapters 2-5, these principles are not themselves perfectionist considerations. They are not claims that paternalism reduces or fails to advance well-being. It may be thought that accepting any of these principles does nothing to damage the case for neutrality. That there may be other conditions that must be met for justified paternalism need not imply that neutrality must not also be respected. There can be more than one necessary condition for justified paternalism. However, these alternative principles are important because if any of them can block a significant range of cases of objectionable paternalism that are supported by perfectionism then the case for the neutrality constraint is weakened. If they succeed, neutrality may be unnecessary as an explanation for the unacceptability of much paternalism.

The principles I examine are that concern for well-being must be balanced against the importance of liberty, that the threat of harm must be certain for intervention to be justified, that a person's consent must be obtained before paternalism may proceed, and that paternalism is justified only with conduct that is substantially nonvoluntary. In the sections below, I argue that these views are all either in various ways problematic or insofar as they are plausible, they need to be seen as complements of neutrality rather than alternatives to it.

7.2 The Balancing View

One possible explanation concerning the justifiability of paternalism is *the balancing view* which its main defender claims is 'the natural view in commonsense morality.'[1] According to this view, there are two values at stake in paternalism. The first is the value of liberty[2], the value of people choosing and pursuing those choices without interference. The other value is individual well-being, which paternalistic interference aims to protect or promote. The balancing view does not see liberty as part of well-being. If it did it would be similar to the views, criticised in chapter 2, that make autonomy a component of well-being. In the balancing view, liberty and well-being are both fundamental values. The two values coincide when a person freely chooses that which is for his own good, but they conflict when a person chooses contrary to his own good. In these conflict cases, the person's good could be furthered by paternalistic intervention that overrides his choice for his own good, or his liberty could be respected at the cost of his well-being. According to the balancing view, if the threat to well-being is sufficiently weighty then paternalism would be justified. If, on the other hand, the restriction of liberty would be significant and the gain in well-being relatively trivial, then paternalism would be unjustified. The balancing view

---

[1] Dan W. Brock, 'Paternalism and Autonomy', *Ethics*, vol. 98, no. 3, 1988, p. 551.

[2] Brock describes this as the value of autonomy or self-determination. I use 'liberty' to avoid the theoretical baggage of the other terms and because the way it is explained by Brock is more consistent with liberty than the more complex notions of autonomy and self-determination.

requires weighing up the liberty and well-being at stake in any given case in order to determine whether paternalism would be justified.

The balancing view seems to supply a plausible explanation in many cases of possible paternalism. Preventing a person from unknowingly crossing a dangerous bridge restricts her liberty in only a trivial way while benefiting her greatly. Similar balancing is appropriate when a person is about to jump off a building in a temporary but intense fit of depression. The importance of protecting well-being in these cases outweighs the value of liberty. On the other hand, people often seem to make mistaken choices that lessen their well-being but where intervention would be inappropriate. People would be better off, for instance, by choosing leisure activities more worthwhile than watching television. Smokers would usually be better off not smoking. Readers of pornography could spend their time more rewardingly reading great novels. In such cases the value of liberty outweighs our concern for their well-being. If perfectionism were to embrace the balancing view, it could hold that the good gained from forcing people into worthwhile conceptions of the good is sometimes outweighed by the value of liberty. Perfectionism could thereby avoid much of its paternalistic implications, making neutrality less attractive as an alternative.

The balancing view, however, may support too much paternalism. In many cases people make decisions that significantly set back their well-being and the restriction of liberty required to prevent this would be trivial on some straightforward measures of the importance of liberty. People sometimes choose to sacrifice their well-being for others or for a cause. They also sometimes mistakenly make decisions concerning far-reaching matters such as where to live, who to live with, and what occupation to pursue, thereby setting back their well-being. Engaging in risky activities such as mountain climbing, hang gliding, and racing car driving threatens well-being. In all these cases, the freedom to make these choices is not clearly important enough to

outweigh the threat to well-being when other less costly, more suitable, or safer and no less rewarding and worthwhile options could be taken up instead.

That conclusion of course depends upon the method used for weighing liberty against well-being. Giving liberty great weight will avoid the implications in the cases above. An obvious difficulty with the balancing view is that it requires some non question begging method for determining the appropriate balance between liberty and well-being when they conflict. We cannot, without begging the question, give liberty great weight merely in order to avoid the conclusions in the previous paragraph. Brock does not provide a method for how the balancing is to proceed.[3] But he argues that in many cases where it would seem that the value of well-being outweighs that of liberty, there is no real conflict between liberty and well-being. Although the people involved may be made better off in one respect if they are prevented from making certain choices, they are worse off overall by being prevented from engaging in freely chosen goals that are important to them. Being able to engage in such goals contributes to well-being.[4] A concern for well-being therefore requires refraining from paternalistic intervention rather than providing a motivation for it. Suppose Brother Francis believes it is immoral for non-human animals to be used in lethal experiments and so volunteers as a subject in an experiment in order to save non-human animals that were to be used.[5] In doing so, Brother Francis risks significant harm to himself. It may be thought that on the balancing view, the risk may be great enough to outweigh the restriction of liberty involved when preventing Brother Francis from volunteering. But Brock argues that Brother Francis is not worse off

---

[3] He does suggest that the importance of liberty turns partly upon 'how central and far-reaching the choice is within the particular individual's plan of life, how strongly the individual wants to make the choice in question for him- or herself ,and so forth' (p. 551), but these considerations seem to conflate the value of liberty to that of well-being.

[4] Brock, 'Paternalism and Autonomy', p. 555.

[5] Brock, 'Paternalism and Autonomy', pp. 554-5. Brock attributes the example to Donald VanDeVeer, *Paternalistic Intervention*, Princeton, New Jersey, Princeton University Press, 1986 (no page number given). In order to make the case one of paternalism, one must set aside a concern for the welfare of non-human animals.

since freely and knowingly choosing self-sacrifice in order to forward a goal that he considers is of great importance advances rather than sets back his well-being.[6]

The problem with this reply of Brock's is that if what he claims is true, then it is impossible for a person to engage in a freely chosen goal that sets back his or her well-being. No matter how harmful the activity, how painful it is, or how much it seems contrary to the person's interests, he or she is actually better off through having freely chosen and engaged in the activity. Brock's reply describes people's interests in a way that is completely at odds with the way that they themselves see them. Brother Francis would not view volunteering to save non-human animals as in his overall interests. He would say that it is *not* in his interests, but that he feels morally compelled to do it anyway. Whilst helping others and pursuing moral goals *may* advance one's own well-being, Brock's argument makes it always the case. But certain goals are worth pursuing for moral reasons that are not accompanied by or reducible to self-interested reasons. Sometimes people freely choose activities and thereby become no better or even worse off. Such decisions may be made for moral reasons, but they need not be. A person may freely choose self-mutilation. He may be better off as a result. Perhaps the mutilation results in a large material reward. Or perhaps he makes the decision in the light of a carefully worked out philosophy which he is devoted to in all other aspects of his life, and such devotion may lead to a rewarding life. But say neither such factors apply. He engages in severe self-mutilation for some cheap thrills and fleeting fame. He also may agree that the decision is against his interests. In this case, it seems perverse to say that the freely chosen activity advances the person's well-being. It is another case of a person freely choosing an activity that makes him worse off.

Brock's attempt to show that the balancing view does not support too much paternalism is therefore unpersuasive. People sometimes can and do freely choose

---

[6] Brock, 'Paternalism and Autonomy', p. 555.

activities that make them worse off, and weighing their liberty to do so against their well-being may support paternalism in many cases where it would be objectionable. Hence we should reject the balancing view, at least as it is interpreted by Brock, as an explanation of the justifiability of paternalism.

But the view still retains some plausibility. When thinking about paternalism we do seem to have to weigh the values of liberty and well-being. I believe its plausibility can be accounted for by the view I am defending. When interference to correct people's choices is motivated by perfectionist considerations, then liberty should be respected at the cost of less well-being. The value of people's freedom to make their own decisions, even though they will be worse off than they would be with interference to guide them into more worthwhile conceptions of the good, outweighs the loss in well-being. But when, and only when, well-being is threatened in some neutral sense, that is, when intervention can be guided by goods that are compatible with a wide range of conceptions of the good, then a concern for well-being may (but does not necessarily) outweigh the value of liberty. In other words, restricting liberty for a person's own good is permissible only when the intervention is guided by neutral considerations. Neutrality is a necessary condition for justified paternalism.

