

Against Autonomy: Justifying Coercive Paternalismⁱ

Sarah Conly

Department of Philosophy at Bowdoin College

We do things that are bad for us—we take risks we soon regret, we thwart our own desires, we undercut our own fulfillment. Should we be stopped?

On the one hand, we value our liberty, and resent being told what to do. On the other, we often regret bitterly the choices that have diminished the quality of our lives. The question I address here is whether society—typically in the form of government legislation—should step in, and make people do what is good for them. I will argue that, in many more cases than we now allow, it should; that preserving our liberty of action is not worth the costs of exercising choice. I argue, then, against autonomy, against what Joel Feinberg has called “[t]he kernel of the idea of autonomy ... the right to make choices and decisions—what to put in my body, what contacts with my body to permit, where and how to move my body through public space, how to use my chattels and personal property, what personal information to disclose to others, what information to conceal, and more.”ⁱⁱ While in some cases autonomous action does no harm, in other cases it does, however “harm” is construed—as detrimental to happiness, detrimental to material survival, or even detrimental to the promotion of autonomous action.

I am arguing, then, for the permissibility of interference in personal lives, interference even in actions a person takes that will affect only himself. (Like most people, I also think we are right to interfere in actions a person takes that are unduly harmful to others; since this is relatively uncontroversial, I won’t argue for that here, although I will make use of the fact that we accept interventions for the sake of third

parties to argue that should accept them in order to prevent harms to the self.) This policy is known generally as paternalism. John Kleinig defines paternalism simply as any case where “X acts to diminish Y’s freedom, to the end that Y’s good may be secured.”ⁱⁱⁱ Under this general rubric there are paternalists of many kinds, and they vary greatly in the extent to which and ways in which they are willing to constrain people’s activities. What I will argue for is a specific and controversial position: that we may, and indeed are sometimes morally obligated to, force people to refrain from certain actions, and to engage in others. I will call this strong position Coercive Paternalism, and will show that it is indeed more acceptable, in some cases, than softer forms of paternalism that may seek to guide rather than constrain.

We are all familiar with, and generally accept, coercive intervention in some contexts. Generally, we think there are two sorts of cases where it is permitted: first, where ignorance of the facts means that the person doesn’t know what it is that he is choosing: this is why we think it is justified to grab a person who’s about to step off the curb in front of a car, or for the government to require prescriptions for some medicines. The second general condition under which we typically allow coercive paternalism is that of incompetence. There are people we think aren’t capable of dealing with the facts, even if they are informed of them, whether because of youth, or mental retardation of a sufficient degree, psychosis, emotional duress such as debilitating fear, and so forth.^{iv}

The question is why we aren’t willing to extend this acceptance of paternalistic interference into other areas. The standard response relies on the differentiation in the mental circumstances of those doing the choosing. Choosing medication requires

knowledge that most people lack, we have normally expected that average people know enough to be able to run their personal lives in the way that best suits them. Similarly for competence: most of us are thought to be capable—calm enough, rational enough, smart enough—to use facts appropriately to get us to our given ends. In most areas of life, it is argued, we should respect a person’s liberty of choice, and allow him to choose for himself.

The more we learn about psychology, though, the more we see that these differences--between areas where one should bow to expertise, and where one’s own comprehension is sufficient to assess a situation, and between the irrationality of the incompetent and the rationality of the competent—is one of degree, not of kind. In the last thirty years or so, behavioral economists and social psychologists demonstrate that failures to reason well are pervasive, as normal a part of psychology (if “normal” may be taken to be mean found in the average person) as any other. The work of Amos Tversky, Daniel Kahneman, and others, has been presented in numerous academic papers^v, and in 2008 the legal scholar Cass Sunstein and economist Richard Thaler teamed up to present many of these findings in their well-received book, *Nudge*.^{vi}

Many of the cognitive biases they report relate the effect on our decision-making of factors that even we, the decision-maker, would consider irrelevant. We are, for example, unduly influenced by framing, the particular description used in the presentation of our options ,unduly prone to think that we are ourselves are less likely than others to suffer misfortune, even of something entirely random, like lightening;^{vii} prone to miscalculate the value of a thing depending upon whether we do or don’t yet own it;^{viii} Smoking, not surprisingly, seems to involve a number of errors of judgment:

people use time-discounting to undervalue how much the future matters, or anchoring, the use of an irrelevant starting point to make comparisons, so that they judge that since the first ten cigarettes haven't hurt them the next ten years' worth won't either, or wishful-thinking, the tendency to re-interpret judgments to make what we are doing look o.k., and to conclude that since they smoke, smoking can't really be harmful.^{ix}

It is hard, in these and other contexts, for us to think effectively about what we need to do to achieve our ends.

