

## **PROTECTING THE FUTURE SELF: AN EXCEPTION TO MILL'S PRINCIPLE OF LIBERTY<sup>1</sup>**

John Stuart Mill, in *On Liberty*, expresses what he calls his one simple principle as follows:

That the only purpose for which power can be rightfully exercised over any member of a civilized community, against his will, is to prevent harm to others. His own good, either physical or moral, is not a sufficient warrant. (Mill, p13)

Although individual liberty may be seen as an ideal goal – a benefit in itself – Mill defends this position from his stance as a utilitarian. He claims that erosion of this principle will be to the detriment of society as a whole. The utility of such tolerance is defended even more vigorously by Wilhelm von Humboldt who maintained that paternalistic intervention stifles independent action and produces weak willed citizens who are of little benefit to the scientific and cultural advancement of the community. Those that seek to justify exceptions to Mill's principle do so on two grounds: that modern circumstances deserve different considerations and that Mill's own utilitarian motives can be shown to breach his principle on certain cases.

Paternalistic laws and regulations are enacted to protect individuals from harming themselves by their own actions or inaction. If these actions or inactions do not 'harm' – a word with contestable definition as we shall see – any other person or institution, then it is reasonable to ask why the state chooses to act in this area at all. Why is the law applied to some areas and not others? Does each such measure act as a platform for the next and if escalation exists, what is the overall effect and is it inevitable?

Although it seems likely that the motivation for paternalistic legislation is a product of the complex world of political lobbying and populist democracy, I explore in this paper the political philosophy that is used to support acts of coercion that purport to be beneficial to the actor under threat of criminal law. Laws that are intended to prevent the use of narcotics, prohibit the sale of beef on the bone, make us wear seat belts in cars and fluoridate public water supplies may be unproductive or

---

<sup>1</sup>This has developed from a similar paper presented to the Graduate Conference on Political Theory at the University of Warwick 22 May 1999.

even counterproductive, but what I am considering here is not an empirical study of the efficacy or negative results of these measures, but the theoretical justifications that have been put forward for their implementation. I find that these arguments divide into two sections and I will briefly summarise the categories in each, concentrate on one which I find particularly engaging and assess its relevance to the others.

Characteristic of this debate is the tendency to drift from one unconcluded category of justification to another. Parliamentary debate is particularly prone to listing, sometimes contradictory, reasons as justification by weight rather than quality. The description of the categories that follow may appear dismissive but the intention is to identify them in order to clarify the limits of the major subject of this paper and not to despatch or prove them. The major subject is the justification of paternalistic intervention based on a philosophical presumption that, in certain ways, one's future self is sufficiently psychologically separate to be considered another identity. This is dissected into an assessment of how various moral claims may be dealt with in the terms of this theory before applying it to paternalistic coercion.

The types of justification for paternalistic intervention may be divided into the paternal and non-paternal. Although these labels have been used before (VanDeVeer), I have found that writing of paternalistic and non-paternalistic justifications for paternalistic intervention has confused some readers searching for an anti-paternalist strand that is not present. To avoid the unwieldy, I have adopted the terms benevolent and civic justifications for paternalism.

One side accepts that the intervention will be classified as paternalistic and seeks to prove that, on balance, the level of paternalism is justified. This stresses that there is a need to protect individuals from themselves in some cases. Mill himself gives exceptions to his principle which allows for many cases to fit into the categories within this branch with a little elastic interpretation. The problem with this approach is that one is called upon to explain why intervention may be justified in one case and not in another without exposing some bias. Where, precisely, does one propose to draw the line? I call this the benevolent justification for paternalism.

The other branch accepts the need to avoid paternalistic intervention unless a clear case is brought forward by the proposer which denies paternalistic tendencies. The justification for intervention on this approach is a case by case claim that an intervention is not paternalistic. It would attempt to show that some actions that are, on the surface, of concern only to the individual, may substantially harm others and

could no longer be considered self regarding. This position appeals to the more acceptable type of intervention: the private harm principle: preventing the individual from harming others. I call this the civic justification for paternalism.