Hence the neutrality constraint provides an interpretation of the balancing view. Moreover, the neutrality constraint seems to be supported by many of the comments Brock makes when he is defending the balancing view. He claims that his interpretation of the balancing view helps explain what any plausible theory should, namely that what is most objectionable about paternalism is that it 'involves the substitution by the paternalistic interferer of his or her conception of what is good for another for that other's own conception of his or her good.'[7] Brock believes that taking the balancing view in conjunction with a subjectivist conception of well-being explains what is wrong with such substitution. But the neutrality constraint also of

[7] Brock, 'Paternalism and Autonomy', p. 559. See also p. 560.

course provides such an explanation: such substitution violates neutrality. Neutrality also seems to be the implication in Brock's claim that his position 'is one important support for a political liberalism that declines to found political institutions on any particular conception of the good life.'[8] Nothing in the balancing view commits someone who accepts it to also accepting neutrality. The balancing view alone does not rule out including perfectionist considerations on the well-being side of the equation. If Brock wishes to support neutrality, then he is better off embracing a principle making neutrality a necessary condition of justified paternalism along with the balancing view. Brock's concerns are catered for by the neutrality constraint and the balancing view is most plausible if interpreted in light of that constraint.

7.3 The Consent Principle

In this section, I examine Vinit Haksar's arguments for his claim that paternalism is justified only if the paternalised person has consented to the interference. Haksar is a perfectionist and claims that an acceptable stance on paternalism must appeal to perfectionist considerations. But he also admits that permitting the state to base its decisions on the inferiority or superiority of forms of life leaves perfectionism open to the charge that it supports too much paternalism. To avoid this difficulty he appeals to the consent principle. Before considering that principle, I first examine another that Haksar considers and rejects. Although he is right to reject it, he does so for mistaken reasons.

7.3.1 The Certainty Principle

Haksar considers Mill's suggestion that interference is justified only when the danger is certain. If the risk of harm is only probable then the person should only be warned of the danger rather than forcibly prevented from taking the risk.[9] As Mill puts it,

---

[8] Brock, 'Paternalism and Autonomy', p. 561.

[9] Vinit Haksar, *Equality, Liberty, and Perfectionism*, Oxford, Oxford University Press, 1979, p. 237.

> when there is not a certainty, but only a danger of mischief, ... [the person ought] only to be warned of the danger; not forcibly prevented from exposing himself to it.[10]

The certainty principle, as I shall call it, permits interference with a person attempting to cross a bridge that he does not know is unsafe, so long as we are certain of the danger. Haksar claims that this principle raises the difficulty of determining when a danger is certain or not, something over which there could be much disagreement. But his main argument against it is that the principle cannot 'be used as a sufficient guide for paternalistic intervention' because it would allow intervention in too many cases such as 'to forcibly liberate people who worship false gods.'[11] The principle however was suggested as a necessary but not a sufficient condition for paternalistic intervention. Whilst this necessary condition is satisfied in the case of people who worship false gods, there could be other reasons why intervention in such cases should be prohibited. Haksar does not show that certainty is not a necessary condition for justified paternalism, just that it is not sufficient.

The certainty principle seems plausible since intervention becomes less justifiable the more remote the risk is. But it ought to be rejected for three reasons. First, stopping a person from crossing a dangerous bridge seems justified even if the danger is only probable rather than certain, if there is no time to warn him of the risk. There may be no certainty that he would fall from the bridge and instead be a chance of him making it across unharmed. But if there is some significant risk and no time for warnings then intervention seems warranted even though the certainty condition is not met.[12] Second, Mill's argument for the certainty principle is unconvincing. He claims that when the danger is less than certain, 'no one but the person himself can judge of the sufficiency of the motive which may prompt him to incur the risk'.[13] I take this to

---

[10] John Stuart Mill, *On Liberty*, Norton Critical Edition (David Spitz, ed.), New York, Norton, 1975, ch. V, par. 5, p. 89.
[11] Haksar, *Equality, Liberty, and Perfectionism*, p. 239.
[12] Haksar makes the same point on p. 239 of *Equality, Liberty, and Perfectionism*.
[13] Mill, *On Liberty*, ch. V, par. 5, p. 89.

mean that people are the best judges of whether the reasons for which they act are sufficiently strong to risk the harm that may accompany their acts. But if they are the best judges, then this seems equally true when the harm is certain as when merely probable. A person can judge whether there are sufficient reasons to incur a harm even if it is a certain harm. Mill's argument therefore fails to provide a convincing case for focusing on certainty as a necessary condition for justified paternalism. Finally, the certainty principle seems unlikely to be justified within a perfectionist theory. Trying to ensure that people live good lives supports intervening when there is the possibility that they are not doing so, not just when there is certainty that they are not. If the possible improvement to well-being is a large one, then perfectionism may support interference if it is probable though not certain that the person will not get the improvement without interference. The certainty principle therefore cannot be used by perfectionism to block paternalism.

7.3.2 Future and Hypothetical Consent

Hence we should share Haksar's rejection of the certainty principle, though for different reasons. Having rejected Mill's suggestion, Haksar presents his consent principle. The consent needed to justify intervention, says Haksar, must be present or prior and actual. It must not be future consent - that is, consent after the intervention - or hypothetical consent - that is, what people would counterfactually consent to under certain circumstances.[14]

Haksar claims that making future consent a necessary condition for justified paternalism is problematic because future consent can be the result of brainwashing.[15] He suggests a way in which this problem may be avoided. We could add Rawls's further test which says that intervention is not warranted unless there has been 'an

---

[14] Haksar, *Equality, Liberty, and Perfectionism*, p. 249.
[15] Haksar, *Equality, Liberty, and Perfectionism*, p. 240.

evident failure of reason or of the will."[16] This test rules out intervening with and brainwashing people who are in possession of their reason and will. But Haksar argues that this test is unacceptable because 'it may also allow us to interfere' with people who 'worship false gods and pursue illusory cures for their problems."[17] This argument, however, is again mistaken. The requirement that there be evident failure of reason and will is another necessary but not sufficient condition for justified paternalistic intervention. Although this condition may be satisfied in the case of people who worship false gods and pursue illusory cures for their problems, there may be reasons why interference in such cases is wrong. The condition may not rule out interference in such cases, but there may be some other condition that does. It is no argument against a necessary condition that it fails to rule out all unjustified paternalism.

Hence Haksar's argument against the suggestion designed to avoid the brainwashing problem is misguided. However, such a suggestion was not needed in the first place. The best way to rule out the possibility of future consent being the result of brainwashing is simply to hold that, for it to justify paternalism, future consent must not be the result of brainwashing. For consent to do justificatory work, it must represent the genuine will of the person involved. It does not do so if it is the product of coercion, or manipulation, and brainwashing is an example of the latter. Future consent must be genuine if it is to justify paternalistic intervention.[18]

Whilst that argument of Haksar's against future consent is unconvincing, his main arguments against future and hypothetical consent are of a different sort. He claims that even if we can get genuine future consent, paternalism may not be justified. For

---

[16] Haksar, *Equality, Liberty, and Perfectionism*, p. 240, quoting John Rawls, *A Theory of Justice*, Oxford, Oxford University Press, 1972, p. 250.

[17] Haksar, *Equality, Liberty, and Perfectionism*, p. 240.

[18] Present or prior consent must also be genuine and not be the result of brainwashing if it is to justify paternalism. Haksar does not seem to consider the possibility that those forms of consent could also be the result of brainwashing.

example, if we force a woman to marry against her will, she may come to genuinely love her husband and consent to the earlier intervention. However, such paternalism nevertheless seems objectionable. With regard to hypothetical consent, people who worship false gods 'would, if they saw things clearly, see the absurdity of their ways of life.'[19] But forcibly liberating them from their way of life would be wrong. Hence, Haksar claims, these tests of justified paternalism are implausible.

On the one hand, these objections are misguided. The suggestion was that future and hypothetical consent are supposed to be necessary conditions only, not sufficient ones. If future consent is necessary for justified intervention, then even if a woman forced to marry gives her subsequent approval, thereby satisfying the condition, there could nevertheless be reasons why such force is morally prohibited. If hypothetical consent is necessary for justified intervention, then even if forcibly liberating people who worship false gods could be supported by their hypothetical consent, there could nevertheless be reasons why such force is wrong. On the other hand though, the objection is pertinent. The fewer cases of unjustified paternalism that supposedly necessary conditions can rule out, the less useful they (the supposedly necessary conditions) are. In failing to rule out the cases Haksar mentions, future and hypothetical consent may remain necessary conditions, but not very useful ones.

