Paternalism, Freedom, Identity, and Commitment

Donald H. Regan
University of Michigan Law School, donregan@umich.edu

Available at: https://repository.law.umich.edu/book_chapters/245

Follow this and additional works at: https://repository.law.umich.edu/book_chapters

Part of the Law and Philosophy Commons

Publication Information & Recommended Citation

This Book Chapter is brought to you for free and open access by the Faculty Scholarship at University of Michigan Law School Scholarship Repository. It has been accepted for inclusion in Book Chapters by an authorized administrator of University of Michigan Law School Scholarship Repository. For more information, please contact mlaw.repository@umich.edu.
Paternalism, Freedom, Identity, and Commitment

Donald H. Regan

Some years ago, I wrote an essay entitled "Justifications for Paternalism."1 That essay is here revised, and expanded by the addition of a new topic. Many readers of the original version did not understand that the two principal sections presented arguments that were quite independent. I would therefore emphasize that in the present version the three principal sections (II, III, and IV) are separable one from another. Not surprisingly, in an essay so disconnected, I reach no general conclusions I have much confidence in. I suspect the reason for the failure is that I have been insufficiently daring in rejecting common premises for thinking about the problem. That, however, is a story for another time.

I

It may be useful to rehearse briefly the main points of what I take to be the standard dialectic of paternalism. In the context of a traditional utilitarian approach to the problem of paternalism, there is one necessary and sufficient condition for paternalistic coercion—namely, that the coercion will result in more pleasure or happiness overall for the person coerced. If some individual will be happier overall if she abstains from cigarettes, or from heroin, or if she wears seat-belts in cars or a helmet when riding a motorcycle, we should coerce her to do all those things. That is all there is to it. But this singleminded pleasure-maximizing approach does not satisfy us. Anyone who suggests that we are always justified in compelling people to do that which will make them happiest is ignoring another value that is not the same as happiness, the value of freedom of choice. Among our intuitions we seem
to find the idea that individuals have a right to make their own choices, even if they are bad ones. In cases in which paternalistic coercion would be justified on utilitarian grounds, two important values, pleasure or happiness on the one hand and freedom on the other, seem to be irreconcilably opposed.

We might try to resolve the conflict by a rule-utilitarian move. It seems plausible to suppose that in many cases, perhaps most, individuals have a better idea what will make them happy than do remote legislators. If that is so, paternalistic legislation may be misguided more often than not, and a general principle forbidding such legislation, possibly enshrined in a constitution, might make good sense. This is a significant argument. In the end, it may be the strongest argument that can be made for a general prohibition on paternalism. However, it leaves the defense of freedom dependent on a contingency. If there is any intrinsic value in freedom, this argument does not give that value its due. I shall therefore eschew this and other rule-utilitarian arguments in what follows. Except where I specify otherwise, I shall consider the question of when coercion would be justified from the point of view of an idealized paternalist, who not only knows everything about the individual he is coercing and the consequences of various choices by that individual but who also has at his disposal means of coercion that can discriminate perfectly between different individuals and different acts. My hypothetical paternalist does not make mistakes; he need not worry about possible overbreadth in general prohibitions; and he operates in a system in which paternalistic legislation had no bad effects on the administration of the legal system generally. I am ignoring serious practical problems, because it seems to me that before we can decide what sorts of paternalism are justified in practice, we need to have some idea of what sorts would be justified for my ideal paternalist.

We might also try to resolve the conflict between happiness and freedom by saying that some considerable degree of freedom is a necessary condition of being happy. Unfortunately, this claim is not obviously true—witness the cases of nuns, soldiers, and others who manage to be happy inside total institutions—and even if it were true it would still seem to subordinate to happiness something which is an independent value of equal stature.

Another move, still within the standard dialectic, is more promising. It is often suggested that paternalism may be justified when the individual coerced lacks relevant information about the consequences of her acts. Presumably this justification for paternalism gets its force from a feeling that ignorance is a sort of unfreedom. Since the person who lacks information is unfree even if we do not intervene to constrain her choice, we are not really decreasing her freedom by intervening, and the conflict between freedom and happiness never arises. Just why being uninformed is a way of being unfree is an interesting question. Certainly it smacks more of a lack of "positive" freedom than of a lack of "negative" freedom. Still, as an excuse for paternalism, ignorance is in reasonably good repute even with advocates of negative freedom.

If our justification of paternalism is simply people's ignorance, it might seem that we have a warrant not for coercion, but only for education. If the reason
we feel justified in forbidding drugs is that we don’t think users realize the danger to themselves, should we not concentrate on informing them of the danger, and then let them do as they please? In fact, there are a variety of claims we might put forth to justify coercion in particular cases. Sometimes there simply will not be time to educate the party coerced, as when someone threatens to act in a way that will do her irreparable damage before we can convince her of what the facts are. In other cases, the party to be coerced may lack the expertise to understand or use the information she should have.

Stretching the concept of information a bit, we might suggest that even when expertise is not in question, an individual might have all the facts within her cognitive grasp and still not really appreciate them. For example, someone might know all the medical facts about cigarettes, emphysema, and lung cancer, and still fail to appreciate just how unpleasant the possible consequences of smoking are. In a similar vein, we might suggest that most people are simply incapable of taking very small probabilities properly into account, and this could be regarded as a sort of ignorance about consequences. With these arguments, we can defend a good deal of paternalism with some persuasiveness on the ground that we are interfering only when people lack information.

Another move, similar to the appeal to ignorance, is the claim that paternalism is justified when the subject of coercion is acting under psychological compulsion or under unusual social pressure. In these cases also, we might claim not to interfere with freedom, since the subject of coercion is already unfree, and we might nonetheless open up very considerable opportunities for benevolent intervention.

We have now reached approximately the point Gerald Dworkin reaches at the end of his deservedly influential essay on paternalism. After canvassing much the same justifications for paternalism as those I have mentioned, Dworkin concludes, in effect, that paternalism is acceptable so long as “we are simply using coercion to enable people to carry out their own goals.” Ignorance, psychological compulsion, or even outright weakness of will may prevent people from achieving their goals, and paternalistic coercion is an appropriate remedy. To be sure, Dworkin does not say we are justified in coercing the ignorant or the weak-willed just to maximize their pleasure or their happiness. He speaks rather of enabling them to accomplish their own goals, whatever those may be. Still, Dworkin's conclusion is troublesome, for two quite different reasons.

On the one hand, it is not clear that even after we limit paternalism as Dworkin would limit it (to cases of ignorance and the like), we have given freedom of choice its due. We are so seldom fully informed of the consequences of our acts, and we are so seldom unaffected by psychological and social pressures of various kinds, that our whole lives might be subject to paternalistic supervision if the approach we have been expounding were taken seriously. Furthermore, this approach fails to take into account a very important fact, which is that making choices, including bad ones, is an essential way in which people acquire information, learn to resist compulsion, and develop strength of will. Freedom, in
the sense of the ability to make a genuinely free choice, is not merely something one has or does not have at any point in time with respect to any particular choice. It is a general capacity that one acquires and improves as one exercises it. The achievement of freedom is impossible unless one is allowed to make some decisions that are not fully free. To be sure, Dworkin's statement about enabling people to carry out their own goals (and other similar statements) could be construed as taking this need for "exercise in freedom" into account. But the possible conflict between enabling someone to make the present choice which will further her goals and enabling her to develop a capacity for improved choice in the future goes unnoticed. A related point, less important but worth mentioning, is that we may not value equally Smith's achieving her goals with paternalistic help and Smith's achieving her goals on her own. In other words, we may think there is an important difference between Smith's goals being achieved and Smith's achieving them.

