Moral Paternalism

Sarah Conly

Associate Professor, Department of Philosophy, Bowdoin College, Brunswick, ME, USA; sconly@bowdoin.edu

ABSTRACT

I present a brief overview of the paternalistic policies that are currently promoted, and then suggest that one broad area has been overlooked: individuals’ interest in behaving in a way that reflects their moral values. Paternalists want to promote individual welfare, generally construed subjectively: welfare is a function of that person’s own goals. Most paternalists have focused on relatively material goals, including good health and financial success. If we examine the idea of subjective welfare more closely we will see that we have other goals as well: most people have moral values, and moral values that are not purely egoistic. People make mistakes in the pursuit of their moral goals that are in some cases very similar to the mistakes they make in the pursuit of material welfare, and similar interventions are permissible.

Keywords: Moral values, paternalism, welfare, moral agency, charity, coercion

1 Welfare

There has been a lot of discussion of paternalism in public policy of late. Rather than trying to recapitulate that in all its complexity, I will present a brief overview of the current discourse, and then suggest that one broad area of paternalist policy has been overlooked: individuals’ interest in behaving in a way that reflects their moral values. I will argue for its inclusion in the things that paternalist policies should promote, and address objections to that.

Contemporary writing on paternalism in public policy, both for and against, has centered on the promotion of personal welfare. Proponents argue that individuals who are self-interested nonetheless make decisions that are at odds with their self-interest, and that institutions such as government can and should intervene to promote choices that in fact will be more fruitful from the
individual’s point of view (Thaler and Sunstein, 2008; Conly, 2013; Le Grand and New, 2015). Opponents, too, focus on personal welfare, but argue variously that people’s choices already reflect what they really want, and thus conduce to their welfare, or alternatively, that even if such choices do not effectively achieve personal goals, policies instituted by third parties will do even worse (Mill, 2003, pp. 121–138). Both sides embrace, thus, what is called a subjective view of welfare: what constitutes an individual’s welfare is a function of what that individual herself wants. As it was described originally by L. W. Sumner, “A subjective theory will map the polarity of welfare onto the polarity of attitudes, so that being well-off will depend (in some way or other) on having a favorable attitude towards one’s life…” (Sumner, 1995, p. 76). Opponents of paternalism, on the other hand, tend to argue that whatever their intentions may be, paternalists are liable to force people into doing things that don’t actually benefit the individual: they will promote health, for example, when individuals don’t value that as much as liberty, and thus impose foreign values on individuals in a way that is both disrespectful and contrary to personal happiness. Both sides take the individual’s desire for or endorsement of a state of affairs as necessary to its achievement contributing to his being better off.

Both, then, typically aim to avoid what are called objective theories of welfare. Those who hold an objective view of welfare believe that what is better for people is not a function of what they themselves want but rather of some objective system of values. The fact that a person consistently wants to live a certain way does not mean that it is better for him to do so: on this account, he may simply have bad values, and he is worse off if he lives in accordance with those. Some such views are quite unattractive to the modern eye, however wholeheartedly they were embraced in the past: those who have held objective theories might hold, for example, that it is better for people to be Christian, or heterosexual, or monogamous, regardless of whether the people in question actually want that, and historically, well-meaning paternalists who tried to impose these supposed objective value have created suffering and sometimes severe psychological harm. Contemporary objective accounts of well-being are often more attractive, promoting less restrictive goals, like education, social relationships, and the acquisition of knowledge as a goods (Hurka, 1993). I will not debate the two approaches to welfare here, although I favor subjective accounts. I want only to contrast the two and to draw attention to the general orientation towards subjective theories in contemporary discussions of paternalism, because if we endorse a subjective theory of welfare, it naturally makes sense to look at what people actually desire and value.

Paternalists have certainly done a great deal of this. There has been a good deal of exploration of how to promote the achievement of things we may naturally take to be subjective goods, like health, both in scholarly articles and in press designed to affect policy more directly. Good health is generally desired both for its own sake and because it typically enables us to pursue our
other goals, whatever they may be, more effectively than if we are unhealthy, and so its achievement has been widely examined (Thaler and Sunstein, 2008; Khullar, 2017). There has been a great deal of discussion of money, in pension plans, credit card debt, or other areas, again because running up unpayable debts or facing old age with insufficient funds strikes most of us as something we would very much like to avoid (Benartzi and Thaler, 2007; Hollanders, 2015). This is all worthwhile and the discovery of more and better ways to promote health and economic well-being is to be encouraged.