## 1. Benevolent Justification

The benevolent justifications fall into four distinct categories. I have grouped them as follows:

### 1. Freedom maximisation

Mill was quite precise about the absolute prohibition of anyone selling themselves into slavery. This, he claims, would throw away the very liberty that is being defended. Some have sought to extend this principle to other areas; Lincoln Allison (p114) highlights the ambiguity around defining whether one is free to enjoy drugs or be a slave to them, freed by colonialism or freed from colonialism.

There are examples of freedom maximisation that would find broad support, where choices are made under subtle duress and the interference of an external agency may be welcomed. Doubts cast on one's masculinity by peer pressure is a commonly given reason for unnecessary risk taking and the coercion by sporting bodies regarding protective equipment or performance enhancing drugs may well receive consensus from those involved. Their specificity to a particular group, however, probably grants them more legitimacy to act in this way than that of Parliament over the general population.

Mill did not suggest, however, the imposition of criminal penalties – the usual legislative deterrent – on anyone that sells themselves into slavery but proposed that any such contract arising from such an agreement should not be enforced. The nature and conditions of contract are the legitimate interest of the state. It might be argued that the slavery exception, selling oneself as a slave, is not properly liberty of the individual in matters that affect no other, but concerns the *implied* liberty of individuals to form mutually binding contracts and the adjudication or acceptance of a contract by an independent arbiter which, inevitably, is an agent of the state.

### 2. Self paternalism

It cannot be denied that we sometimes act in a paternalistic manner towards ourselves, although it is clear that in these cases we are judging our own best interests and not raising contentious issues with either the paternalist or the libertarian.

Examples might include the action of moving the alarm clock to the other side of the room when expecting, for whatever reason, a reluctance to get up in the morning.

Another often quoted example (Bayles, Dworkin, Elster, Parfit 1973) is that of Odysseus (or Ulysses) from Greek mythology who, wishing to hear the call of the sirens but fearing the consequences, orders his crew to lash him to the mast and plug their own ears until the danger has past. In anticipating a future weakness, he makes provision to safeguard the ship whilst his mind may be unbalanced by extreme temptation. A similar situation may be if you were to ask your friends not to give you cigarettes even if specifically asked.

These last two examples have been used in the future self argument – that your future self may be harmed by your current actions and therefore causes harm to another person. However, these examples seem to belong quite clearly in the category of self paternalism as the personal identity, for these purposes, remains the same: it is merely an attempt to override the potential weaknesses, temptation or addiction of oneself in the future.

The acceptance of self paternalism only becomes a legislative issue when it is suggested that in authorising a democratic government to intervene in our lives for the collective and individual good of us all generally, we defer to them the right to intervene in a more general manner in ways that would be in our best interests. A small sacrifice of liberty is a part of the deference to governance which is in itself a self paternalistic act.

### **3. *Balancing of harms and benefits***

In stretching principles by degrees, there inevitably appears a set of scales to balance a loss of liberty against a gain in some other sense. It is difficult to argue against a life saving measure on the grounds of minor inconvenience but easier to defend a slightly dangerous activity if it is essential to your enjoyment or way of life.

If the overwhelming majority have no interest in, or disapprove of a certain activity, then the value attached to it by them will be negligible, far less, presumably, than the actor involved who would value it more. On the other side, the benefit gained by the actor in terms of health or longer life are, I argue elsewhere, often founded on popular wisdom and the weight of evidence necessary to tip the scales is dependent on the cultural and numerical status of the actors affected. In the words of CL Ten.

But where we disapprove of an activity, or cannot appreciate it, we tend to think that the agent himself derives little benefit from it. In these ways the practice of strong paternalism easily becomes a cloak for the imposition of our values on those who are coerced. (Ten, p117)

#### **4. Knowledge of one's objective interests**

A further justification for intervention is the notion, so effectively denounced by Isaiah Berlin, that we may coerce individuals in the name of a particular goal that they would pursue themselves were they more enlightened. Once we take the view that we know their objective interests better than they do themselves, we may, Berlin argues, “bully, oppress, torture them in the name and on behalf of their ‘real’ selves” (Berlin p133).

This assertion of a superior knowledge of someone else’s objective interest may be justified when considering a parent’s authority over their child – an exception accepted by Mill – but not if judging all actions of a minority group as the majority probably, “interferes wrongly and in the wrong place” (Mill, 1859, p83). Once in place, an interference may become a precedent in order to dismiss the views of that group in future and label it as permanently unaware of its objective interests.