This last thought can be turned to the advantage of the view I am defending and against future and hypothetical consent as alternatives to the neutrality constraint. If future and hypothetical consent are only necessary though not sufficient conditions for justified paternalism, then there must be further conditions that explain why cases of unjustified paternalism that would pass the tests set by future and hypothetical consent are unjustified. That is, if there are cases of paternalism that are unjustified even though they pass the tests of future and hypothetical consent, then there must be further principles of paternalism that explain why those cases are unjustified. One of

---

[19] Haksar, *Equality, Liberty, and Perfectionism*, p. 250.

these principles, I suggest, is set by the neutrality constraint. Forcing a woman to marry against her will is likely (though not necessarily) to be interference that is guided by a conception of the good. And forcing people out of their false religions for their own good is also paternalism that is guided by perfectionist considerations. These cases may be consented to hypothetically or in the future but fail the test of neutrality. If we accept the constraint of neutrality, both these types of paternalism would be unjustified. The problem for future and hypothetical consent is that they need to be supplemented with the neutrality constraint in order to make them more plausible. They are not alternatives to neutrality but complements of it.

Take first the case of hypothetical consent. Donald VanDeVeer considers a Principle of Hypothetical Rational Consent (PHRC) which makes paternalism justified if the paternalised person would consent to the intervention if he were fully rational and aware of the relevant circumstances.[20] VanDeVeer rejects this principle as being too wide, supporting interference with gamblers who are not aware of the true odds of winning, with those holding false religious beliefs, and with mountain climbers unaware of the risks involved. These people would, if fully rational, consent to interference. Whilst a necessary condition need not rule out all unjustified paternalism, allowing so much makes it a principle that does not do much work.[21] Instead, VanDeVeer endorses a Principle of Hypothetical Individualized Consent (PHIC) which makes paternalism justified if the paternalised person would consent if he were aware of the relevant circumstances and given his beliefs, settled preferences, and goals. The idea is to discover what the paternalised himself, rather than a fully rational individual, would counterfactually consent to.[22]

---

[20] VanDeVeer, *Paternalistic Intervention*, p. 71.

[21] VanDeVeer discusses PHRC as sufficient, rather than as necessary for justified paternalism.

[22] VanDeVeer, *Paternalistic Intervention*, p. 75.

VanDeVeer believes that PHIC avoids the too wide objection that PHRC faces. He explains how PHIC does this in two ways. First, we are to imagine what the person towards whom intervention is being considered would decide were he aware of certain evidence and arguments, rather than were he to hold different beliefs, which is what VanDeVeer takes PHRC to ask us to imagine. So instead of asking what holders of false religious beliefs would consent to if they did not have those false beliefs, we are to ask what they would consent to if they were merely aware of the evidence and arguments against their faith. In the latter case, says VanDeVeer, it is not clear that a Jehovah's Witness would consent to a blood transfusion. But the difficulty with this argument of VanDeVeer's is that even if we interpret PHIC the way he suggests, it still may not prevent PHIC from being too wide. Although it may not be as clear that compulsory blood transfusions for Jehovah's Witnesses are supported by PHIC as that they are by PHRC, they still might be supported in many cases under PHIC. If Witnesses were, hypothetically, aware of the evidence and arguments against their beliefs, they may agree to transfusions even if they are not as certain to as if they held contrary beliefs. So though this argument of VanDeVeer's makes PHIC narrower than PHRC, it is still too wide.

The second way VanDeVeer explains PHIC's advantages over PHRC (although VanDeVeer is uncertain whether it is separate from the first) is that under the former we are to take as given beliefs that are 'central to the subject's conception of his own personal identity'[23]. These could include religious beliefs and - at a stretch - beliefs in the worthwhileness of gambling and mountain climbing. Under PHIC we are to ask what someone would consent to without abandoning these central beliefs. Given this stipulation, PHIC would avoid supporting intervention in the cases that made PHRC subject to the too-wide objection.

---

[23] VanDeVeer, *Paternalistic Intervention*, p. 81.

This argument of VanDeVeer's is, I suggest, an argument for neutrality. It claims that since certain beliefs are central to, in VanDeVeer's words, a person's 'conception of how to live her life and ... her basic, life-defining and life-guiding, views and attitudes'[24] or what we can call her conception of the good, her conduct guided by those beliefs ought not to be paternalistically interfered with. Whether this is a sound argument for neutrality is not important here. What is important is that VanDeVeer accepts the conclusion of the argument, that is, the neutrality constraint. PHIC is simply PHRC with the additional assumption of the neutrality constraint, resulting in a more plausible position. Hypothetical consent, therefore, under its most plausible interpretation is not an alternative to neutrality but a principle that works alongside it.

A similar argument can be made with regard to future consent. Someone forced into a more worthwhile way of life could subsequently consent to the forcing. So could someone forced to appreciate high art rather than low culture. A further assumption, the assumption of neutrality, is needed to rule out these cases of paternalism. The counter-examples may strike the reader as unlikely. We are confident that people would not subsequently consent to intervention that prevented them from leading a life in conformity to their own conception of the good. A person who is content watching game shows on television is not going to agree afterwards with being forced to attend operas instead. But such 'conversions' can occur. Any intuition about the unlikelihood of the counter-examples, I suggest, is not based on any deep understanding of what people would or would not agree to after being the subjects of paternalism. Rather, it is based on the thought that such intervention would be wrong even if the prediction of future consent is an accurate one. And the wrongness of this intervention, I have been arguing, is explained by the neutrality constraint.

7.3.3 Actual and Present/Prior Consent

---

[24] VanDeVeer, *Paternalistic Intervention*, p. 81.

Hence Haksar is again correct, though again for mistaken reasons, to reject future and hypothetical consent, understood as alternatives to the neutrality constraint, as necessary conditions for justified paternalism. Instead of hypothetical or future consent, Haksar argues that to be justified, paternalism must be accompanied by actual and present or prior consent of the paternalised individual. The consent principle rules out as unjustified many cases of paternalism, such as forced religious worship or forced marriage, which will not be consented to.

It may, however, be thought to rule out too much. In emergency cases (and the cases are paternalistic since the people are rescued for their own good without being given any choice), such as stopping a person from unknowingly crossing a dangerous bridge, pulling a person out of the way of an oncoming truck, or preventing a person from leaping from a building in a fit of depression, paternalism is sometimes justified at least to ensure that the person knows what he or she is doing, even though there is no consent. But according to Haksar, the consent needed to justify intervention 'need not be explicit, it may only be tacit.'[25] In some cases there is no explicit consent to the intervention, but Haksar claims that tacit consent may be presumed since there is a practice of helping people in certain situations, and since most people would like to be helped in those situations and do not object to such practices.[26]

But it is not clear that people can be said to consent, even tacitly, to intervention in these cases. The fact that there is a practice of helping people in certain situations and that I have not objected to that practice does not mean that I have consented, even tacitly, to intervention when I find myself in the relevant situation. Consent, whether explicit or tacit, is an act to intentionally give others, by virtue of that act, a right that they did not have before. There is an intention to change the moral situation.[27] That is

---

[25] Haksar, *Equality, Liberty, and Perfectionism*, p. 240.
[26] Haksar, *Equality, Liberty, and Perfectionism*, p. 240.
[27] This aspect of consent is noted by writers on the notion. See A. John Simmons, 'Consent' in Edward Craig (general editor), *Routledge Encyclopedia of Philosophy*, London and New York, Routledge,

why letting a friend buy me a drink, when him doing so in the past has always been followed by me buying him one in return, amounts to tacitly consenting to buy him one in return. That is also why, contrary to Locke's claim, merely inhabiting a country does not mean that one consents to obey its government.[28] Merely not objecting to social practices of helping people in emergencies does not involve an intention to change the moral situation and is for that reason not a case of consent.

But despite the absence of consent, paternalistic intervention still seems acceptable in emergency cases. Even if one thinks, with Haksar, that in such cases there is tacit consent, imagine a case where a person unknowingly about to cross a dangerous bridge definitely cannot be presumed to have consented to intervention because he is unaware of a general practice of saving people in such circumstances and so was unable to choose whether to register his approval or disapproval of such a practice beforehand. Intervention would nevertheless not seem objectionable. Hence the consent principle cannot account for these seemingly acceptable cases of paternalism. The neutrality principle, on the other hand, can. Paternalism in emergency cases is not guided by any particular conception of the good. The good involved, the avoidance of death, is compatible with a wide range of (if not all[29]) conceptions of the good. Hence paternalism in those cases passes the test of neutrality and would, so long as it complies with other conditions for justified paternalism (see 6.5 for these), be justified.