My other objection to the approach summarized in the quote from Dworkin is that it assumes each person has, or can be regarded as having, something describable as "her" goals. This is mildly problematic even if we consider only the moment of the choice we propose to interfere with—few people have goals that are fully worked out, and the vaguenesses and gaps might be important. What really troubles me, however, is the implicit assumption that people's goals are stable over time. People change over time, and their goals change with them. Two of the three sections that follow will consider ways in which this fact affects the analysis of paternalism.

II

In the preceding section, we considered one standard approach to minimizing the conflict between our paternalistic impulses and our inclination to value freedom. We saw that if we regard as unfree acts that are performed in ignorance, or under pressure, or as a result of weakness of will, we can indulge in a great deal of paternalism without interfering with freedom at all. This approach, however, threatens to justify more paternalism than we are comfortable with.

In this section, I shall consider a quite different approach to minimizing the conflict between our paternalistic impulses and our concern for freedom. The approach is suggested by John Stuart Mill's discussion of a contract for slavery. Mill says that although we generally enforce contracts, out of respect for individuals' free choices, we should not enforce contracts for slavery. Mill writes of the would-be slave:

[B]y selling himself for a slave, he abdicates his liberty; he foregoes any future use of it beyond that single act. He therefore defeats, in his own case, the very purpose which is the justification of allowing him to dispose of himself. He is no longer free; but is thenceforth in a position which has no longer the presumption in its favor, that would be afforded by his voluntarily
remaining in it. The principle of freedom cannot require that he should be free not to be free. It is not freedom to be allowed to alienate his freedom.⁵

There is a certain looseness here. A person might have what seems to her a very strong reason for wanting to sell herself into slavery. It might, for example, be the only way she could secure the money for an expensive operation necessary to save the life of her child. In such a case, a prohibition on contracts for slavery would be a genuine barrier to achieving the person's goals. It would be a genuine denial of freedom.⁶

Still, Mill has a point. It seems quite natural to argue that even if the act of selling oneself into slavery can be an exercise of freedom, the act in question destroys much more freedom than would be destroyed by prohibiting the act. The person who wants to sell herself into slavery and is not allowed to is less free in regard to her present desire, but she will be freer, in the long run, overall. Therefore, our valuing freedom does not weigh against prohibiting contracts for slavery. It supports it. Whatever the basis of our original impulse to forbid the contract for slavery, that impulse does not conflict in this case with our concern for freedom. The apparent conflict is illusory.

What we have done, in effect, is to substitute for the deontological principle that an agent's freedom must not be interfered with, a teleological principle that the agent's total freedom should be maximized over time. Of course, neither the deontological principle nor the teleological principle could be regarded as the only principle relevant to cases of the sort we are discussing. Freedom is not all that counts. But the point remains that a deontological principle weighs against every paternalistic interference with the freedom of the moment, whereas the teleological principle weighs against some such interferences and supports others. The teleological principle will conflict less often with impulses to paternalism from other sources.

It might be objected that in substituting the teleological "freedom-maximization' principle for the deontological "freedom-respecting' principle, we are suggesting an unacceptable "utilitarianism of rights." Whatever the general force of this criticism, it is surely blunted here by the response that we are not invading one person's rights to enhance another person's rights. Rather, we are invading one person's right (if we concede so much to the objection) to enhance that same person's total enjoyment of the same right over time.⁷ In this essay, at least, I do not recommend interpersonal freedom-maximization.

I have introduced the freedom-maximization approach in the context of the slavery case because Mill's discussion of that case so clearly suggests it. I do not claim that once we have formulated the freedom-maximization principle, we have said all there is to say about the slavery case. A full discussion of contracts for slavery would take us far beyond the topic of paternalism. Let us therefore turn to a simpler problem.
Consider cigarette-smoking. I am inclined to think that it would violate no one's rights (and that it would probably be a good thing) if the sale and use of cigarettes were prohibited. It may be that a ban on cigarettes can be justified by the arguments considered in the first section of this paper. Many cigarette-smokers remain ignorant about the risks of smoking. Many smokers who are aware of the risks in some sense do not imagine them vividly or do not respond appropriately given the small probabilities involved. Many smokers acquired the habit under social pressure. And so on. But surely there are some smokers who are covered by none of these arguments. Perhaps we ought really not to interfere with those remaining smokers, or my ideal paternalist with his powers of perfect discrimination ought really not to interfere with them. I have a lingering feeling, however, that it may be permissible to prevent cigarette-smoking even by a smoker who has no family, who is as clear-headed and as free of neuroses as a person can be, who is well informed about the chances of getting cancer or emphysema and the general diminution of life expectancy, who has seen close-up the effects of cancer and emphysema, and who just doesn't give a damn.

The freedom-maximization approach allows us to explain why it is permissible to coerce the well-adjusted, well-informed, would-be smoker. To be sure, the smoker does not destroy her freedom completely, as Mill's would-be slave is trying to destroy his. But cigarette-smoking will diminish the smoker's freedom, at least statistically speaking. It will shorten her life-expectancy and increase the likelihood of debilitating disease. (The appeal to statistics is somewhat inconsistent with my intention to consider the problem of paternalism from the viewpoint of an omniscient ideal paternalist. I would concede that if the ideal paternalist can separate those smokers who will suffer bad effects from those who will not, he should coerce only the former. Since we in the real world are unable to make this distinction, it is convenient to talk in terms of statistics and in terms of all smokers suffering a statistical harm, although this raises problems that I am going to ignore.) It seems to me that the expected loss of freedom caused by smoking (taking into account both the magnitude of the various possible losses and their probability) is greater than the loss imposed directly by the prohibition on smoking.

It could be objected that smoking does not cause any loss of freedom. Disease and death are not ordinarily regarded as "unfreedom." Whatever their ill effects, the issue of freedom is not involved. My answer to this is that we are not trying to minimize unfreedom but to maximize freedom, and what I mean by "freedom" in this connection includes abilities, capacities, and in general whatever is a precondition for any human activity. What we desire is that the largest number of people should have the widest possible range of effective choice about what to do with themselves. From this point of view, it is clear that death and injury and disease all diminish freedom.

The objector might persist, with the suggestion that a loss of freedom from disease, say, is less important than an equal loss of freedom attributable to direct paternalistic intervention. I do not necessarily deny this. I shall suggest a reason
shortly why it might be true. For the moment, I would say that the expected loss from smoking is so much greater than the loss from paternalistic intervention to prohibit smoking that even if the latter is given some extra weight to reflect the fact that it flows from external intervention and not from the natural consequences of chosen actions, the balance still favors the paternalistic course.

An obvious difficulty with the freedom-maximizing approach is that it assumes freedom is quantifiable, at least to some extent. It assumes that we can compare “bundles” of freedom, and (barring omniscience) lottery tickets with bundles of freedom as payoffs. Now, I cannot give complete directions for comparing bundles of freedom, or lottery tickets, but I can say a few things about how freedom is measured. First of all, it is clear that we do not determine the extent of a person’s freedom just by counting up the actions available to her at all relevant times and saying that the greater the number of actions, the greater her freedom. For one thing there is no obviously satisfactory criterion for individuating actions. More important, however, is the fact that freedom to do some things is much more important than freedom to do others. Any criterion based on mere counting would ignore such differences. In deciding how great a person’s freedom is, we need to consider not merely how many different things she can do, but what the things she can do are. Freedom to do X will presumably count for more than freedom to do Y whenever X is more pleasurable to the particular individual than Y, or more highly valued by a rational individual than Y, or more essential to the individual’s sense of personal identity than Y, and so on. The considerations just listed do not exhaust the possibilities, and each must be taken as including the qualification “other things being equal,” if only because the considerations listed might conflict for particular values of X and Y.