Where identifiable classes of individuals can be shown to be *likely* to manifest choice behaviour inconsistent with their preferences as those preferences can be unproblematically attributed to them, the odds may change in favor of interfering with their personal liberty. (Sartorius, p157) [my emphasis]

## **2. Civic Justification**

We now turn to the civic justifications for paternalistic intervention

### **I. Burden on the State**

The first claim on civic justification would point to the burden on state funded provisions such as health care or social services resulting from irresponsible action on the part of individual actors. It may be claimed that this argument is used most where the effects are least – relying on appeals to common assumption rather than evidence – and that measures that would result in a substantial and identifiable reduction in health care costs are normally of the sort that would not be politically acceptable.

## **2. State investment costs**

Secondly, the state has made an investment in people as working units and are entitled to a productive return and not have that life rashly curtailed. We educate and provide for all whilst opportunities to ‘repay’ society are limited. There are only a finite number of roles within which one may repay debts to society and many will remain an unemployment, pension and elderly health service burden. It seems unreasonable to complain that a few lives shortened by rash behaviour have left a debit on an individual's account with the rest of us.

## **3. Effects on those close to the actor**

An emotional appeal to view self harming action as other regarding is one which reflects on the suffering of friends and relatives after the premature death or injury of a loved one. Those who witness their painful death after a smoking related illness, those that give up parts of their lives to cope with the drug addiction of someone close to them and those that may have to care for a paralysed relative as the result of a risky act of bravado or failure to protect themselves adequately do indeed suffer a harm.

The failure of this appeal is that the harm claims to be relevant even if it is less than that born by the actors themselves. If this is accepted, then it appears to have no readily drawn boundaries. Surely every action anyone takes will, to some degree, affect someone else, whether or not we would normally define that act as self regarding. The combination of those two points would lead many to argue that an unacceptably wide range of restrictions would result. This can be taken to mean that the state is to protect everyone from the effects of another’s actions and that in the end nothing is self regarding.

## **4. The future self argument**

Finally, the future self argument, which I pursue in more detail. I shall describe the theory in the terms of Derek Parfit (1973, 1975), whose work on personhood and identity may be adapted to view paternalistic intervention.

The argument, put simply, that as one’s character changes over time, and particularly as a result of traumatic events, each change in character constitutes the creation of a fresh identity and therefore a later self. This argument suggests that your later self is, for the purposes of Mill’s basic premise, another person: thereby denying

claims that a particular act is entirely self regarding. It is interesting to note that acceptance of the principle is used to argue both for and against intervention (reference CL Ten later in this paper).

The psychological link between ourselves now and what we were some years ago is in its nature one of varying degrees. Some views we hold or targets we seek may be identical and others may be quite different. If however we take the view that each day we live is directly connected with the previous one and, barring an attack of total amnesia, we are for all purposes the same person, then we speak of a continuity that makes us the same person at all times. Derek Parfit, in ‘Personal Identity’ and ‘Later Selves and Moral Principles’ describes what he calls the complex view which holds that although the continuity view is an all or nothing assessment, it involves ‘connectedness’ which may hold on a stage by stage basis, but over time will be a matter of varying degrees.

There are many cases that could be considered differently depending on whether the simple or complex view is held. I shall outline the cases put forward by Parfit, those on desert, promises and distributive justice, before moving on to paternalistic coercion, and highlight inconsistencies if this theory is extended. These are presented in the following subsections.

### **1. Desert**

The complex view on desert may suggest that a criminal whose psychological connection with the crime is weak, deserves less punishment. John Locke has suggested this in *An Essay Concerning Human Understanding* (book II chapter XXVII section 26).

Locke’s assessment of variable punishment is taken further by Donald Regan who, in ‘Justifications for Paternalism’, presents us with the example of an embezzler who, discovered after ten years, fully regrets what he has done and has a different set of values. It could be argued that he is so different that it would be unfair to treat him now as the sort of person who would embezzle. However, if they were also to be guilty of a separate crime – Regan’s suggestion is a violent assault – which has not been psychologically separated, how do we deal with this? Regan uses the future self argument to imply that we now have two different persons in the same body at the same time. It is clearly not possible to treat them as different people for the purposes of inflicting a punishment.