There is a deeper problem with the consent principle. It is a principle of neutrality in the sense that it creates an area within which the state may not interfere to promote

---

1998, vol. 3, and John Plamenatz, *Consent, Freedom, and Political Obligation*, second edition, Oxford, Oxford University Press, 1968, p. 18.

[28] John Locke, *Two Treatises of Government*, (Peter Laslett, ed.) New York, Signet, 1965, second treatise, ch. 8, section 119, p. 392.

[29] Some conceptions of the good may involve not avoiding death. In these cases, neutrality requires not forcing the avoidance of death on the person. But in most emergency cases, it would be safe to assume that rescuing the person would not violate neutrality.

perfectionist aims, at least not without the consent of those interfered with. Within this area the state must leave people free from interference to pursue their own conceptions of the good. As such, the principle sits uncomfortably within a perfectionist theory such as Haksar's. We need to know why consent should have such importance in perfectionism. If a person is pursuing a mistaken conception of the good and intervention could guide him into a more worthwhile conception which would enhance his well-being, why should the absence of his consent bother a perfectionist?

Some of Haksar's comments provide an answer. It could be argued (though Haksar does not explicitly argue this) that we must accept the consent principle if we believe that an autonomous life is a worthwhile way of living. Haksar's conception of autonomy is roughly the one that was examined in chapter two, that of an individual being free to choose and pursue, without coercion or manipulation, her own plan of life.[30] The consent principle may be viewed as a way of respecting the ideal of autonomy. By prohibiting interference without consent we rule out interference with a person's pursuit of her chosen way of life. Hence there is an autonomy-based argument for the consent principle. Moreover, appealing to autonomy may give us a variation on the consent principle that is more acceptable. That is, the argument from autonomy could result in something like the condition that paternalistic intervention is justified only if consented to or is necessary to respect a person's autonomy. The second part of this principle allows us to accept that paternalistic intervention to rescue people putting their lives or health in danger is justified even though not consented to since by protecting their lives and health, we would be protecting their autonomy.

---

[30] Haksar, *Equality, Liberty, and Perfectionism*, pp. 161, 178.

Haksar claims that autonomy is an intrinsic good.[31] Assuming so and assuming that it entails the consent principle, then perhaps consent rather than neutrality provides a better explanation of the wrongness of paternalism. However, the argument faces the same difficulties, presented in chapter two, that face a direct appeal to the value of autonomy in order to rule out paternalism. Autonomy cannot be the only good and it is possible that in certain cases autonomy may be outweighed by other goods. That is, the state may be able to promote a person's own good by over-riding his autonomous choice and forcing him into a worthwhile activity. The resulting gain in well-being may outweigh the loss of in autonomy. This would mean that the consent principle would not be a necessary condition for paternalistic intervention.

One way of avoiding this is to say, as Haksar does, that autonomy is an 'essential constituent of human well-being.'[32] If so, then any invasion of autonomy for the sake of other goods would lessen well-being by undermining the essential constituent. The consent principle would consequently be a necessary condition for justified paternalistic intervention.

However, to restate the argument of 2.6, it is implausible to think that the autonomous life is essential to well-being, even for people in societies, such as most of the Western countries, where the environment encourages autonomy. It is implausible to believe that someone who lives a non-autonomous life, say by pursuing a conception of the good that she did not autonomously choose but simply adopted without reflection, will have no well-being at all even though she may pursue it with much success. Moreover, the person who pursues many worthwhile activities with great success though without being autonomous, probably has a better life than someone

---

[31] Haksar, *Equality, Liberty, and Perfectionism*, pp. 182, 199.

[32] Haksar, *Equality, Liberty, and Perfectionism*, p. 184. See pp. 189-190, 221 for similar claims. This is not the only way of avoiding the over-riding of autonomy objection. It could be held that autonomy, while not necessary for well-being, is so important that it can seldom be over-ridden by other goods. The argument against autonomy as necessary, explained below, holds against this alternative view as well.

who fails at what are worthless projects even if the latter has autonomously chosen and pursued those projects.

Haksar suggests another way of avoiding the possibility that the state could improve someone's life by forcing her out of an autonomously chosen inferior way of life into one that is valuable enough to outweigh the loss of autonomy (and thus over-ride the consent principle). This is to hold that people cannot autonomously choose lower forms of life. Such forms are only chosen by non-autonomous people. Autonomous people will always choose valuable forms of life.[33] This would rule out exceptions to the consent principle where the state could improve someone's life by interfering with her autonomous choice of an inferior form of life, since such a choice is impossible.

However, it is neither conceptually nor empirically true that autonomous people never choose inferior forms of life. Since the autonomous person, on Haksar's account, is one who chooses, without coercion or manipulation, and pursues her own plan of life, it is conceivable that such a person could mistakenly choose a worthless way of life. Autonomy is a procedural not substantive notion. Many examples could be pointed to to show that in actual fact autonomous people have sometimes chosen inferior ways of life. It may be true that autonomous people are less likely to mistakenly choose the worthless than non-autonomous people, since the former have more choices and more opportunity to deliberate about what would be in their best interests. But this is not enough for the argument to go through.

Moreover, even if it were true that autonomous people always choose the worthwhile and never choose the worthless, such a view would still not support the consent principle which requires consent before paternalism is justified. The view that autonomous people always choose the good could still allow the state to intervene

---

[33] Haksar, *Equality, Liberty, and Perfectionism*, p. 293. Haksar suggests but does not clearly endorse this idea.

without the consent of the paternalised. It could do so in order to force non-autonomous people who would choose inferior forms of life into superior ones. And it could do so in order to force autonomous people who choose worthwhile ways of life into even more valuable alternatives.

Since the argument from autonomy fails to establish the consent principle, it is still not clear why perfectionists such as Haksar can endorse a consent-based prohibition on paternalism. A principle of consent is more in keeping with the neutral position on paternalism. I do not however believe that consent is necessary for justified paternalism, even if neutrality is accepted, since it is faced with the counter-examples of interventions in emergency cases where consent cannot be obtained. But although consent is not necessary for justified paternalism, it is - along with other plausible conditions such as that the consent be voluntary, that the paternalist be reasonably competent, and that no third parties are harmed - sufficient. If the paternalised person consents to intervention for his own good then it is justified. A smoker may consent to his friends helping him to kick the habit by hiding his cigarettes. Someone may consent to being forced to save a fraction of her income to support her in old-age. Such cases, if they count as paternalism at all[34], are justified paternalism.

Moreover, they are justified even if guided by perfectionist considerations. Intervention to force a couch potato to experience high culture is justified when the couch potato has given his consent. Hence there are a class of counter-examples to the claim that neutrality is a necessary condition for justified state paternalism. When paternalism is consented to, it is justified even though the neutrality constraint is not satisfied. My claim must therefore be adjusted to the proposition that neutrality is a necessary condition for justified *unconsented-to* state paternalism. I take this to be a

---

[34] They are perhaps not accurately described as paternalism. The reason that the paternalist intervenes may be merely that the person has consented and not because he (the paternalist) thinks it is for the paternalised's good. This point was suggested to me by Jerry Cohen. But the paternalist may be motivated by both reasons, and insofar as his motivation includes the person's good then his behaviour is paternalistic.

minor adjustment. With most, if not all, instances of state paternalism, it is not possible to choose whether to consent to it or not. Neither the fact of residing in a country, nor that of voting in elections establishes that a person consents to any paternalistic policies that the state may enact. There are some imaginable cases of state paternalism that may be consented to, such as state-run organisations that force those who sign up to their programmes (thereby consenting) to attend a museum at least once a week, or state insurance plans applicable only to those who agree to them, where once you have agreed you cannot afterwards refuse to pay the premiums. But most state paternalism is not of this nature. Hence I will continue to talk of neutrality as a necessary condition for justified state paternalism without mentioning the qualification that condition is only necessary for unconsented-to paternalism, since the latter covers almost all imaginable state paternalism.

7.4 The Voluntariness Principle

A highly influential principle of paternalism is set out and defended by Joel Feinberg. He claims that

> the state has the right to prevent self-regarding harmful conduct when but only when it is substantially nonvoluntary, or when temporary intervention is necessary to establish whether it is voluntary or not.[35]

Feinberg claims that people's choices are voluntary if they make them 'when fully informed of all relevant facts and contingencies, with their eyes wide open, so to speak, and in the absence of all coercive pressure.'[36] If this principle is accepted, it may provide perfectionism with a means of ruling out wrongful paternalism without conceding to a principle of neutrality. According to the voluntariness principle, so long as a person's decision is voluntary then interference is prohibited, even though

---

[35] Joel Feinberg, *Harm to Self,* New York and Oxford, Oxford University Press, 1986, p. 126. See Feinberg's 'Legal Paternalism', *Canadian Journal of Philosophy*, vol. 1, no. 1, 1971, p. 113 for a similar statement.