It may seem that in the last paragraph I replace a nearly hopeless problem (how to count up actions) with an utterly impossible one (how to evaluate bundles of freedom in terms of the relative importance, according to various criteria, of the available actions). I do not think that is the case. We are no closer to a well-defined procedure for ranking bundles of freedom, but the new problem is more amenable to acceptable intuitive judgments than the old one. If the criterion for ranking bundles of freedom is simply the number of available actions, my intuition says that almost all the bundles that arise in practice are going to contain the same infinite number of actions, and therefore be equally valuable. My intuition also says, however, that the conclusion that almost all bundles of freedom arising in practice are equally valuable is ludicrous. Once the criterion is expanded to include reference to the importance of the actions, I find that I can make some intuitive judgments, like the one I have already revealed regarding where the “balance of freedom” lies if we are considering prohibiting cigarettes.

The notion of freedom-maximization may require judgments that are controversial and incapable of definitive establishment, but in this respect it does not differ from the notion of utility-maximization. Many people have been utilitarians, and many more have agreed that utility-maximization should play some role in moral decision-making, despite uncertainty about just what actions
maximize utility, or even just what utility-maximization means. If the notion of freedom-maximization strikes a responsive chord from somewhere among our moral intuitions, perhaps we should struggle along with it just as we do with other notions equally vague.

The freedom-maximization idea seems to me useful in other cases besides the slavery case and the case of cigarettes. Laws against drugs like heroin and laws requiring the use of seat-belts in cars or the wearing of helmets when riding motorcycles present much the same problem as laws against cigarettes. My intuition is that, in each of these cases, paternalistic intervention is likely to maximize freedom overall.

I do not suggest, however, that we would be justified in forbidding all risky activities. Consider mountain-climbing. Although there are substantial risks involved in mountain-climbing, the freedom that would be lost if mountain-climbing were forbidden looms much larger, to my mind, than the freedom that is lost if cigarettes are prohibited or seat-belts required. For one thing, mountain-climbing is likely to be much more important to people who want to climb mountains than cigarettes are to people who want to smoke cigarettes. Climbing is likely to be a source of greater pleasure, especially when we consider not merely the time spent on the mountain, but time spent planning and preparing for trips, talking to other climbers, and so on. Also, climbing is more likely to be closely linked with the would-be participants’ sense of identity. A person might describe herself as a mountain-climber in the way that someone else would describe herself as a chess-player, a gardener, or a philosopher. Who would think of describing herself as a cigarette-smoker in the same way?

Beyond all of that, I am inclined to think that mountain-climbing is intrinsically a more valuable activity than cigarette-smoking. It is no accident that people think of themselves as mountain-climbers in a way they do not think of themselves as cigarette-smokers. There is something about the activities themselves that accounts for this difference in attitude. Indeed, while mountain-climbing is plainly an "activity," that word seems out of place as applied to cigarette-smoking.

The claim that some activities are intrinsically more valuable than others is controversial. I lack both the space and the arguments to defend the claim here, although I hope to defend it at some time in the future. Nonetheless, it is part of what accounts for my differing reactions to mountain-climbing and cigarette-smoking; I suspect it contributes to some other people’s differing reactions; and I suggest that the reader who reflects on the ways we talk about mountain-climbing and cigarette-smoking will be led to see some of the differences that I think make the one intrinsically more valuable than the other.

One other problem that invites attention in a discussion of freedom-maximization is the problem of suicide. In all probability, most of my readers will share the view that suicide should not be unconditionally forbidden. And yet it would seem that there can be no act more destructive of freedom than suicide. If we allow suicide, what has become of freedom-maximization?
People want to commit suicide in various circumstances, and I have no room here for a detailed discussion of the problem. Speaking briefly and generally, there are two points to be made on how freedom-maximization may be reconciled with allowing suicide. First, in any case where the would-be suicide's desire to commit suicide persists (and I am prepared to admit that it is only in such cases, if we can identify them, that suicide should be allowed), the freedom that will be preserved by forbidding suicide will be substantially devalued by the possessor's lack of interest in it and lack of inclination to put it to good use. Second, the person who is prevented from committing suicide (and whose desire persists) is likely to be very unhappy during her remaining life. Considerations of utility-maximization argue in favor of allowing the suicide, and may simply outweigh considerations of maximizing freedom. In other cases, we have often implicitly assumed that our paternalistic impulses were prompted by utility-maximization, so we have been arguing in effect that freedom-maximization and utility-maximization pointed in the same direction. But they need not always do so, nor need either always prevail over the other when they conflict.

One final observation: Although I have been discussing freedom-maximization as a principle which focuses on maximizing the range of choice available to an individual over time (restricting choice now in order to preserve opportunities later on), the same phrase, "freedom-maximization," could encompass an idea I mentioned in the preceding section of the essay, namely, the idea that we want to allow, or perhaps even help, individuals to develop the ability to make choices. We are interested both in people's having opportunities for choice and in their having capacities for making well-thought-out choices and sticking to them. Unfortunately, a paternalistic intervention that will maximize opportunities for choice in the long run may also interfere with the development of the ability to choose.

I do not conclude that we should abandon range-of-choice-maximization. In cases like the cigarette case, for example, the potential for an individual to learn from her own experience is small, just because the major costs of smoking are likely to appear long after the important initial decisions. Even here, of course, one person may learn from the experience of another, and for any individual who will not smoke, it is almost certainly better, on "ability-maximizing" grounds, that she make the decision and stick to it by herself than that she do so with paternalistic assistance. In general, this ability-maximizing aspect of freedom-maximization may account for the feeling noted earlier that a loss of freedom as a natural consequence of some earlier choice is not as objectionable as the loss of an equal amount of freedom, in the range-of-choice sense, as a result of paternalistic intervention. The first loss is more likely than the second to be educational. The point is just that the idea of freedom-maximization is complex. Even so, it is a way of giving content to the value (or values) of freedom that we cannot ignore.
III

At the end of the first section of this essay, I suggested that one weakness of traditional arguments about paternalism is an implicit assumption that people do not change over time. In this section and the next, I shall explore some arguments that take change over time specifically into account.

The central tenet of most arguments against paternalism is Mill’s proposition 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.’’

Paternalism presents a problem because it involves exercising power over an individual, not to prevent harm to others, but to prevent harm to the individual herself. If there were some good reason for regarding the harm done as done to a person other than the agent, the problem would disappear. In many cases in which paternalism seems justified there is a good reason for thinking of the harm as done to someone other than the agent. To illustrate that reason, it will be useful to consider a few cases, starting with one that has nothing directly to do with paternalism.

Suppose that, ten years after the occurrence of an act of embezzlement, we finally discover the identity of the embezzler. Suppose also that we hold a retributivist view of punishment. A crime has been committed. We have at last identified the criminal. It would seem that she should be punished. But when we consider the person before us, the ‘‘criminal’’ we are supposed to punish, we discover that she is a different person from the person she was ten years ago. She squandered all her ill-gotten gains in the first six months. Since then she has lived a blameless life. She has punctiliously fulfilled all obligations of trust. She has not (in this hypothetical case) repaid the money she originally embezzled, but that is because her blameless life has been a modest one, and she has had no funds to spare. I think we would be most reluctant to punish in this case. Although the criminal was (and perhaps timelessly is) deserving of punishment, the criminal is no longer accessible to us. Inhabiting the criminal’s body and social role, we find a new woman.

Consider next a slightly different case. Imagine a smoker who smokes for twenty years, then quits for ten, and then turns up with lung cancer, of which his smoking is a causal antecedent. Would it not seem to us that nature has been unfair? Perhaps someone who has smoked for twenty years has no ground for complaint if he gets cancer, but in this case the person who gets the cancer is an abstainer of ten years’ standing. That ought to count for something.

To be sure, accusing nature of unfairness is only a manner of speaking, but it is a manner of speaking that suggests a bridge between the problem of punishment and the problem of paternalism. It seemed unfair to punish the reformed embezzler. It seems ‘‘unfair’’ to ‘‘punish’’ the reformed smoker. The reason is that in each case the person who suffers is not the person who (arguably) deserves to. But if the reformed smoker is not the same person as the unrepentant nicotine-fiend, that may be relevant to the issue of paternalism. If we step in and prevent the
would-be smoker from smoking, can we not claim that we are protecting a different person, the smoker's later self? (It may occur to the reader that I am skipping over some problems involving the question of whether, or under what conditions, the smoker will or would reform. I shall discuss these problems presently.)