Regan's proposals raise several problems regarding desert for different future selves in the same body which, although they are applied to other regarding acts, may also be applied to paternalistic measures in a deeper way as the one who is being protected is also the one being coerced. Punishing someone on the grounds of protecting a potential future self, is also punishing the other person with many other plans who has made promises which may not be kept due to the punishment inflicted.

## 2. Promises

With regard to promises, if we claim to have some commitment to our relations because they are related to us, do we have a lesser commitment to someone who is less closely related? If we follow the complex view we may believe our commitment varies by matters of degree. This assessment might then be applied to the idea of future selves. Promises, however are not easily adapted to a scalar view: you cannot be bound by a promise to a greater or lesser degree, you either are or are not. Although conditions put upon the promise may introduce variables, the original agreement is either binding or not – it cannot be binding to a greater or lesser degree.

If a promise is made to someone now, how much should your future self be bound to it? It seems fair to believe that if you make a promise to someone, your future self cannot release you from it, but the recipient may and therefore the future self of that recipient may also.

If you promise a mother that, after her death, you will care for her child, you are obviously bound by that promise even though the recipient is not the one to whom the promise was made. There should therefore be no obvious objections to the proposal that you are bound to a promise made to any future self, including your own, no matter how psychologically separate they may become.

An interesting scenario is drawn by Parfit (1973 pp145-6) as follows. A nineteenth century Russian Socialist, due to inherit land in the future, decides that it should be given to the peasants and draws up a document to that effect. In the event of his socialist ideals changing over the years he makes up the document to say that only his wife can rescind the commitment. He then asks his wife to promise that she will not do so no matter how much he may plead.

When the time comes, he does change his views and pleads with his wife to change the document. Should his wife believe that if his views have changed, then he is surely a different person to the one whom she married many years ago and to whom

she made the promise. If so, he is unable to release her from that promise.

Alternatively, she may feel that he is the same person and may legitimately release her from her responsibility.

This is not to be confused with ‘self-paternalism’ mentioned previously which elicits a promise to prevent one from actions that might occur under the influence of addictive cravings or extreme temptation. Examples of these might be a promise not to let me have any cigarettes even if I beg for them or the case of Odysseus and the Sirens from Greek mythology. These suggest that the one to whom the promise is made is still the same person, but under a predicted duress; in the first, a craving for tobacco which they obviously intend to give up, and in the second, a temptation by the call of the Sirens to dash the ship against the rocks during a temporary diversion from reason.

### **3. Distributive Justice**

Parfit’s application of the complex view of identity to distributive justice softens the distinction between the utilitarian view of distribution and that adopted by John Rawls in *A Theory of Justice*. For the purposes of distribution, the utilitarian considers the divisions between lives as insignificant; it matters little who enjoys the benefit that results from a burden that has been endured by somebody as long as the benefit is greater than the burden. To the Rawlsian, this is unfair; if a burden is endured by somebody, then the reward should come to the same individual. There are clearly strong boundaries that separate lives for the Rawls viewpoint that do not exist for the utilitarian.

The complex view of identity breaks down this distinction. If our future selves are less psychologically connected to us, then the sub-divisions within lives compete with the divisions between lives.

### **5. Paternalistic Application of the Future Self**

The case of the Russian noble presents an interesting problem for the complex view of the future self; he may be considered a different person as his values have changed so much, but the promise was made to cover that very eventuality. Whilst this gives an example pertaining to commitment, it also shows the opposite view when considering paternalistic intervention. The priority in the promises example is given to the ideals of the younger man and the values of his later self are considered to be overridden. In using the future self argument in matters of intervention, the

paternalist argues that if the consequences of what you do now are likely to be regretted by a future self then they must be curtailed. This is noted by Bernard Williams with regard to the Russian.

Why should I hinder my future projects from the perspective of my present values rather than inhibit my present projects from the perspective of my future values? (Williams p206)

Presumably, Williams' question can be inverted to ask why the paternalist should treat present values as subservient to the future ones. Without a time machine, we can never extract the opinion of the former self on the values and judgements of a later self, but discussion between a teenager and parent may give some guide.