[36] Feinberg, *Harm to Self*, p. 104.

the decision may be mistaken and harmful. Hence much of the unacceptable paternalism supported by perfectionism would be ruled out since many instances of forcing people into more worthwhile activities would interfere with their voluntary choices. People can and do make choices on the basis of misguided conceptions of the good and though these choices are mistaken in that sense, they are voluntary and therefore, according to Feinberg, immune from intervention.

The voluntariness principle sets out a necessary condition for justified state paternalism. Feinberg sometimes seems to regard it as also a sufficient condition for justified interference. This, however, is implausible since there are a number of other conditions that would have to be met before paternalism would be warranted and a number of considerations that would override the case for paternalism even when nonvoluntariness exists. These include (but are not limited to) the following types of cases. As Feinberg himself discusses, there must be a significant risk of harm incurred or benefit foregone for interference to be justified. Interference to prevent a trivial harm would be unjustified even if it is incurred nonvoluntarily.[37] As Feinberg also points out, someone may deliberately pursue a course of nonvoluntary choices in order to live a life that is spontaneous and unexamined, and requiring him to be more sober and calculating would be unwarranted.[38] Also, making choices, even those that are not fully voluntary, is an essential way that people acquire information, resist compulsion, and develop strength of will, and that it is important for people to be able to make mistakes and learn from them.[39] A policy of preventing some people's nonvoluntary decisions may be unjustified overall because of its effects on those who are pursuing the activity that is being interfered with with full voluntariness. Other paternalistic policies may be too administratively cumbersome to be worth

---

[37] Feinberg, *Harm to Self*, pp. 105, 118-9.

[38] Feinberg, *Harm to Self*, p. 109.

[39] Donald Regan, 'Paternalism, Freedom, Identity, and Commitment', in Rolf Sartorius (ed.), *Paternalism*. Minneapolis, University of Minnesota Press, 1983, pp. 115-6. It should be noted that Regan does not explicitly discuss Feinberg's work. But it is clear that he is talking about a notion similar to or the same as voluntariness.

implementing. The principle of voluntariness, therefore, is best understood as setting out what is necessary though not sufficient for paternalism to be justified.

Feinberg sets out a precise account of a perfectly voluntary choice. Such a choice is one where (A) the chooser is competent, i.e. not an infant, insane, or severely retarded, (B) the choice is not made under coercion or duress (which includes evils from natural sources, and coercive offers), or (C) more subtle manipulation such as subliminal or post-hypnotic suggestions, (D) the choice is not made because of ignorance or mistaken belief of the factual circumstances or the likely consequences of the choice, and (E) that the chooser does not choose in circumstances that are temporarily distorting. This final category rules out choices made impetuously, while fatigued, excessively nervous, or agitated, under the influence of powerful passions or moods such as rage or depression, under the influence of mind-numbing drugs, in severe pain, in a neurotically compulsive state, or under severe time pressures.[40]

Most choices, Feinberg admits, fall short of this standard of perfect voluntariness to some extent. To categorise all decisions that fall short of perfect voluntariness as nonvoluntary and therefore open to intervention would permit paternalism in too many cases where we would normally think it unjustified. But, as Feinberg points out, voluntariness is a matter of degree, with *perfectly voluntary* decisions at one end of the continuum, and *completely involuntary* acts, such as when one is completely controlled by external forces, at the opposite end. *Fully voluntary* choices, says Feinberg, are those that come close to the 'perfect' end of the continuum, whilst *nonvoluntary* choices are those that fall substantially short of the ideal.[41] It is only nonvoluntary choices, rather than all those that fall short of perfect voluntariness, that may legitimately be interfered with. We would not think that interference would be warranted when a person makes a decision when slightly fatigued, or slightly nervous,

---

[40] This model of the perfectly voluntary choice is set out on p. 115 of *Harm to Self.*

[41] Feinberg, *Harm to Self*, p. 104.

or with a mild headache. These would all, however, make the decision less than perfectly voluntary. But, given that the factors listed are all a matter of degree, it is only if they occur to a substantial degree, such as if there is severe fatigue, nervousness, etc., that the choice fails to be fully voluntary and may therefore be subject to intervention.

The principle of voluntariness can account for various types of justified and unjustified paternalism. It permits paternalism toward children and the mentally ill, and towards people about to make decisions that will involve great harm, without being aware of the potential harm or being in certain psychological states such as severe depression or emotional stress. More importantly, it rules out paternalism in a number of cases where intervention seems supported by perfectionist considerations. Forcing couch potatoes into more active lifestyles, for example, would usually be unwarranted under the voluntariness principle since people who choose an inactive lifestyle do not do so as a result of coercion, manipulation, factual mistake or any other voluntariness-vitiating factors. Mistakes about the value of activities do not count as factors that make decisions nonvoluntary, on Feinberg's (and any plausible) account of voluntariness. People's voluntary choices that depend upon their conceptions of the good are to be left free from interference, regardless of the truth or falsity of those conceptions.

Hence the notion of voluntariness seems to provide an account of paternalism that can provide a principled limit to intervention motivated by perfectionist considerations, without recourse to the neutrality constraint. I will argue, however, that accepting the voluntariness principle makes one committed to also accepting the neutrality constraint. Rather than being alternatives, the two notions are complementary. It is not clear, to begin with, what the factors that Feinberg lists as establishing voluntariness have in common. They are not implied by the common sense definition of 'voluntary'. Whilst voluntariness clearly rules out being forced, coerced, and

manipulated[42] (factors B and C in the list above), it does not so obviously cover the others. Infants and the mentally (A) retarded can act voluntarily. So can those who are ignorant of factual circumstances (D). Some of the circumstances grouped under category E, concerning temporarily distorting circumstances, may make an act nonvoluntary, such as a neurotically compulsive choice, but also sometimes will not, such as decisions made under time pressure or when swayed by emotions.

What general explanation, then, can be given for the various factors that Feinberg lists under his expanded notion of voluntariness? The account I shall suggest is a straightforward one. Voluntary decisions, understood as those where categories A to E are satisfied to a significant degree, are those that are likely to be correct, that is, for the person's good. Nonvoluntary decisions are more likely to be against a person's interests. Such decisions occur in circumstances where the person is likely to choose in a way that will cause him harm. Children, the mentally ill, people ignorant of relevant facts, or those influenced by drugs, or certain psychological states, or those subject to coercion or manipulation[43] are more likely to choose against their interests than competent, fully informed, sober, calmly deliberating adults. The circumstances that Feinberg lists under the notion of voluntariness are all exceptions to the rule that people are good judges of their own interests.

This explanation can account for why voluntary decisions must not be interfered with, but nonvoluntary ones may be. The latter are likely to be mistaken and require intervention to protect the person's interests, whilst the former, being less likely to be mistaken, will not require intervention. But there is one class of circumstances that are likely to result in mistaken decisions that are not included in Feinberg's list of

---

[42] Severely manipulated that is, such as hypnotised or brainwashed. Mild manipulation such as persuasive advertisements do not render acts influenced by them nonvoluntary.

[43] Coerced or manipulated people are likely to choose the right option in their circumstances, such as handing over the money rather than losing their life, but the decision would usually, in the absence of the coercion or manipulation, be against their interests.

voluntariness-vitiating factors. These are decisions made in the light of a misguided or worthless conception of the good. A person can be competent, free from coercion and manipulation, fully informed of the relevant factual circumstances, not be subject to any distorting circumstances, and yet choose a couch potato lifestyle or choose cheap pop over classical music or pornography over Proust. But such choices are against the person's interests. He or she is guided by a mistaken conception of the good and would choose differently if he or she had a sound conception. Why aren't such cases classified by Feinberg as nonvoluntary along with the other types of cases where people are likely to choose mistakenly against their own interests? The answer, I suggest, is that interference in such cases would violate the neutrality constraint. Decisions made in the light of a misguided conception of the good may be against a person's interests and that person's interests would be advanced if they decided on the basis of a better conception of the good. But they should be protected from state intervention because paternalism ought not to be guided by perfectionist considerations. Although similar to nonvoluntary decisions in that they are likely to be mistaken, decisions made on the basis of misguided conceptions of the good ought to be classed with voluntary decisions, and therefore protected from intervention. Without the neutrality constraint, choices made on the basis on misguided conceptions of the good would be classified, along with the other circumstances that tend to result in mistaken choices, as nonvoluntary and therefore may be subject to paternalism.