Actually, I think the suggestion that the smoker becomes a different person is an essential prop for a more standard line of argument in favor of paternalism, at least if the standard line is to amount to anything more than straight utility-maximization. Consider, for the sake of variety, another case in which a cyclist is bound by a statute that makes it an offense to ride a motorcycle without wearing a helmet. Such statutes are frequently attacked on the ground that if the cyclist wants to run the risk of serious injury, that is her own business. A defender of the statute is likely to respond as follows: "Anyone who rides a motorcycle without a helmet risks serious injury. If she is seriously injured, she is likely to become a public charge. She will be cared for in a public hospital or, even if she can afford private care, she will end up unemployed and drawing public compensation. This may not happen in every case, but certainly in statistical terms the helmetless cyclist imposes a burden on public assistance funds. Since public funds must be raised by taxation, the helmetless cyclist hurts someone besides herself."

The argument just stated is not very satisfying. For one thing, when we consider what the statistical burden the cyclist imposes on the public treasury comes to, it may well be that the harm the cyclist does to others by this route is outweighed by the utility to her of riding without a helmet. Further, the tenuousness of the connection between the conduct and the harm gives the argument something of the false ring of rationalization. In any case, the opponent of the helmet statute, in order apparently to avoid the force of the argument, has only to steel himself and say: "You go too fast. You say that the cyclist will be a burden on public assistance funds, but the cyclist never asked for public assistance. The cyclist I have in mind values her freedom, and she realizes that the price of freedom is to suffer the consequences of her choices. If she suffers a serious injury, leave her to manage as best she can. Leave her to private charity, or let her die in the street. So long as you are prepared to do that, her riding without a helmet doesn't hurt anyone but herself."

At this stage, the defender of the statute might reply: "We can't leave her in the street. That would be inhuman. It would cost us more in suffering to leave her in the street than it would cost to care for her properly. So you see, she has harmed us, either way." The obvious retort is: "If you value freedom at all, you must admit that one person's mere emotional distress at another's behavior is no justification for making that behavior a crime. If the cyclist insists on dying in a public thoroughfare, let us remove her, as we would any other offensive exhibitionist. But so long as she is out of the public eye, she is as entitled to die as to read a dirty book."

In rebuttal, the defender of the statute will say: "It's not simply a matter of squeamishness that makes me want to help the injured cyclist. I have a moral
obligation to. Being denied assistance when one is injured is a punishment too
great to visit on anyone’s head just for making a foolish choice, even if the choice
was precisely to risk that punishment. The cyclist may have made her original
choice with full knowledge, but she must regret it now. My general duty to help
people in need would be satisfied here only if the cyclist did not want help at the
time she needs it. The fact that she decided to do without help before she needed it
is quite irrelevant.’’ The defender of the statute adds that being put in a position
where one must undertake some burden or expense if one is to satisfy a moral
obligation (as opposed to reacting on the basis of feelings of pity or horror) is harm,
so the cyclist has harmed someone after all.

The defender of the statute is now in a strong position. He may have won the
argument. What remains to be observed is that if he has won the argument, he has
done so by hitting on a suggestion that makes most of the argument as I have
described it superfluous. If we really have an obligation to help the injured cyclist
regardless of whether, when she decided to ride without a helmet, she expected
help or wanted it, the most plausible explanation is that the cyclist before the
accident and the cyclist after the accident are in some sense different people. That
is the simplest explanation of why an initial willingness to forego aid is not
definitive. But if the cyclist is a different person at the later time, then the cyclist at
the time of her original decision has harmed another person. The decision to run
the risk was not the original cyclist’s “own” business.

It is all very well to suggest that the embezzler, the smoker, or the cyclist
may each be different people at different stages of their lives. What exactly makes
them different? To say that it is mere passage of time—to say that each of us is a
different person on every day of her life—would completely subvert our ordinary
notions of personhood. If the mere passage of time does not make the difference,
what is it? How do we decide whether the embezzler, for example, has changed in
such a way that she no longer deserves to be punished?

I would say, roughly, that the embezzler is a different person when we
discover her ten years later if she is no longer the sort of person who would
embezzle if placed in the same situation in which she did so originally. What this
means, of course, is far from clear. Many moralists, concerned about preserving
freedom of the will, might want to hold that the question “Would Jones, if placed
in the following situation, embezzle?” often has no well-defined answer. My
suggestion about retributive punishment seems to require that the question “Is
Jones the sort of person who, if placed in the following situation, would
embezzle?” should have an answer. In that case, the second question, about what
sort of person Jones is, cannot be the same as the first question, about what Jones
would do. Perhaps the second question is about what Jones would probably do, or
would be strongly disposed to do, or might do without greatly surprising those who
knew Jones well, or something along those lines.

Assuming that further reflection would provide us with a satisfactory sense
of “the sort of person who . . . ,’’ we now observe another important point. Our
embezzler may, after ten years, be the same person for some purposes and yet be a
different person for other purposes. Thus, suppose that at the time she committed embezzlement, the embezzler also committed an unrelated aggravated assault. It is at least conceivable that after ten years the embezzler has grown much more conscious of duties of trust without becoming any more disposed to control a volatile temper that produces occasional physical aggression. In such a case, I think we might hold it inappropriate to punish the embezzler now for her embezzlement, but appropriate to punish her for the assault. 11

What about the cyclist who rides without a helmet? What makes her a different person after her accident? The answer, I think, is that the cyclist is a different person, in the relevant respect, if she is no longer the sort of person who would ignore her future well-being for the sake of small increments of present utility. Of course, it is not certain that having the accident will produce any such change in the cyclist. But it seems likely to. In many cases, I should think, the cyclist will not merely wish she had behaved differently in the past. She will have a new appreciation of the virtue of prudence and will alter her attitude toward risk in the future. If the cyclist changes in this way, she is a different person, who deserves protection against the foolish behavior of her earlier self.

The reader may have noticed something odd about that last sentence. If the cyclist learns prudence as the result of an accident, she may then be a person who deserves protection, but it is too late to protect her. The harm she deserves to be protected against has already occurred. Looking at the matter from the other end, if we compel the imprudent original cyclist to wear a helmet, the other, prudent, cyclist may never come into existence. Can we really interfere with the first cyclist for the benefit of a later cyclist who may never exist and whose existence our interference is intended to make less probable?

There are a number of questions here. Some of them arise because we do not know whether the cyclist will be injured if she rides without a helmet and whether she will become a new person if she suffers injury. In this essay I am generally ignoring such questions. I am interested in how the situation looks to an ideal paternalist. The ideal paternalist does know whether, if the cyclist rides without a helmet, she will be injured, and he does know whether, if she is injured, she will develop a new attitude to risk as a result.