The question is what should be done in the face of this evidence. The answer I embrace is that we need external guidance—constraints on our actions through regulation, law, and institutional design. This, though, is not popular. We fear that if we give other people control, they'll do a worse job than we ourselves, inhibiting our liberty and frustrating our desires. On the other hand, being bankrupt, addicted to cigarettes, or too poor to retire, much less retire as we like, are frustrating, liberty-inhibiting conditions, too. The question is how we can avoid these and other debilitating circumstances that liberty of action brings us to.

An alternative to coercive paternalism has been articulated by Richard Thaler and Cass Sunstein in their recent book, *Nudge*. Thaler and Sunstein endorse what they call “libertarian paternalism.” Their suggestion is that we help people do what is best for them by making the right choice easier for those with cognitive biases. If it is a question of pension plans, we make the most advantageous plan (for the employee) the default option: if you do nothing, you end up with (what is at least generally) the best choice. If you need help choosing healthy food, we put the healthy fruit at eye level, and the fried pork rinds on the bottom shelf, since people are more likely to choose whatever is at eye

level. We give you a nudge in the direction that is at best for you. However, while we change the “choice architecture,” we don’t actually eliminate options. You can still get those bar-b-q flavor pork rinds by bending over, and you can still opt for the pension plan that makes it most likely that you will end up dependent and poor. As they see it this preserves your autonomy, and thus allows us to have our cake (beneficial consequences) and eat it too (as we respect liberty of choice.)

There are two important things to notice about this, though. One, Libertarian Paternalism is manipulative. It does not suggest that we rely on rational argument in order to persuade you to change your ways. Sunstein and Thaler are not opposed to this, but they don’t think it is enough to get you to choose efficiently, because of the cognitive deficits they have described. The point of the nudge is to push you in ways that bypass your reasoning. They use your cognitive biases, like the tendency to go with the default option, to bring about good effects. There is a sense in which, then, they then fail to respect people’s decision-making ability. Second, libertarian paternalism is less likely to achieve its goal, benefit to those who choose, than is the more intrusive system of coercive paternalism. More freedom to choose means more people will choose badly. It is true that since a libertarian paternalist system allows individuals the ability to act contrary to the nudge, those for whom the default option, etc. are not good choices could bypass the nudge to hit upon a choice more appropriate to their own particular case. Libertarian Paternalism might be the ideal choice if the manipulative nudges worked for those who would otherwise make foolish choices, while the remaining option to act differently allowed only those who are choosing the most rational means to their ends to deviate from the direction into which they are nudged. Those for whom pork rinds have

no adverse health or aesthetic effects (or, whose ends really don't include or require good health or attractive appearance) would make the slight extra mental and physical effort required to get the pork rinds. Those who are prone to heart disease and yet wish to live long lives would be nudged into getting the fruit, thus achieving their goal of better health.

However, when you allow people the option to choose contrary to the direction of the nudge, this freedom isn't preserved exclusively for those who are going to use it to do what is best for them. Some people would refuse the pension plan because they have wild ideas about retirement (they want more to spend on lottery tickets) that will never yield the results they want. In other words, libertarian paternalism does is not simply preserve the option of better choices for those who, for some reason, are different from the norm. It preserves options for those who have stronger motivations than do others, or for those who have crazier convictions than the norm. It preserves the option for error. The danger is, then, that libertarian paternalism may end up neither having its cake nor eating it—it doesn't really respect choice, in the sense of thinking that people should be left to their own devices in deciding what to do. And, while it would no doubt save many people from foolish actions by nudging them in a better direction, it will leave many others to suffer the consequences of their bad thinking. We may end up with neither of the valuable things libertarian paternalism hoped to promote.

Coercive paternalism takes a different position. Rather than leaving us to sink or swim, as does the status quo, or engaging in mental manipulation, as does libertarian paternalism, the coercive paternalist will simply say some things are not allowed. I don't know that this is more respectful of people than is manipulation, but I don't see that it is

less respectful. In either case, we are trying to control people on the grounds that their own decision-making is not to be trusted. And, coercive paternalism is more likely to get us good results, because certain behaviors, like smoking, will be out of the question. So, instead of simply educating people about the dangers of smoking, as do liberals, or disincentivizing smoking by making it very expensive, the coercive paternalist gets rid of cigarettes. We know that leaving people to fend for themselves is too often simply not successful in getting people where they want to go. Instead of letting people languish in the misery caused by their own decisions, why not intervene, as we do with prescription drugs, as we do with seatbelts, and help people out?