Someone suffering a smoking related illness will probably regret their previous smoking habit, but, unless they were born a very long time ago, they would have been aware of the risks and fully informed of the consequences throughout their smoking life. The fact that they started smoking must have been a decision made as a sane and rational adult. In advocating coercion to prevent someone taking risks which they feel worthwhile on the basis of the opinion of a later self is implying that the older member of society is always wiser than the younger. The existence of this position is confirmed by Robert Goodin.

The risk-courting preferences of youth are a characteristic product of a peculiarly daredevil phase that virtually all of them will, like their predecessors, almost certainly outgrow. (Goodin, p119)

If we suggest that the younger person does not give proper consideration to the welfare of a later self, we are perhaps endorsing Parfit's complex view in believing, like the undeterred criminal, that they do not care about their later selves as they do not feel closely related to them.

There are many problems that surround the concept of the future self constituting another to a greater or lesser degree when we also consider the competence of judges. It might be suggested that a later self, damaged by smoking related diseases for example, is in a better position of judgement due to their relevant experience to travel back in time and stop their former self from smoking. But we can not go back in time, and smoking carries a risk, not a certainty, of debilitating illness and death. That risk is assessed by someone not necessarily ignorant of the risk and gravity of the case, but with different attitudes towards risk. That attitude is not necessarily wrong but different.

It is suggested that one should refrain from a certain action or behaviour

because one's future self, who may be damaged by such action or behaviour, is so psychologically separated from you as to justify calling them a different person: it would therefore be other regarding action. It might then be expected that, if such a psychological separation exists, the judgements of the earlier self will not place great value on the desires and plans of their later self – assuming they are even known.

To an elderly person, a few extra years on their life span may be of prime concern: good health may be of secondary concern, with all other desires of a minor importance. To a younger person, these may not be so important. Danger itself is something sought by younger people: fun is more important than long term health. It may well be judged that a hallucinogenic experience promised by a drug may outweigh the risk associated with taking it. The mistake of the paternalist is to assume that longevity is the sole objective.

Just as Mill asserted that a nation will stagnate if individuality is stifled and conformity is enforced, perhaps society will also cease to function if individual values become common across the ages; how would the armed forces cope if young people refused to do anything dangerous? Watching television would become more important than climbing mountains and crossing oceans.

There are further complications when we consider that the future self is born of a single traumatic event. CL Ten, in *Introduction to J.S. Mill's On Liberty*, draws attention to the reliance of this theory on a future self, with different attitudes towards risk, being born of an event. Using the examples of motorcycle crash helmets and car seat belts, it is accepted that, in the event of an accident, it is the accident itself that creates the later self, not the wearing of a helmet or seat belt. If death results from an accident when no protective devices are used, then no future self is born: it is entirely self regarding.

Death, at least instant death which leaves no time for regret, ensures that there is no birth of the later self, and so death makes the accident self-harming but not harmful to another person....It is paradoxical that activities that are likely to cause these types of serious harm fall outside of the scope of Regan's and Kogan's new defence of paternalism, whereas activities which are likely to cause lesser harm can be interfered with. (Ten, 1980, p121-2)

On the future self justification, we seem to have a reason for discouraging the use of devices that will prevent death in order to avoid an injured future self full of remorse and regret.

What I would like to suggest is that it is inevitable that one's values and goals

change with age and that to interpret that change as good is to equate experience with wisdom. If we accept that the later self constitutes a different person, then any judgement of previous action as regrettable is not valid as the judge has different opinions and values and may be considered a different person. They may even belong to a different culture. He or she is not ‘qualified’ to judge.

The probably misplaced belief that recreational drug use occurs only amongst the young, reinforces the idea. If people stop using drugs at the age of twenty, They have perhaps, with maturity, judged that activity as regrettable. It is then, with an extension of the assumed irrationality of children, easy to prohibit drugs on the basis of an expected later consent.

### **3. Conclusion**

I hope that I have explained that there are many reasons given to justify the application of paternalistic coercion and that they fall in to two sections with different ideological foundations: one accepts Mill’s principle and the other seeks to breach it. One could not therefore look to one section for supporting argument without having first given up on the other – even though they sometimes seem to be making the same case.

The debates around personal identity make the future self argument the most thought provoking and sadly inconclusive. The theory claims to justify both intervention and non-intervention on behalf of a future self. It also explains the actions of the individuals who are the objects of coercion. People smoke tobacco knowing that it increases the chances of a painful death. They also know that most people who smoke regret having taken it up. Only the complex view of identity can explain why we care so little about our future selves: less even than others around us.