Now, on the theory I am suggesting, there is no ground to interfere with the cyclist unless she will both be injured and be changed by the experience. 12 Suppose that the ideal paternalist is confronted with a cyclist who will be injured, if she rides without a helmet, and who will be changed. Can the ideal paternalist intervene? The answer is still not obvious. We assume that if the paternalist intervenes, the cyclist will not be injured and will not change. But this means that if the paternalist intervenes, he will not be protecting any actual person. He will be protecting a possible later self of the cyclist, whom he prevents by his intervention from becoming actual. If he doesn’t intervene, the possible person he could protect will exist and will suffer the effects of a serious accident. But if he does intervene, she will not exist at all.
I think the ideal paternalist can intervene to protect the prudent cyclist, even though she is only a possible person, whose existence the paternalist’s intervention will prevent. What is appropriate to do on behalf of possible persons is a complicated issue, and one that I shall not discuss at length even though perhaps I should. There are two points that I think distinguish this case from others in which intervention on behalf of possible persons seems more problematic. First, the possible person the paternalist can protect here is one who will exist and suffer unless the paternalist interferes with another agent’s decision. This is quite different from, say, the case of forbidding birth-control, where the possible person on whose behalf we might act will not exist unless we interfere with the potential parents’ choice. It seems that we have more reason to act on behalf of possible persons who will exist without our help than to act on behalf of possible persons who need our help to exist in the first place. Second, the possible person we are concerned with in the cyclist case, if she exists, will be connected by physical continuity with an actual person (the original cyclist) who exists regardless of the paternalist’s decision. The possible person we are concerned with in this case is unusually well-rooted in actuality, as possible persons go. She deserves protection.

One final point merits brief discussion. Let us refer to the view I have been expounding, on which different temporal stages of one physically connected “person” may be different persons for certain moral purposes, as the “time-slice view.” It might be urged against the time-slice view that agents have a greater right to impose risks or harms on their future selves (or time-slices) than to impose risks or harms on unconnected others. It is worth noting that this can be true without destroying the force of the time-slice view as a defense of paternalism. Most people think that there are situations in which it is wrong to impose harms on unconnected others, even though the act imposing the harm would maximize utility. If this is correct, it is a stronger principle against harming others than I think the time-slice approach justifies with regard to one’s future selves. I would not suggest that an agent must not harm a future self even if overall utility is thereby maximized. I would suggest only that an agent should not harm a future self and diminish overall utility into the bargain, and that an agent who threatens this may be prevented. Even that is enough to justify more paternalistic intervention than is justified on the view that the agent should be left alone so long as she harms no one but “herself,” present or future.

It might be suggested, of course, that the position I have just indicated is untenable. The argument goes as follows. If we believe that one person may not harm another even though she thereby maximizes utility, and if we genuinely regard future selves as other persons, we must be prepared to protect a future self from harm by a present self even though the present self would maximize utility by her choice. If we are not prepared to protect future selves in this way, we do not truly regard them as other persons. That this argument is correct is not self-evident. Might there not be different degrees of otherness, with different moral consequences? If that is held to be impossible—if all others must be treated the same as
unconnected others—I think there is as much to be said for the view that unconnected others should not be protected against acts that harm them but that maximize utility, as there is to be said for the view that unconnected others deserve a greater protection than this and that future selves deserve either the same greater protection against the agent’s present self or no protection at all.

IV

In the preceding section I suggested that we could prevent an imprudent cyclist from riding without a helmet and harming a later, prudent, self. In effect, I suggested that the earlier cyclist’s choice should not be held to bind the later cyclist. It is the inability of the earlier cyclist to bind the later that opens the way to “paternalistic” interference.

There is a different sort of case in which it is often suggested that the justifiability of paternalistic interference depends precisely on the fact that a person can bind herself. A standard example is the story of Odysseus and the Sirens. I shall argue that attempts at self-binding are no more effective in cases of this sort than they were in the cases of the preceding section. Paternalistic interference that can be justified only by reference to self-binding decisions is not justified at all.

Whereas the general tendency of the two preceding sections was to provide new arguments in favor of paternalism, the general tendency of the present section (not necessarily the only tendency, as we shall see) is to undercut an accepted argument for paternalism.

Consider Odysseus. In order to be able to hear the Sirens’ song without being tempted to his death, Odysseus commands his crew to bind him to the mast of his ship and not release him until the Sirens have been passed and left behind. It is often suggested that even if, with the Sirens’ song in his ears, Odysseus pleads to be released, the crew are justified in ignoring his pleas because of his original command. In brief, Odysseus’ original command is regarded as binding him, and as justifying the crew’s protective action, until the command expires by its own terms.13

I agree that Odysseus’ crew are right to keep him tied to the mast, even if we consider only Odysseus’ interests and not the crew’s need for a captain. But if we ask why the crew are right, Odysseus’ command has very little to do with it. The command has some relevance, for reasons I shall discuss below, but it is neither a necessary nor a sufficient condition for the rightness of the crew’s behavior.

To see that the command is not a necessary condition, we have only to suppose that Odysseus has not given the command, believing that he can sail by the Sirens, listen to their song, and preserve himself by sheer strength of will. If Odysseus has attempted this and is mistaken, and if the crew see him poised to jump overboard and swim to his death, surely they are justified in stopping him, tying him to the mast if need be, with no by-your-leave.
To see that the command is not a sufficient condition, we must alter the facts a bit more. Suppose that the Sirens' song casts no irresistible spell, but is merely extremely beautiful. Suppose that the song is most beautiful when heard from near the Sirens' isle, which is surrounded by dangerous rocks. And suppose that if Odysseus swims to the rocks, he will not be killed but will suffer some physical injury. Surely if we bring all these assumptions together in the right relation, we can imagine a situation in which either swimming to the Sirens or not swimming to them would be a choice a reasonable person might make. Suppose now that Odysseus, approaching the Sirens' isle, decides he wants to sail by. Not trusting himself to stick to this decision, he commands the crew to tie him to the mast, as in the standard version of the tale. When he hears the Sirens' song, however, he changes his mind, not because of any sinister compulsion in the Sirens' song, but just because the song is more beautiful than he imagined, or because he discovers that he cares more about hearing it than he realized. If we remember that the decision to swim to the Sirens can be a reasonable one, then surely Odysseus is entitled to change his mind. (If we like, we can hypothesize further that Odysseus will have no other opportunity to hear the Sirens' song. It is now or never.) The crew ought to release him, despite the earlier command.

If what I have just said is correct, the command is neither necessary nor sufficient to justify the crew's decision to keep Odysseus aboard. What justifies their decision in the standard version of the tale is the fact that swimming to the Sirens is understood to be an irrational choice, with fatal consequences, made under preternatural compulsion.

I shall say more about why Odysseus is entitled to change his mind, but first I should tie up a loose end. I mentioned earlier that the command has some relevance to what the crew should do. One reason is that, in practice, the crew do not know Odysseus perfectly. Even if swimming to the Sirens would be reasonable for some people, the crew must make a decision, when Odysseus asks to be released, whether it is reasonable for him. We could understand their supposing that what Odysseus thought before he was tied to the mast is important evidence concerning what is reasonable for Odysseus now. In addition, if the crew decide wrongly not to let Odysseus swim away when he wants to, they are less to blame if there was an earlier command than if there was not. They can be more easily forgiven for choosing incorrectly between Odysseus' inconsistent commands than for restraining him improperly on their own unprompted initiative.

If we set out to explain more fully why Odysseus is entitled to change his mind, provided that his new decision is a reasonable one, we encounter a difficulty. To speak of Odysseus' changing his mind seems to assume that Odysseus is the same person throughout the episode we are considering. The discussion of the preceding section, however, suggests the possibility that in "changing his mind" Odysseus is really changing his identity. Does it matter how we describe the change in Odysseus? Ultimately I do not think it matters, in this kind of case, how we describe the change in Odysseus. I think the later choice, whether it be of the original Odysseus or of a new one (and provided always that it
is a reasonable choice at the time it is made), should control. But the fact that we can describe the change in Odysseus in two different ways complicates the issue. Some people, I predict, will disagree with my conclusion and will believe both that Odysseus does not change identity and that his not changing is an essential part of the explanation why his first choice should control. Others will disagree with my conclusion but will believe that Odysseus does change and will believe that his changing is an essential part of the reason that his first choice should control. I shall deal with both positions in turn.

First, on the assumption that Odysseus is the same person, is he entitled to change his mind? Many people seem to have the intuition that someone may reasonably want to bind herself for the future by a present decision, and that provided we really are talking about a single person—that is, about a present person binding a future person who is still herself—such self-binding decisions ought to be recognized and, where appropriate, enforced.