The initial answer is that to many, using coercion to stop people from doing what they have decided, however foolishly, that they want to do, seems to give them less than the respect they deserve. What adequate respect consists in—indeed, what respect itself consists in—is a difficult question. At the least, though, to respect something seems to mean to recognize that thing's value, and to act in a way that is consistent with that value. Respect does not, however, entail letting you do whatever you want. We agree that the government can stop your neighbors from bashing you over the head when your loud music irritates them, no matter how competent as agents they are. Furthermore, we often stop others from doing something to you even if you consent to doing that (like, raising money by agreeing to let other people beat you to a pulp.) On the other hand, you are permitted to consent to be dropped into the dump tank at the fair to raise charity—what might otherwise be assault in this case *is* rendered permissible through your consent. The difference is the degree of harm you will suffer. And, we sometimes require that others

act positively for the sake of your welfare, rather than simply refraining from injuring you: we can make them pay taxes for uses you will benefit but from which they won't.

The fact that others may be required to do, or refrain from doing, something to you does not seem disrespectful. Why is it different when we require you to do, or refrain from doing, something when that is in your own interest? Some might say that in this case we prevent you from doing what you want to do, and in the other cases we only stop them from doing what you don't want, but that is simply not so. For one thing, as mentioned above, sometimes we will stop them from doing something to you that you want them to do. For another, there is a sense in which paternalism in this case *does* help you do what you want to do. Admittedly, paternalistic action prevents you from doing something you want to do at that moment, but it does this for the sake of helping you obtain something you want more, but that your short-sighted action will make more difficult to achieve. So, the difference between the permissibility of third person restrictions and first person restrictions is not whether or not the action is one you want.

Inequality:

Some will object that one difference between constraining your harmful action against another person, and constraining your harmful action against yourself, is that in paternalism there is a substitution of judgment: one party assumes that his judgment about what you need is superior to your own judgment as to what you need, to the point where, in coercive paternalism, he can force you to do what he thinks is best rather than what you think is best. This, in turn, is said to involve treating people unequally: one person's judgment about himself is held to be inferior to the other person's judgment about him.

On a political scale, accepting the legitimacy of this substitution of judgment is often taken to posit significant inequality among humans. The question frequently asked about paternalism is who will determine what rules there should be, and the fear is that there will be one class of persons who make the rules, and another class who obey them. In fact, however, it is not an assumption of the superiority of one group to another, but of shared fallibility, that moves us to paternalism. There is no evidence of demarcation in education or IQ that distinguishes who is, and who is not, prone to the sorts of errors which can prevent us from reaching our goals.^x The sorts of errors the paternalistic intervention promoted here addresses are a function of circumstance rather than the kind of person doing the thinking. What we need, then, is for these sorts of decisions to be made under a different set of circumstances, not by a distinct class of people. Rules do not need to be made by a superior bunch of thinkers, but by any and all of us when we are doing our superior thinking.

Degradation:

Even if we assume that we are equal in this regard, though, this will not satisfy everyone who thinks that paternalism is essentially disrespectful. Some people will argue that it is degrading to say of people that they are not fit to make all the decisions required for running their own lives, and if this is true of all people, rather than just some people, so much the worse.^{xi} The question, though, is what degradation consists in. Someone may degrade me if he refuses to acknowledge my abilities, perhaps because he assumes that someone of my race or sex simply cannot have them. He degrades me, because he refuses to acknowledge that someone with these characteristics can be as smart, as able,

as valuable, as I am. He puts me down, trying to assign me to a category to which I don't belong, and inferior to his own.

When someone accurately assesses my abilities, though, and finds me lacking in some respects, it is very hard for me to argue that I have been degraded or disrespected. We regard it as insulting to treat humans like animals, but we don't regard it as insulting to treat animals like animals. We regard it as insulting to act as if women aren't capable of balancing their checkbooks, because they can. We don't regard it as insulting to assume that the man on the street can't do quantum mechanics, because he can't. The coercive paternalist believes the facts that suggest a change in the status we accord people, a change from what we might have thought about ourselves to a more realistic acceptance of our inabilities. Indeed, it is more demeaning to pretend to have a quality that you don't than to admit to not having one you might like to have. It suggests that what you've got left when you eliminate the disputed property isn't worth much—but there is no reason to think of this of people. We remain as we have been, as we have experienced ourselves, and have appreciated ourselves, and this is clearly valuable, whatever stature we may lose.

Coercive paternalism deprives us of some of our liberty, taking decisions out of our hands and entrusting them to others. However, it helps us achieve long-term goals without manifesting disrespect, positing inequality, or degrading our dignity. Interference in individual choice is thus justified in far many more cases than we now admit.

ⁱ This paper is in part a synopsis of my forthcoming book, *Against Autonomy: Justifying Coercive Paternalism*, Cambridge University Press, publication expected in the autumn of 2012..