Hence, perfectionism cannot make use of the principle of voluntariness since that principle, on Feinberg's understanding, implies neutrality. On a different, common sense, understanding it would not imply neutrality but the principle would permit interference with much conduct thereby defeating the reason that made it an appealing principle for perfectionism to embrace.

That the voluntariness principle requires acceptance of the neutrality constraint is supported by some of Feinberg's comments. He claims that when the conditions of

voluntariness are satisfied, that is, when factors A to E are met and the conduct is perfectly voluntary, the choice that a person makes is 'the faithful expression of the settled values and preferences of the actor, or an accurate representation of him in some centrally important way.'[44] When the choice is nonvoluntary, then 'it does not come from his own will, and might be as alien to him as the choices of someone else.'[45] I take these comments to suggest an argument for neutrality, or at least reflect a concern for it. They suggest an ideal conception of the person as someone who is able to choose and pursue her own view of a good way of life. Respecting this ideal, so the argument goes, requires not interfering with a person on the basis that her chosen way of life is misguided. Hence the state ought to be neutral in its interventions. This argument for neutrality from a conception of the person is, in my view, a flawed one. But the soundness or otherwise of the argument is not important here. The important point is that Feinberg, regardless of his grounds, seems to share the conclusion - neutrality - that I am suggesting is implied by his voluntariness principle.

7.5 Conclusion

Of the principles examined in this chapter, none have provided a convincing alternative to the neutrality constraint as a necessary condition for justified state paternalism. The balancing view was unable, without further interpretation, to avoid supporting too much paternalism. The neutrality constraint provides a way of interpreting the balancing view that makes that view more plausible. The balancing view therefore, rather than supplying an alternative to neutrality, must be supplemented by it. The consent principle is too strong, ruling out paternalism in cases where it is acceptable, and the main argument on perfectionist grounds for the consent principle does not support that principle. Moreover, to be plausible, the notions of future and hypothetical consent require the assumption of neutrality. The

---

[44] Feinberg, *Harm to Self*, p. 117.
[45] Feinberg, *Harm to Self*, p. 12.

voluntariness principle also needs to be supplemented by the neutrality constraint in order to make it plausible. Without the assumption of neutrality, there would be no grounds for classifying as voluntary, and therefore immune from intervention, many decisions people make in the light of misguided conceptions of the good that ought not to be interfered with. The voluntariness principle is also not an alternative to the neutrality constraint. The neutrality constraint therefore remains a necessary condition for justified state paternalism.

Conclusion

The last chapter examined some paternalism-limiting principles that could be embraced by perfectionists in order to avoid the paternalistic implications of their theories. But I argued that those principles all ran counter to the perfectionist ideal, having to be understood in conjunction with neutrality if they are to be plausible. In chapters 2-5, I argued that various arguments drawing on perfectionist components and understandings of well-being are also unsuccessful in ruling out paternalism. Chapter 6 argued that a plausible neutral position on paternalism is able to avoid the difficulties that plague perfectionism, as well as avoiding a few of its own.

These considerations together give strong support to a two-fold conclusion. The first is a conclusion about the neutrality-perfectionist debate: neutrality is required for paternalistic state action whilst perfectionism may be acceptable in the non-paternalistic sphere.[1] From this follows a conclusion about state paternalism: it is only justified if guided by neutral considerations.

It might be objected that I have done nothing to *prove* that neutrality is a necessary condition for justified paternalism. Rather, I merely assumed it from the start and then used it to criticise perfectionism. But this would misunderstand what has been my strategy. I assumed that there were cases of unacceptable paternalism and sought whether neutrality or perfectionism could best account for them. Many of the cases of unacceptable paternalism involve the imposition of conceptions of the good on those who do not accept them but I did not assume that this was the sole feature common to

[1] Note that I say that perfectionism *may* be acceptable outside the paternalistic sphere. I have given no reason to suggest that it is or that it is not.

all such cases which accounts for their wrongness. It could have been that those cases undermined autonomy or ignored the endorsement constraint or any of the other explanations that perfectionists offer. But I have argued that these explanations are not convincing. Neutrality, on the other hand, can account for these cases of unacceptable paternalism.

The conclusion has important practical implications. As was noted in 1.3.4, the neutrality constraint entails that where there is no neutral justification for state policy, that policy is illegitimate. If neutrality is a requirement in the sphere of paternalism, then cases of perfectionist paternalism ought not to be implemented by the state. If the only sound justifications for laws against drugs, prostitution, pornography, gambling, and euthanasia are perfectionist paternalistic ones, then there ought not to be such laws. Douglas Husak argues that there are no good arguments (either paternalistic or non-paternalistic) for the criminalisation of drug use that satisfy the constraint of neutrality.[2] If so and if my conclusion is correct, then the only legitimate rationales left for an advocate of such criminalisation are perfectionist non-paternalistic ones. If there are no sound rationales of that sort, then such criminalisation is illegitimate.

My arguments are not conclusive. It is still possible that there are some other strategies that defenders of perfectionism could pursue in order to ensure that it does not support too much paternalism. There are many strands of argument in Mill's *On Liberty* that I have not been able to attend to here. Or perhaps paternalism violates some fundamental right. Or, according to an interesting suggestion made by David Archard, even if autonomy is not valuable it may be impossible for human beings to recognise that it is not.[3]

---

[2] Douglas N. Husak, 'Liberal Neutrality, Autonomy, and Drug Prohibition', *Philosophy and Public Affairs*, vol. 29, no.1, 2000.

[3] David Archard, 'For Our Own Good', *Australasian Journal of Philosophy,* vol. 72, no. 3, 1994, pp. 292-3.

Another strategy that I have not considered is to embrace a subjectivist conception of well-being. Subjectivists about well-being believe that a person's good is a function of his desires. A person's well-being is determined by his getting what he believes to be worth getting. Such a conception of well-being may admit some possibilities of mistake, such as when a person is ignorant of factual circumstances. A person wanting to drink wine without knowing it is poisoned will not advance his well-being by having that desire satisfied. But any more fundamental mistakes about what is for one's own good, according to informed-preference subjectivists, are impossible. It may seem a glaring omission of my thesis that this view's implications for paternalism are not addressed. Perhaps it is the most promising strategy for a perfectionist wishing to resist supporting paternalism to take. A state wanting to help people lead good lives should merely, if subjectivism is correct, let people do what they want.

I will restrict my comments on subjectivism to the following. One way, perhaps the best, of understanding this view is to see it as a view for political morality only. Activities and objects may have value independently of whether people want them, but from the political point of view, it is best if the state is guided by a subjectivist conception of the good in deciding what laws and policies to enforce.[4] But this approach turns what seems to be a perfectionist theory into a neutralist one. It admits that there are values people could be mistaken about but excludes them from serving as the basis for political decisions. Instead the state should just be guided by people's own beliefs of what is worthwhile and be neutral towards the truth or falsity of those beliefs in the justification of state action. Subjectivism is a form of neutrality rather than perfectionism and since there is no need to go into different forms of neutrality (other than what I say in chapter 1) I have not examined the subjectivist approach in the course of my argument.

---

[4] This is the view of Richard Arneson in 'Neutrality and Utility', *Canadian Journal of Philosophy*, vol. 20, no. 2, 1990.

I end with some sketchy comments about the theoretical implications of my conclusion. Perhaps a conclusion stronger than neutrality being required for paternalism follows from my arguments. If perfectionism requires the paternalistic policies that I have been arguing it does, and if such policies are unacceptable, as I have been assuming, then does not that mean that we must reject perfectionism completely and not only with regard to paternalism? I hesitate to endorse this conclusion. Neutrality has been subjected to powerful critiques in recent years[5] and we would need replies to these critiques, replies that I do not have, before accepting the stronger conclusion. But then if we have no replies to these critiques why should neutrality apply even in the sphere of paternalism? Perhaps we must simply accept the paternalistic implications of perfectionism. As was noted in the Introduction, my arguments give no reason why someone undisturbed by perfectionism's paternalistic implications should accept my conclusion. And even someone who is disturbed by those implications will want a more general explanation for why perfectionist considerations are excluded from state paternalism but may be acceptable in other state action. I have not given one.