My intuition runs the other way, and I see two general reasons that support the freedom to change one’s mind. One reason is that, other things being equal, the later decision is likely to be a better one. The agent is likely to have more information and to have had more time to reflect on her goals. To be sure, the first decision may have been better. It may have been more carefully considered, or the agent may simply have forgotten something important in the time between the decisions. Nonetheless, in the absence of extrinsic evidence about which decision is more deliberate or better-informed, the mere passage of time suggests considerations which favor the second.

In addition, allowing changes of mind will tend to develop strength of purpose, part of which is precisely the ability to resist vacillation without outside help. I have suggested before that we might sometimes want to let an individual do something foolish, which would harm her in the future, in order to encourage her to develop the ability to make good choices. I now suggest that we may want to inform the agent that we will not, despite her present request, undertake to prevent her doing something foolish, or rather something she currently regards as foolish, in the future. She should master herself. Her fate is in her own hands.

The reasons I have suggested for allowing changes of mind may seem significant but not overwhelming. Is there anything more to say? It is natural to look for examples, discussion of which will move possessors of the conflicting intuitions closer to agreement. Unfortunately, it is not easy to produce examples that shed light on the question. Most examples that present the question seem merely to present it, without illuminating it. Perhaps it is worth mentioning some legal examples, since they are a sort of example about which we have codified views (though not necessarily correct ones). We allow people to revoke or rewrite wills. We allow people to retract offers before they are accepted. Even after an offer has been accepted and a contract formed, we allow the contracting parties to modify the contract or rescind it by mutual agreement. Legislatures, persons of a sort, can repeal legislation. I could go on.
Unfortunately, listing cases in which we allow changes of mind suggests a similar list of cases in which we do not. If there are revocable licenses, there are irrevocable licenses also, often in the form of easements. There are revocable, but also irrevocable, trusts. Gifts, once completed, are irrevocable by the donor. Offers become contracts upon acceptance, binding the offeror unless she is released by the offeree. On a higher level, one might suggest that a constitution embodies decisions by which the whole body politic binds itself.

If we ask what distinguishes decisions that are revocable from decisions that are not, the proponent of the intuition that favors self-binding is apt to suggest that the crucial factor is precisely whether the decision is intended to be self-binding. When a decision is intended to be self-binding, it is.

The truth of the matter is complex, and whether one intends to bind oneself is no doubt relevant to whether one actually does so. But I think an intent to bind oneself is far from a sufficient condition for self-binding. I cannot discuss all the examples at length. I shall comment briefly on two of them.

Regarding constitutions, the key point is that a constitution is not ordinarily intended to bind (or at least does not ordinarily bind) the same entity that adopts it. A constitution is adopted by the people at large for the purpose of binding the organs of day-to-day government. The people at large, who adopt the constitution, can always change their minds by amending it.\footnote{\label{footnote}Contracts are different. A contract does bind the agent who enters into it. It is tempting to say that a party to a contract is bound only because someone else’s interests are involved. If this were a fully adequate explanation of the bindingness of contracts, the contract case could be dismissed as bearing no analogy at all to the case of Odysseus. Unfortunately, I do not think the proponent of Odysseus’ freedom to change his mind can rest here. The bindingness of a contract is not fully explained just by saying someone else’s interests are involved. Someone else’s interests are involved because people rely on contracts. People rely on contracts in part because contracts are known to be binding. In short, bindingness and reliance are a package, and must be explained together. Not until we have the right explanation of the package can we be sure that the reason contracts bind, whatever it is, does not cover Odysseus’ decision as well.

The believer in effective self-binding might now suggest that we make contracts binding because the agent who enters a contract benefits from being able to bind herself. Her ability to bind herself is the prerequisite to her getting something she wants. In the contract case, to be sure, being able to bind herself is a prerequisite to getting something she wants from the other party. But there might be cases in which ability to bind herself is a prerequisite to her getting something she wants even though no other party is involved. Remember Odysseus: When he asks to be bound to the mast, he wants not to swim to the Sirens. The only way he can be certain of achieving this is by having himself bound and having a crew who will ignore any change of mind. If we make contracts binding so that people can get what they want from other parties, why not make purely self-regarding decisions binding (when they are intended to be binding) so that people can get what they want (at the time of the decisions) from themselves?
One possible answer to the question just posed is that there is a difference in the way self-binding is "necessary" in the contract case and in the self-regarding case. Odysseus *could* get what he wants when he tries to bind himself simply by being strong-willed and by not changing his mind. In the contract case, it is not enough that one not change one's mind; the other party needs to *know* that one will not, or cannot effectively, change one's mind.

There is another difference between the two cases, not necessarily deeper, but more interesting. The standard contract case we call to mind to justify our belief that contracts should be binding simply does not involve a change of mind, or even a threatened change of mind, of the kind that occurs in the Odysseus case. I shall explain: Generally speaking, the reciprocal obligations of a contract are performed by the parties at different times. (This is not invariably true, but if contractual performances were always exchanged contemporaneously, we would have much less use for contracts than we have.) The most obvious case in which one party would like not to be bound by her contract is the case in which the other party has performed and she has not. But here it is entirely the time-displacement that accounts for this party's "change of mind." There has been no change in her beliefs or her interests. She just wants to get something for nothing. If we view the transaction as a whole, it is as much in her interest at the later time when she wants out as it was when the contract was made. The question of whether she should be able to "change her mind" does not arise in the form in which it arises regarding Odysseus.

To be sure, the desire to breach one's contract *may* result from a genuine change of mind about the transaction as a whole. The party who wishes to breach may be willing to return what she has received under the contract, if that is possible. In this case, if we hold her to the contract, we will prevent an Odysseus-like change of mind. But it is not a desire to prevent Odysseus-like changes of mind, and it is not this case, that leads us to accept the idea of binding contracts in the first place. If people's beliefs and preferences were generally so volatile that Odysseus-like changes of mind were the commonest reason for wanting to break contracts, we would have a very different doctrine of contract from what we have. We would enforce many fewer contracts. Instead of encouraging reliance, and protecting it by making contracts binding, we would save people from relying to their detriment by making it clear in advance that enforcement was not the norm. The doctrine we have, and our standard intuitions about contract, depend on the fact that in ordinary cases parties do not change their minds *in the way Odysseus does*. For that reason, the contract analogy does not suggest that Odysseus should be forbidden to change his mind.

The lesson of the legal examples is ambiguous, but I believe that in the end these examples neither undermine my arguments in favor of freedom to change one's mind nor cast doubt on the intuition that freedom to change one's mind, even after a decision intended to be binding, is the "normal" state of affairs.

I have been arguing that Odysseus should be able to change his mind, provided his later choice is a reasonable one, on the assumption that he is at all
relevant times the same person. I turn now to the suggestion that Odysseus is not the same person if he changes his mind, and to the claim that, precisely because Odysseus is not the same person, his later self should not be able to undo what his earlier self hoped to accomplish. Actually, this suggestion seems quite implausible in the case of Odysseus. I shall therefore consider a somewhat different example, in which the suggestion has more appeal. The example is Derek Parfit's: "Let us take a nineteenth-century Russian who, in several years, should inherit vast estates. Because he has socialist ideals, he intends, now, to give the land to the peasants. But he knows that in time his ideals may fade. To guard against this possibility he does two things. He first signs a legal document, which will automatically give away the land, and which can only be revoked with his wife's consent. He then says to his wife, 'If I ever change my mind, and ask you to revoke this document, promise me that you will not consent.'" I shall call the Russian "Boris." Suppose that as Boris gets older, his ideals do fade. When he inherits the estates, he wishes to keep them. He asks his wife's consent to the revocation of the original document. What should his wife do? Parfit suggests that we can view the older Boris as a different person from the idealistic young Boris who signed the document. If we do, the older Boris's request cannot release the wife from her commitment to the young Boris. The young Boris no longer exists. There is no living person by whom the wife can be released.