The proposal is that Parfit’s complex view of the future self is a justification for legal intervention on the basis that the future self is ‘another’ for the purposes of denying that a particular act is self regarding. If we accept this proposal, a number of consequences arise that have significance elsewhere.

If the later self is sufficiently changed by traumatic experience to be considered a separate person for the purpose of utilising the private harm principle, then they may also be separate for the purposes of assessing the qualification as a judge of utility. Although Bentham might argue that the later self is qualified as a judge due to the experience of both the early benefit and subsequent harmful consequences, this invites us to abandon, temporarily, the separation of the identities

that is so essential to the proposal. It may be that the competence of the later self to make judgements is lessened by the very different framework of values within which that judgement is made. This does no damage whatsoever to Parfit's complex view that a self regarding act may not be considered as such if traumatic events or debilitation give birth to a later self: a separate identity. However, it challenges the assumption that the values around which the later self makes judgements should be given precedent and imposed upon the young.

In addition to a form of justification, the future self argument can be usefully used as a form of behavioural explanation. When the future self argument is applied to distributive justice as noted earlier, it suggests subdivisions within lives in addition to the divisions between lives. In this way it can be explained why someone might behave in a selfish way in relation to their later self; they feel less psychologically connected to them than to their contemporaries.

## BIBLIOGRAPHY

- Allison L *Right Principles: A Conservative Philosophy of Politics* (Oxford, Blackwell, 1984).
- Bayles Michael D ‘Criminal Paternalism’ in *The Limits of Law: Nomos XV*. Yearbook of the American Society for Political and Legal Philosophy (New York, Lieber-Atherton, 1974) pp174-188
- Beauchamp Tom L ‘A critique of paternalism’ Tom Beauchamp and Terry Pinkard eds *Ethics and Public Policy: an Introduction to Ethics*. (New Jersey, Prentice Hall, 1983)pp86-96.
- Berlin Isaiah ‘Two Concepts of Liberty’, *Four Essays on Liberty*. (Oxford University Press, 1969) pp118-172.
- Dworkin G ‘Paternalism’ Beauchamp and Pinkard eds. *Ethics and Public Policy: an Introduction to Ethics*. (New Jersey, Prentice Hall, 1983)
- Elster John *Ulysses and the Sirens: Studies in Rationality and Irrationality*, (Cambridge, Cambridge University Press, 1979).
- Goodin R Amitai ‘Permissible Paternalism: In Defense of the Nanny State’ Etzioni ed. *The Essential Communitarian Reader* (Lanham Maryland, Rowman and Littlefield, 1998) pp115-123
- von Humboldt W *The Limits of State Action*. ed JW Burrow (Indianapolis, Liberty, 1993)
- Locke J *An Essay Concerning Human Understanding*. ed Peter Nidditch (Oxford, Clarendon, 1975).
- Mill JS 1989). *On Liberty*, 1859, Collini S (Cambridge University Press,
- Parfit Derek ‘Later Selves and Moral Principles’, Alan Montefiore ed. *Philosophy and Personal Relations: An Anglo-French Study*. (London, Routledge and Kegan Paul, 1973) pp137-169.
- Parfit Derek ‘Personal Identity’, John Perry (ed), *Personal Identity*, (Berkeley, University of California Press, 1975) pp199-220.
- Rawls J *A Theory of Justice*. (Oxford University Press, 1992).
- Regan Donald ‘Justifications for Paternalism’ *The Limits of Law: Nomos XV*. Yearbook of the American Society for Political and Legal Philosophy (New York, Lieber-Atherton, 1974) pp189-210.
- Sartorius R *Individual Conduct and Social Norms: A Utilitarian Account of Social Union and the Rule of Law*, (California, Dickenson, 1975).
- Ten C L *Mill On Liberty*, (Oxford, Clarendon, 1980).
- Van de Veer D *Paternalistic Intervention: The Moral Bounds of Benevolence*, (Princeton University Press, 1986).
- Williams B (ed) ‘Persons Character and Morality’, Amelie Oksenburg Rorty *The Identity of Persons* (Berkeley, University of California Press, 1976).