Another possible complication is this. My arguments give us reason to suspect perfectionism because of the motivation it gives for paternalism. But that gives us no reason to reject it as a motivation for *anti*paternalism. Can we have a theory that says perfectionist reasons are excluded in the motivation *for* paternalism, but that they are legitimate reasons *against* paternalism (and in the non-paternalistic sphere)? Although I argued that perfectionist notions such as autonomy and endorsement are unable to rule out much perfectionist paternalism, they do some work and could be useful in a full explanation of antipaternalism. But can we pick and choose which perfectionist considerations are to count like this? Surely if they are to be counted

---

[5] See for example Joseph Raz, *The Morality of Freedom*, Oxford, Clarendon Press, 1986, chapters 5 and 6; George Sher, *Beyond Neutrality*, Cambridge, Cambridge University Press, 1997; Steven Wall, *Liberalism, Perfectionism, and Restraint*, Cambridge, Cambridge University Press, 1998, part 1.

when they go against paternalism, they must also be counted when in favour. I am not sure about the plausibility or even coherence of such a view but just mention it as a possibility.

Finally, I suggest one more possible theoretical implication. My arguments may say something about the value of freedom. Is freedom valuable solely because of its contribution to people's well-being? My arguments may imply that that cannot be the complete explanation. The neutrality constraint prohibits paternalism even when perfectionist guidance would advance the paternalised's well-being. Perhaps this means that sometimes people ought to be left free from interference even though they would be better off with that interference. The value of freedom may be at least partly independent of its contribution to people's well-being.

Bibliography

Ackerman, Bruce, *Social Justice in the Liberal State*, New Haven, Conn., Yale University Press, 1980.

Archard, David, 'Paternalism Defined', *Analysis*, vol. 50, no. 1, 1990.

Archard, David, 'For Our Own Good', *Australasian Journal of Philosophy*, vol. 72, no. 3, 1994.

Arneson, Richard, 'Mill versus Paternalism', *Ethics*, vol. 90, no. 4, 1980.

Arneson, Richard, 'Paternalism, Utility, and Fairness', *Review of International Philosophy*, vol. 43, no. 170, 1989.

Arneson, Richard J., 'Human Flourishing Versus Desire Satisfaction', *Social Philosophy and Policy*, vol. 16, no. 1, 1999.

Arneson, Richard J., 'Perfectionism and Politics', *Ethics*, vol. 111, no. 1, 2000.

Arneson, Richard, 'Neutrality and Utility', *Canadian Journal of Philosophy*, vol. 20, no. 2, 1990.

Barry, Brian, *Justice as Impartiality*. Oxford, Clarendon Press, 1995.

Brighouse, Harry, 'Neutrality, Publicity, and State Funding of the Arts', *Philosophy and Public Affairs*, vol.24, no.1, 1995.

Brock, Dan W., 'Paternalism and Autonomy', *Ethics*, vol. 98, no. 3, 1988

Buchanan, Allen, 'Medical Paternalism', *Philosophy and Public Affairs*, vol. 7, no. 4, 1978.

Caney, Simon, 'Consequentialist Defences of Liberal Neutrality', *Philosophical Quarterly*, vol. 41, no. 165, 1991.

Caney, Simon, *Liberal perfectionism*, D.Phil. thesis, University of Oxford, 1994.

Christman, John, 'Constructing the Inner Citadel: Recent Work on the Concept of Autonomy', *Ethics*, vol. 99, no. 1, 1988.

Clarke, Simon, 'Contractarianism, Liberal Neutrality and Epistemology', *Political Studies*, vol. 47, no. 4, 1999.

Clayton, Matthew, *Educating Liberals: an argument about neutrality, equality of opportunity, and parental autonomy*, D.Phil. thesis, University of Oxford, 1997.

Cloonan, Martin, *Banned! Censorship of popular music in Britain 1967-92.* Aldershot, Arena, 1996.

Coval, Simon, *Liberalism v. Perfectionism: The Personal Ethics Debate*, D.Phil. thesis, University of Oxford, 1992.

Crisp, Roger, 'Raz on Well-Being', *Oxford Journal of Legal Studies*, vol. 17, no. 3, 1997.

Dan-Cohen, Meir, 'In Defense of Defiance', *Philosophy and Public Affairs*, vol. 23, no. 1, 1994.

DeMarneffe, Peter, 'Liberalism, Liberty, and Neutrality', *Philosophy and Public Affairs*, vol. 19, no. 3, 1990.

Dworkin, Gerald, 'Paternalism', in Peter Laslett and James S. Fishkin (eds.). *Philosophy, Politics, and Society: a collection.* 5th series, New Haven, Yale University Press, 1979.

Dworkin, Gerald, *The Theory and Practice of Autonomy,* Cambridge, Cambridge University Press, 1988.

Dworkin, Gerald, 'Paternalism', in Lawrence C Becker (ed.), *Encyclopedia of Ethics.* New York, Garland, 1992.

Dworkin, Gerald, 'Paternalism: Some Second Thoughts', in Joel Feinberg and Hyman Gross (eds.), *Philosophy of Law*, Belmont California, Wadsworth, 1995.

Dworkin, Ronald, 'Liberalism' in his *A Matter of Principle*, Cambridge, Massachusetts, Harvard University Press, 1985.

Dworkin, Ronald, 'Can a Liberal State Support Art?' in his *A Matter of Principle*, Cambridge, Massachusetts, Harvard University Press, 1985.

Dworkin, Ronald, 'Liberal Community', *California Law Review*, vol. 77, no. 3, 1989.

Dworkin, Ronald, 'Foundations of Liberal Equality', in Grethe B. Peterson (ed.), *The Tanner Lectures on Human Values*, IX, Salt Lake City, University of Utah Press, 1990.

Dworkin, Ronald, Thomas Nagel, Robert Nozick, John Rawls, Thomas Scanlon, and Judith Jarvis Thomson, 'Assisted Suicide: The Philosophers' Brief', *New York Review of Books*, March 27 1997.

Dworkin, Ronald, *Sovereign Virtue: The Theory and Practice of Equality,* Cambridge, Massachusetts, Harvard University Press, 2000.

Elster, Jon, 'Self-Realization in Work and Politics: The Marxist Conception of the Good Life', *Social Philosophy and Policy*, vol. 3, no. 2, 1996.

Farrell, Martin D., 'Autonomy and Paternalism: The Political Philosophy of Joseph Raz', *Ratio Juris*, vol. 4, no. 1, 1991.

Feinberg, Joel, 'Legal Paternalism', *Canadian Journal of Philosophy*, vol. 1, no. 1, 1971.

Feinberg, Joel, *Harm to Self (volume three of The Moral Limits of the Criminal Law)* New York and Oxford, Oxford University Press, 1986.

Fotion, N., 'Paternalism', *Ethics*, vol. 89, no. 2, 1979.

George, Robert, *Making Men Moral*, Clarendon Press, Oxford, 1996.

Gert, Bernard and Charles Culver, 'Paternalistic Behavior', *Philosophy and Public Affairs*, vol. 6, no. 1, 1976.

Gert, Bernard and Charles Culver, 'The Justification of Paternalism', *Ethics*, vol. 89, no. 2, 1979.

Goodin, Robert E., 'Liberalism and the Best-Judge Principle', *Political Studies*, vol. 38, no. 2, 1990.

Green, Leslie, 'Un-American Liberalism: Raz's Morality of Freedom', *University of Toronto Law Journal*, vol. 38, no. 3, 1988.

Griffin, James, *Well-Being: Its Meaning, Measurement, and Moral Importance.* Oxford, Oxford University Press, 1986.

Haksar, Vinit, *Equality, Liberty, and Perfectionism.* Oxford, Oxford University Press, 1979.

Hodson, John D., 'The Principle of Paternalism', *American Philosophical Quarterly*, vol. 14, no. 1, 1977.

Home Office, *Supporting Families; A Consultation Document*, London, HMSO, 1998.

Hurka, Thomas, *Perfectionism*, New York and Oxford, Oxford University Press, 1993.

Hurka, Thomas, 'Indirect Perfectionism: Kymlicka on Liberal Neutrality', *Journal of Political Philosophy*, vol. 3, no. 1, 1995.

Husak, Douglas, 'Paternalism and Autonomy', *Philosophy and Public Affairs*, vol. 10, no. 1, 1980

Husak, Douglas N., 'Recreational Drugs and Paternalism', *Law and Philosophy*, vol. 8, no. 3, 1989.

Husak, Douglas N., 'Liberal Neutrality, Autonomy, and Drug Prohibitions', *Philosophy and Public Affairs*, vol. 29, no.1, 2000.

Index on Censorship, issue 6 of 1998.

Jones, Peter, 'The ideal of the neutral state', in Robert E. Goodin and Andrew Reeve, *Liberal Neutrality*, Routledge, London and New York, 1989.