Parfit suggests that the commitment is unreleasable. He does not say that because the commitment is unreleasable it must be honored. Parfit is primarily concerned, at this point in his essay, to establish the possibility of viewing the older Boris as a different person from the younger. The general tendency of his essay, as I read it, would support the weakening of all obligations based on commitments. Nonetheless, it could be suggested that the unreleasable commitment to the young Boris is a binding commitment, which should lead the wife to refuse the older Boris's request. This view I shall call the "rigorist" view.

Note that the rigorist view presupposes Parfit's approach to personal identity (which I share) and applies to cases in which there is an actual change of identity between an earlier and a later occupant of the same body. The rigorist view is quite consistent with the view that so long as there is no change of identity, an agent should be able to change her mind and to release others from commitments to her. Thus, one who holds the rigorist view may believe that in our earlier example, Odysseus can change his mind. This raises a question. Suppose that Odysseus says to his crew: "Bind me, and do not release me. If I ask to be released, ignore me. The request to be released will be proof that I have changed, and that my then-existing self cannot release you from the commitment you now give me." If the crew must accept Odysseus' claim that any change of mind would represent a change of identity, the rigorist view entails that Odysseus can always bind himself firmly, so long as he uses the right formula. I take it, however, that if we accept the rigorist view, the question of whether the original Odysseus has been replaced by a new one is not to be settled simply on the basis of what Odysseus at any time says.
What Odysseus says is significant. His assertions about which of his beliefs or preferences or ideals are essential to his identity cannot be ignored. But his assertions, even his *sincere* assertions, are not conclusive.

So much for description of the rigorist view. Is it reasonable? Why, when the older Boris asks his wife to revoke the document, should she care about her commitment to the young Boris at all? The young Boris no longer exists. He can neither enjoy nor even be aware of the satisfaction of his preferences. Why should his preferences, or the commitment to him, count?

In response we might note that people frequently have preferences about what happens after their death or, more generally, about what happens without their knowledge. It is not easy to explain why the satisfaction of such preferences should not be regarded as valuable without falling into a theory which locates all value in mental states. I shall not argue that the preferences of the "deceased" young Boris should not count at all. On the other hand, they might count without counting as heavily as the preferences of the Boris who is still alive. This would tend to undercut the rigorist view. I think there are reasons for devaluing the young Boris's preferences, at least in this case.

Parfit suggests by implication that the wife's promise to the young Boris is like a deathbed promise to a parent to help his or her children. Certainly the promises are alike in that the promisors cannot be released. But they are very different in another respect. The promise to the dying parent is to help the children. The promise to the young Boris is to *oppose* Boris the Old. (It might be suggested that the promise to the young Boris is a promise to help the peasants. In a general analysis of whether to keep the promise, that would surely be relevant. But it does not eliminate the difference between the promise to the parent and the promise to Boris that I wish to focus on.) On the face of it, a deceased's preference that the goals of his or her survivors (which we assume are reasonable and morally permissible goals) be opposed seems entitled to less weight than a deceased's preference that the goals of her or his survivors be furthered.

Young Boris is unlike the dying parent because there is a conflict between the goals of the younger Boris and the older. Suppose that we ask directly which Boris ought to prevail. To say that the older ought to prevail just because he is alive might beg the question, although I am not sure it would be a mistake. There may, however, be another way to look at the matter. It would not be implausible to suggest that if the wife honors her promise to the young Boris, she will harm the older Boris. "Harm" is a problematic concept. It might be said that merely not inheriting vast estates cannot be a harm. On the other hand, if we say that the older Boris is deprived of his patrimony, that does sound like harm. At least, the use of the word "harm" is not implausible. It is much less plausible, however, to say that if the wife gives in to the older Boris, she *harms* the younger. Even if postmortem preferences (preferences about what happens after one's death) count, it does not seem that the dissatisfaction of one's postmortem preferences can be a harm. Harms are peculiarly grievous losses. Death may not close the accounts on a life completely (and it does not if postmortem preferences matter), but death arguably...
puts one beyond the reach of really serious loss. If one way of resolving the conflict between the younger and the older Boris imposes what can plausibly be regarded as a harm, and if the other way does not, should not the latter way (favoring the older Boris) be preferred?

Once we emphasize the fact that the young Boris's specific intention is to oppose the preferences of a later person, we cannot help but see that the proliferation of binding and unreleasable commitments of this sort would be undesirable. One response, short of abandoning the rigorist view, would be to weaken our view of the bindingness of commitments generally. If we wish to retain the idea that commitments are strongly binding and that commitments such as that to the young Boris are as binding as any others, then surely there is a different concession we must make. We must admit that such commitments should not be undertaken lightly, and often should not be undertaken at all.

It is tempting to say that Boris's wife is free to promise what the young Boris asks if she chooses to. After all, it is only Boris's interests that are affected. But this will not do. There may be two Borises. Indeed, the young Boris's request makes sense only because he fears that there will be another Boris, whose interests he wishes to prejudice in advance. In these circumstances, Boris's wife is not free to give her promise just because the young Boris asks for it.

The wife might perhaps give a promise limited to the 'life' of the young Boris. Such a promise could not become unreleasable. But then, such a promise could not accomplish what the young Boris wants.

The final point against the young Boris and the rigorist view is this. If the older Boris exists (as a person different from the young Boris), and if the rigorist view is rejected, then the older Boris will have certain interests of the young Boris within his control, to the young Boris's detriment. But the young Boris has it largely within his control whether the older Boris exists at all. Young Boris can murder Old Boris in his cradle if he will merely develop the strength to maintain his own convictions. That opportunity, surely, is all young Boris deserves.

In sum, where no unconnected person's interests are affected, neither the agent who changes his mind (Odysseus) nor the agent who changes his identity (Boris) ought to be bound at a later time by an earlier choice. 17

V

The arguments of this essay obviously justify no simple conclusion. I do have some final comments.

First, I would emphasize that the freedom-maximization argument of section II and the time-slice argument of section III are independent, even though they may seem most appealing as applied to the same range of cases, involving smokers, motorcyclists, and so on. Some readers may find themselves inclined to meld the two arguments into a third argument to the effect that we can interfere with the smoker, for example, in order to prevent a loss of freedom to the smoker's later, different, self. This argument differs from my time-slice argument if it
assumes that the only relevant harms to later selves are losses of freedom, and it differs from my freedom-maximization argument in assuming the time-slice view of personal identity. There is in fact some conflict between the general tendencies of the freedom-maximization argument and the time-slice argument. The time-slice argument breaks people down into successive selves and emphasizes discontinuities, whereas the very notion of freedom seems to require some degree of continuity, of integration of personality over time.

If we compare the time-slice argument of section III with the argument against the possibility of binding self-regarding commitments in section IV, there is an apparent similarity. In both cases, it seems, the later self (or else the later decision of a continuing self) prevails. But the parallel is not perfect. Under the time-slice view of section III, as I have described it, the later self does not prevail if the earlier self wishes to make a choice that, although it harms the later self, maximizes utility. The point of the time-slice argument is not that the later self always prevails. The point is that the later self counts, and deserves some protection against the earlier. In situations in which we are considering whether to enforce a commitment intended to be self-binding, the later self does prevail. The reason, roughly, is that the earlier self needs help to achieve its goals (unlike the earlier self in a motorcycle case, which needs only to be let alone), and because the earlier self has no persuasive claim to our help once its time has passed.