2 Moral Values

However, if we examine the idea of subjective welfare more closely we will see that there are other permissible and indeed welcome interventions. While it is reasonable to pay attention to health and finance, we should be aware that we have other goals as well, because the achievements of goals that pertain to our own material well-being are only part of what we care about. Most people have moral values, and moral values that are not purely egoistic. Most people care, for example, about the welfare of others. Most people extend this concern for others to those who will live in the future, and thus hope that we do not, for example, bring about environmental devastation. These may not be their priorities, but they are nonetheless values that are important to them. Just as people fail to act efficiently and effectively on their material goals, however, many fail to act effectively in the service of their moral goals. In some cases, of course, this is because their interest in their own material welfare is greater than their interest in others, and insofar as that is true, intervention on paternalist grounds would be unjustified. However, many times such failures don’t seem to result from a conflict between a greater interest and a lesser moral interest, but from the same sort of defects in rationality that we see in other areas. Just as we order another fat-and-salt burger despite our knowledge that our life would be more enhanced by a salad, we do things that clearly undercut the achievement of our moral goals. Thus, if paternalist interference is justified in the achievement of material welfare, it seems as if it should be equally justified when it comes to helping people avoid moral failures—if that is in fact something that we can efficiently achieve.

This is an area paternalists have avoided. One fear may be that talk of promoting morally right action for paternalistic reasons may hearken back to the oppressive policies of those who held objective theories of welfare, discussed above. Modern paternalists want to avoid this, and for that reason, perhaps, hesitate to talk about individuals’ moral beliefs. Imposing foreign standards of personal life upon unwilling individuals is not what is being discussed here, however. Instead, the idea is that a paternalist who holds a subjective view of welfare, where the only thing that makes people better off is living according
to their own goals and values, may endorse intervention that allows people to
do that more effectively even when the values involved are moral values and
not just material values.

We know, after all, that reason often fails when we try to implement our own
moral values, just as it fails when we try to do what is prudentially best for us.

We know that people are inordinately moved by one vivid example—a
picture of a starving child—and irrationally unmoved by statistical accounts of
severe hunger, even when they accept those accounts as correct. Disasters that
are videotaped (the Indonesian tsunami) get a more compassionate response
than disasters we read about. Emphasizing that a very large number of people
are affected by a problem—a natural disaster, a disease—actually makes people
less willing to contribute to its solution (Tsipursky and Slattery, 2016). And,
of course, we respond more emotionally to disasters that occur to those who
look like us and speak our language than to others (Chiao and Mathur, 2010).
Equally irrationally, people accept the status quo in activities as sufficient for
meeting their moral goals, merely because that is the status quo, and do not
consider that in many cases the status quo—the activities they have always
participated in, or social policies they regard as normal—are antithetical to
their moral values, while not being necessary to the achievement of their
material aims. Just as we allow smoking because we grew up with it, even
though we know that today the FDA would never permit the introduction of
cigarettes, we accept driving gas vehicles with poor mileage even as we deplore
global warming. Hybrid cars function just as well, but we stick overwhelmingly
to gas-driven cars, despite the fact that this doesn’t promote either our material
or moral interests (Consumer Reports, 2015). Just as people choose poorly
when taking actions that will affect their own future material welfare, they
choose contrary to their own values because they do not, in the sense that
has been widely discovered and discussed by Kahneman, Tversky, Thaler, and
others, see that their actions are so contrary. Thus, in many moral cases we
see faults in thinking leading to outcomes antithetical to the agent’s own goals.
It should, then, be equally justifiable to impose paternalistic policies geared
to the achievement of moral goals. All such interventions are conditional on
an acceptable cost-benefit analysis, but if such an analysis shows reason for
interference, that may permissibly be done.

3 Dissimilarities

It may be argued, however, that the two areas of interest—moral and material—
are not sufficiently similar for an intervention permissible in one to be permis-
sible in another.

1. First, a moral philosopher might well argue that a morally right action
that is prompted through external intervention, when that is anything more
than a helpful suggestion, is no longer a morally right action. Even if someone does what would normally be considered a morally right action—helping the needy—that person himself hasn’t acted rightly if he wasn’t motivated appropriately. That his action is coextensive with what duty requires is a coincidence if he was motivated by something other than a regard for doing what is morally right. If, by turning on a light switch, I, unknown to me, disrupt the electric grid such a terrorist bomb fails to detonate, it’s a good thing that happened, but no one judging my moral agency would say I acted in a morally creditable way. I was acting neither morally rightly nor morally wrongly—morality was not involved in my action.

There are roughly three sorts of intervention that may be considered in light of this. Some (not all) paternalists argue that it can be permissible to coerce people into doing what is good for them, when that is the most efficient way to bring about the desired end (Conly, 2013). When we look at material welfare, for example, banning cigarettes altogether would be a coercive strategy for preventing people from smoking. Second, one may change the costs and benefits associated with outcomes in a way that affects the individual’s choice. Choosing the worse outcome is still an option, but given that the costs have changed the person is less likely to choose it. When we use this paternalistic method in the material realm we might, for example, raise taxes on cigarettes, or give breaks on insurance costs to those who don’t smoke, all to make it more likely that that a person will choose not to smoke.1 Third, we may use methods that might be described as merely psychological—not changing the costs and benefits associated with particular choices but merely making some choices look more, and others less, attractive. We might re-frame options to make the personally beneficial option more appealing. Or, we might propagate simple suggestions or recommendations, designed to put the desired alternative vividly in the individual’s mind, and to suggest, perhaps, that this is the choice endorsed by our society. Warnings on cigarette packages would fall into this category, or graphic representations of good nutrition, like the old food pyramid or the contemporary “MyPlate” design.