ⁱⁱ Jules Feinberg, *Harm to Self: The Moral Limits of the Criminal Law*, Oxford University Press, 1986, p. 54.

ⁱⁱⁱ Kleinig, John, *Paternalism*, Rowman and Allenheld, 1983, p. 18. A recent definition by Peter de Marneffe includes another condition, that the person who is affected by the policy would prefer that his choices not be so limited: "...a government policy is paternalistic toward A if and only if (a) it limits A's choices by deterring A from choosing to perform an action or by making it more difficult for A to perform it; (b) A prefers A's own situation when A's choices are not limited in this way; (c) the government has this policy only because those in the relevant political process believe or once believed that this policy will benefit A in some way; and (d) this policy cannot be fully justified without counting its benefits to A in its favor." He also cites Richard Arneson and Gerald Dworkin as having believed a policy is only paternalistic if it is unwanted by the person to whom it applies. I am not unfriendly to this condition, properly understood. It doesn't seem paternalistic if a policy makes us do what we already want to do, although it might make such a policy redundant. As seen below, though, there is some ambiguity involved in determining what we want, since we may want conflicting things, including wanting both to achieve an end and to take a means that does not achieve that end. "Avoiding Paternalism," *Philosophy & Public Affairs*, Winter 2006, vol. 34, # 1, 68-94

^{iv} I will note here that more moderate anti-paternalists also accept paternalism in some cases where the chooser may be neither ignorant nor incompetent: those where the dangers to the agent are very high, and very proximate, and very probable. Thus, some who generally oppose paternalism will allow it in cases, like motorcycle helmet laws. (See Kleinig, *Paternalism*, ch. IV and again pp. 109 ff, and Feinberg, op. cit., chs. 20 and 21. They have some difficulty in arguing that paternalism is permissible in these cases but not more generally. I will suggest that the reason moderate anti-paternalists allow paternalism when the costs of liberty of action are high and definite is really a function of a cost/benefit analysis. There is some cost to controlling people's actions, so when it doesn't matter, we don't think we should do it. But when it does matter enough, we think we should interfere. We are now seeing that there are more cases than we had thought where liberty of action results in sufficient disutility to justify paternalism.

^v See, for example, Daniel Kahneman and Amos Tversky's "Prospect Theory: An Analysis of Decision under Risks," "Advances in Prospect Theory: Cumulative Representation of Uncertainty," "Loss Aversion in Riskless Choice: A Reference-Dependent Model," "Rational Choice and the Framing of Decisions," all in *Choices, Values, and Frames*, Kahneman and Tversky, eds., Cambridge University Press, Cambridge, 2000.

^{vi} Richard H. Thaler and Cass W. Sunstein, *Nudge: Improving Decisions about Health, Wealth, and Happiness*, Yale University Press, New Haven and London, 2008.

^{vii} David M. DeJoy, "The Optimism Bias and Traffic Accident Risk Perception," *Accident Analysis and Prevention*, vol. 21, issue 4, 1989, 333-340; Neil D. Weinstein, "Optimistic Biases About Personal Risks," *Science*, vol. 246, #4935, Dec. 8, 1989, 1232-1233; Dale

Griffin and Amos Tversky, "The Weighing of Evidence and the Determinants of Confidence," in *Heuristics and Biases: The Psychology of Intuitive Judgment*, edited by Thomas Gilovich, Dale Griffin, and Daniel Kahneman, Cambridge University Press, New York, 2002.

^{viii} Daniel Kahneman, J.L. Knetsch, and Richard Thaler, "Experimental Tests of the Endowment Effect and the Coase Theorem," *Journal of Political Economy* 98, 1325-1348, 1990

^{ix} See Robert Goodin, "The Ethics of Smoking," *Ethics*, vol. 99, #3, 1989, pp. 574-624, for a detailed account of the cognitive biases typically involved in smoking.

^x It is true that in the U.S., less educated people smoke more than do more educated people, and some might take this to mean that less educated people are, as a class, more prone to make irrational decisions. (See the Center for Disease Control, www.cdc.gov/tobacco/data_statistics) However, the same sources tell us that men also smoke more than do women, and that there are significantly different rates of smoking among different races, and no one suggests that these differences auger different degrees of rationality. We see no evidence that one group is generally more able to avoid cognitive bias than another.

^{xi} For example, Gerald Dworkin says, "What does have intrinsic value is not having choices but being recognized as the kind of creature who is capable of making choices. That capacity grounds our idea of what it is to be a person and a moral agent equally worthy of respect by all." *The Theory and Practice of Autonomy*, Cambridge University Press, Cambridge , 1997, p. 80