The argument of section IV can be viewed as one of freedom-maximization, like the argument of section II, but again there are differences. The argument of section II is primarily designed to maximize freedom as range of choice. The argument of section IV seems more likely to maximize freedom as ability to choose, by encouraging people to develop strength of will. In section IV, the range of choice the later Odysseus gains if we refuse to regard commitments as binding seems essentially the same as the range of choice the earlier Odysseus loses. It could be argued, of course, that the later Odysseus gains more than the earlier Odysseus loses. The later Odysseus gains a freedom to jump overboard, while the earlier Odysseus loses only the freedom to bind himself not to jump overboard, which may seem less worthy. Similarly, the freedom gained by the later Odysseus may count for more than the freedom lost by the earlier if, as I have suggested, there is some general reason to count the interests of those who are present (at the time when our decision must be made) more heavily than the interests of those who are past. Perhaps it is the timing of the decision about intervention that accounts more than anything else for the different feel of section II cases and section IV cases. But some difference there certainly is.

The arguments of all four of the preceding sections reveal that the sources of our difficulty about paternalism are more various than is usually recognized. The sources are at least three. One source is the conflict between freedom and other values, such as pleasure or happiness. A second source is the complexity of what we value as freedom, which includes both range of choice and ability to choose rationally and effectively. A third source is the discrepancy between the ideal and the actual. We have an ideal of a person who is both rational at each moment and
integrated over time. Actual people are neither. The question arises: Ought we to treat actual people as if they embodied the ideal, or ought we to intervene in their lives in hopes of bringing them closer to the ideal in the future?

Notes


2. This argument is well developed in Rolf E. Sartorius, "The Enforcement of Morality," *Yale Law Journal*, 81 (1972), 891-910.


6. It might be suggested that the nonenforcement of contracts for slavery involves no real denial of freedom because it involves only governmental inaction, and because there can be no objection on antipaternalistic grounds to mere failure to act. This will not do. First, since we generally enforce contracts (at least partly because freedom to contract enhances freedom in general, increasing the agent's control over her life), the refusal to enforce a particular kind of contract cannot be passed off as mere nonintervention. Second, in a case in which the buyer of the slave was able to enforce the contract without judicial assistance by physical coercion, we would not be inclined (if we opposed contracts for slavery in the first place) to recommend government nonintervention. We would hold open the nonwaivable possibility of a civil action for assault or false imprisonment on the part of the slave, and we would regard the buyer as subject to criminal prosecution regardless of the slave's preference in the matter.

7. The argument I will make in the next section suggests a response to the claim at this point in the text. The response is that we are invading one person's right in order to enhance the freedom of that person's later self, who may be a different person for the relevant moral purposes. I am content to note that in order to make this response to the argument of this section, the opponent of paternalism must open himself up to a different attack in the next.

8. I do not suggest that a perfectly discriminating ideal paternalist should ban smoking by every would-be smoker. There may be smokers who, if they could not smoke, would be so persistently wracked by nervous tension they could not otherwise assuage that they would be fit for no other activity. Such smokers should be allowed to continue. I assume such smokers are rare.


11. I shall say no more about punishment in the text. There are two points which merit some supplementary discussion.

1. Some readers may think something like the following: "It's all very well to decide not to punish an embezzler after ten years, but what about, say, Adolf Eichmann? Surely nothing could affect the appropriateness of punishing him, wherever and whenever he was found." To the extent that
punishment is justified in terms of deterrence, or in terms of appeasing public desires for vengeance, or in terms of other utilitarian goals, it may well be appropriate to punish Eichmann at any time. But on a retributivist view, whether one makes desert a justification for punishment or one that makes desert merely a necessary condition for punishment, I think it is a serious question whether Eichmann ought to be punished after twenty years. On the theory I have suggested, the answer ultimately depends, of course, on whether and how Eichmann has changed. I shall not speculate about all the possibilities. For the reader who insists that Eichmann must be punishable, I note that it may be possible to accommodate even that claim within my general approach. Eichmann’s crime was enormously evil. It may be part of our retributive theory that there is a maximum rate at which eligibility for punishment can be shed over time. If that is so, then it is plausible to believe that an embezzler could shed all eligibility over ten years, while Eichmann could not be cleansed in many lifetimes. A related suggestion would be that a crime such as Eichmann’s must have resulted from moral defects much more fundamental and therefore much harder to change than whatever was the cause of an ordinary embezzlement.

2. An untoward consequence of the theory I am suggesting is that it might place a great strain on our adherence to another principle, the prohibition against punishing innocents even when good consequences would result. My theory, if taken seriously, might require nonpunishment of many persons for crimes that earlier inhabitants of the bodies they are attached to unquestionably did commit. The theory might therefore entail that any system of judicial punishment would convict so many innocent persons that it would be insupportable. My own inclination, in the face of this difficulty, is not to abandon the suggested view of personal identity, but to reconcile myself to punishing more innocents (as defined by my theory) so long as there really are good consequences to be achieved.

12. There is actually a further complication, suppressed in the text. It might be the case that the cyclist will be injured in an accident, will not alter her attitude toward risk as a result of the accident, but will alter her attitude toward risk at some later point in her life. Does this cyclist deserve protection? It is tempting to say that in the circumstances described, the cyclist who suffers the injury is still the same person as the imprudent cyclist (even though she is going to change later) and therefore is not entitled to protection. But what about the still-later, prudent, self? If she suffers any after-effects of the injury, the imprudent cyclist will have hurt her too. For that matter, the post-accident-but-still-imprudent cyclist could be viewed as hurting her later, prudent, self by maintaining an attitude that, if the suggestion just made is correct, deprives all selves after her of protection from the first self. In the end, I think we must protect any later selves who will have learned prudence, from whatever experience or reflection. The imprudent cyclist can be interfered with unless no later self will be prudent.

13. The reader may wonder whether the Odysseus example really raises the problem of paternalism at all. At the point when the crew bind Odysseus to the mast, they are doing so with his consent. At the point when they refuse to release him, they are arguably not doing anything to him at all. The fact that the crew need only ignore Odysseus at the crucial point, instead of actively restraining him, makes it somewhat easier to accept their behavior. But I think it would make no difference to the result if Odysseus’ original command were not, “Tie me to the mast, and do not release me,” but rather, “If I start to jump overboard, grab me and stop me.” This command would require active intervention at the crucial time, but the “paternalistic” course, intervention, would be no less justified. For what it is worth, in the Homeric version of the story, the crew are instructed that if Odysseus asks to be released, they should not ignore him but should add to his bonds. He does, and they do.

14. I have here excised, as too extravagant, a long footnote in the Conference version of the paper on unamendable constitutional provisions.


16. Ibid., p. 144.

17. In the text, I argue in effect that paternalism cannot be justified simply by the prospective consent of the person to be coerced. For similar reasons, retrospective consent is neither a necessary nor a sufficient condition for the justifiability of paternalistic intervention. A good and justified effort to educate a child may simply fail, and fail so utterly that the child never sees that the parents’ attempts were appropriate. Therefore, retrospective consent is not necessary. On the other side, brainwashing
might produce retrospective consent, so retrospective consent is not sufficient. It might be objected that brainwashing would not produce a "free" retrospective consent, and that a free retrospective consent would be sufficient. This is certainly the most plausible claim that can be made for the necessity or sufficiency of retrospective consent, but it still fails. Consider an Odysseus case, in which either swimming to the Sirens or not might be reasonable. Suppose that when the ship is sailing past the Sirens, Odysseus wants to swim to the Sirens but the crew, on their own initiative, restrain him. Some days later, Odysseus announces that he has reconsidered his views about the relative importance of aesthetic pleasure and physical integrity, and he thanks the crew for having held him back. This announcement by itself does not establish that the crew were right. To be sure, retrospective consent may be relevant, for some of the same reasons as prospective consent. But it is not determinative. Nor does the time-slice view of the previous section suggest that retrospective consent settles the issue in the cases there considered. Whether an act threatens or causes unacceptable harm to a reformed later self of the agent, and whether the agent will come to be glad to have been prevented from performing the act, are related but distinct questions.