Kleinig, John, *Paternalism*. Manchester, Manchester University Press, 1983.

Knott, David G., *Liberalism and the justice of neutral political concern*, D.Phil. thesis, University of Oxford, 1989.

Kraut, Richard, in 'Politics, Neutrality, and the Good', *Social Philosophy and Policy*, vol. 16, no. 1, 1999.

Kymlicka, Will, 'Liberal Individualism and Liberal Neutrality', *Ethics*, vol. 99, no. 4, 1989

Kymlicka, Will, *Liberalism, Community, and Culture*. Oxford, Clarendon Press, 1989.

Kymlicka, Will, *Contemporary Political Philosophy*. Oxford, Clarendon Press, 1990.

Larmore, Charles, *Patterns of Moral Complexity*. Cambridge, Cambridge University Press, 1987.

Larmore, Charles, 'Political Liberalism', *Political Theory*, v. 18, no. 3, August 1990.

Larmore, Charles, *Morals of Modernity*, Cambridge, Cambridge University Press, 1996.

Lockard, Craig A., *Dance of Life: Popular Music and Politics in Southeast Asia*. Honolulu, University of Hawai'i Press, 1998.

Locke, John, *Two Treatises of Government*, (Peter Laslett, ed.) New York, Signet, 1965.

Macleod, Colin M., 'Liberal Neutrality or Liberal Tolerance?', *Law and Philosophy*, vol. 16, no. 5, 1997.

Mason, Andrew, 'Autonomy, Liberalism and State Neutrality', *Philosophical Quarterly*, vol. 40, no. 161, 1990.

Mill, John Stuart, *On Liberty*, Norton Critical Edition (David Spitz, ed.), New York, Norton, 1975.

Mill, John Stuart, *Principles of Political Economy*, London, Longmans, 1896, Book 5, chapter XI.

Moore, Margaret, 'Liberalism and the Ideal of the Good Life', *Review of Politics*, vol. 53, no. 4, 1991.

Mulhall, Stephen and Adam Swift, *Liberals and Communitarians*, Blackwells, Oxford, second edition, 1996.

Murphy, Jeffrie G., 'Incompetence and Paternalism', *Archives for Philosophy of Law and Social Philosophy*, vol. 60, no. 4, 1974.

Nagel, Thomas, 'Moral Conflict and Political Legitimacy', *Philosophy and Public Affairs*, vol. 16, no. 3, 1987.

Nagel, Thomas, *Equality and Partiality*. Oxford, Oxford University Press, 1991.

Nickel, James W., 'Why Basic Liberties Are Bilateral', *Law and Philosophy*, vol. 17, nos. 5-6, 1998.

Norman, W. J., 'The Autonomy-Based Liberalism of Joseph Raz', *Canadian Journal of Law and Jurisprudence*, vol. II, no. 2, 1989.

Otsuka, Michael, *Equality, neutrality, and prejudice; a critique of Dworkin's liberalism*, D.Phil. thesis, University of Oxford, 1989.

Parfit, Derek, *Reasons and Persons*, Oxford, Oxford University Press, 1984.

Plamenatz, John, *Consent, Freedom, and Political Obligation*, second edition, Oxford, Oxford University Press, 1968.

Rawls, John, *A Theory of Justice*, Oxford, Oxford University Press, 1971.

Rawls, John, *Political Liberalism*, New York, Columbia University Press, 1993.

Raz, Joseph, *The Morality of Freedom*, Oxford, Clarendon Press, 1986.

Raz, Joseph, 'Facing Up: A Reply', *Southern California Law Review*, vol. 62, nos. 3 and 4, 1989.

Raz, Joseph, 'Duties of Well-Being', in his *Ethics in the Public Domain*, revised edition, Oxford, Clarendon Press, 1996.

Raz, Joseph, 'Liberty and Trust', in Robert P. George (ed.), *Natural Law, Liberalism, and Morality*. Oxford, Clarendon Press, 1996.

Regan, Donald, 'Authority and Value: Reflections on Raz's *Morality of Freedom*', *Southern California Law Review*, vol. 62, nos. 3 and 4, 1989.

Regan, Donald, 'Paternalism, Freedom, Identity, and Commitment', in Rolf Sartorius (ed.) *Paternalism*. Minneapolis, University of Minnesota Press, 1983.

Reilly, Irma, 'Supporting families like a rope supports a hanging man', *LM* (Living Marxism), no. 116, December 1998/ January 1999.

Richards, David, *A Theory of Reasons for Action*, Oxford, Clarendon Press, 1971.

Richards, David, 'Commercial Sex and the Rights of the Person', *University of Pennsylvania Law Review*, vol. 127, no. 5, 1979.

Richards, David A. J., 'Human Rights and Moral Ideals', *Social Theory and Practice*, vol. 5, no. 3-4, 1980.

Richards, David, 'Rights and Autonomy', *Ethics*, vol. 92, no. 1, 1981.

Richards, David, *Sex, Drugs, Death, and the Law*, Totowa, New Jersey, Rowman and Littlefield, 1982.

Richards, David, *Toleration and the Constitution*, New York and Oxford, Oxford University Press, 1986.

Sadurski, Wojciech, 'Joseph Raz on Liberal Neutrality and the Harm Principle', *Oxford Journal of Legal Studies*, vol. 10, no. 1, 1990.

Sadurski, Wojciech, *Moral Pluralism and Legal Neutrality*, Dordrecht, Kluwer Academic Publishers, 1990.

Sankowski, Edward, '"Paternalism" and Social Policy', *American Philosophical Quarterly*, vol. 22, no. 1, 1985.

Sartorius, Rolf (ed.), *Paternalism*. Minneapolis, University of Minnesota Press, 1983.

Scoccia, Danny, 'Paternalism and Respect for Autonomy', *Ethics*, vol. 100, no. 2, 1990.

Scoccia, Danny, 'Moral Paternalism, Virtue, and Autonomy', *Australasian Journal of Philosophy*, vol. 78, no. 1, 2000.

Sher, George, *Beyond Neutrality*, Cambridge, Cambridge University Press, 1997.

Simmons, A. John, 'Consent' in Edward Craig (general editor), *Routledge Encyclopedia of Philosophy*, London and New York, Routledge, 1998, vol. 3.

Spitz, David, 'Freedom and Individuality: Mill's *Liberty* in Retrospect', in John Stuart Mill, *On Liberty*, Norton Critical Edition (David Spitz, ed.), New York, Norton, 1975.

Ten, C. L., 'Mill on Self-Regarding Actions', in John Stuart Mill, *On Liberty*, Norton Critical Edition (David Spitz, ed.), New York, Norton, 1975.

Valdes, Ernesto Garzon, 'On Justifying Legal Paternalism', *Ratio Juris*, vol. 3, no. 1, 1990.

VanDeVeer, Donald, *Paternalistic Intervention*, Princeton, New Jersey, Princeton University Press, 1986.

Wall, Steven, *Liberalism, Perfectionism, and Restraint*, Cambridge, Cambridge University Press, 1998.

Waldron, Jeremy, 'A Right to Do Wrong', *Ethics*, vol. 92, no. 1, 1981.

Waldron, Jeremy, 'Legislation and moral neutrality', in Robert E. Goodin and Andrew Reeve, *Liberal Neutrality*, Routledge, London and New York, 1989.

Waldron, Jeremy, 'Autonomy and Perfectionism in Raz's *Morality of Freedom*', *Southern California Law Review*, vol. 62, nos. 3 and 4, 1989.

Weinstock, Daniel M., 'Neutralizing Perfection: Hurka on Liberal Neutrality', Dialogue, vol. 28, no. 1, 1999.

Wilkinson, T. M., 'Dworkin on Paternalism and Well-being', *Oxford Journal of Legal Studies*, vol. 16, no. 3, 1996.

Wilkinson, T. M., review of Robert George (ed.), *Natural Law, Liberalism, and Morality*, in *Utilitas*, vol. 11, no. 1, 1999,

Wolfe, Christopher, 'Liberalism and Paternalism: A Critique of Ronald Dworkin', *The Review of Politics*, vol. 56, no. 4, 1994.

Wolfe, Christopher, 'Being Worthy of Trust: a Response to Joseph Raz', in Robert P. George (ed.), *Natural Law, Liberalism, and Morality*, Oxford, Clarendon Press, 1996.

Zimmerman, Michael J, 'Voluntary Acts' in Lawrence C Becker (ed.), *Encyclopedia of Ethics*. New York, Garland, 1992.