When we are promoting material interests we may not care exactly what the motivational structure of the moral agent is as long as the method used gets her to do what is in accordance with her long-term values. In morality, though, it is different: motivation matters to the evaluation of an action. The argument is that that when it comes to moral action, if the impetus to the action comes from outside the agent, the action has no moral value. Thus, you can’t “help” someone to perform a morally right action. Morally right actions—and any actions that are subject to moral evaluation—must be prompted by the will of the agent.

1 If these costs are so made so high that no reasonable person would pay them, this would become a coercive method. The line between coercing and incentivizing is a function of the amount of pressure placed on the individual to choose in a certain way.
There are several answers to this. First, if we adopt the third approach, using soft, merely psychological methods, we may altogether avoid charges that moral agency has been unacceptably undercut or eliminated. Much as some believe in the power of the rational will and its necessary role in any action properly subject to moral analysis, no one seems to object to education: to teaching facts that are relevant to our moral decisions. Public service reminders as to how much driving affects air quality, for example, do not prevent choice. That there should be greater knowledge of the effects of our actions on others cannot be controversial. Moral education is often regarded as an essential means to becoming a mature moral agent, so if this education occurs through public suggestions and warnings rather than merely from parents or schools, no harm is done to agency.

Incentives and coercive measures though, may seem to be very different: here we do not rely merely on the (properly informed) will but change the costs of the action in question. Can a person who acts conscientiously only in these circumstances be said to be acting in the way that would justify praise for right action?

It might seem not. We think that the person who gives to charity only to place himself in a lower tax bracket isn’t due moral praise, because he isn’t acting out of a regard for what he believes is morally right. Even if he does what might be considered by some to be an objectively morally right action, like helping the needy, we do not evaluate his character as morally good because he is motivated only by self-interest.

However, that is not what is going on in the case of someone who does what he thinks is right only when paternalist strategies help him to do that. Positive moral evaluation can be appropriate when incentives were necessary to the action in question being taken, or even when the action was coerced.

Consider how we evaluate moral character, and in particular, consider the way we evaluate failures to do what the person in question believes is morally right. We consider such people morally flawed for either of two reasons. First, we may consider that the individual failed to care enough about doing the right thing. Whatever lip service he may have given to the duty to help the needy, he didn’t rank that duty high enough to outweigh his selfish desires to spend money on his own comfort. He’s just a person who doesn’t care much about doing what he considers right, and so we judge him as a flawed moral agent.

Second, we often claim that some people who do care about doing the right thing, all things considered, fail to act rightly on a given occasion because they suffer from weakness of will. They are tempted to do what is wrong, to steal when the money is right in front of them even having lived a life of probity. Instead of fighting that temptation, as a good person would, they give in, doing what they believe to be wrong. (Consider Hurstwood, in *Sister Carrie*, who after a life of honesty is tempted by the sight of ready cash into
stealing it, even though he regrets it as soon as he has done it) (Dreiser, 1900). We may well be more sympathetic to the person who acts weakly, who runs from battle when confronted with danger, than we are to the person who just doesn’t care much about probity or duty to his country, but we still commonly regard weakness as a moral failure and the weak action as culpable.

When we look at what goes on in the person who is helped to do what he himself thinks is right by paternalist strategies, however, neither of these failures is present. First, we are assuming here that the person aided by a paternalist stratagem does hold the value in question, and holds it as enough of a priority that it outweighs the alternative options he might pursue. That is why such an intervention would be paternalistic, assuming the modern take on of paternalism where it is geared to help people achieve their own values. As discussed above, today’s paternalists accept a subjective view of welfare, where a person is made better off by a policy only if it accords with his own values and the relative strengths those values have for him. We can assume, then, that if we are helping someone to pursue a value it is because he actually holds that value—not like the person who gives to charity only to get himself in to a lower tax bracket, but rather as someone who genuinely wants to help those who are needy. If the paternalist helps him pursue that value at the expense of material interests, that is only because the individual himself holds material interests to be secondary to this value. The paternalist aims to reflect what the agent values and the degree to which he values those things. We aren’t pushing him to do things he really doesn’t want to do.

What about weakness of will? Does the paternalist supply artificial backbone to someone who left to his own devices would lack the strength of character, self-discipline, or whatever else it takes to stick to one’s principles even in the face of temptation?

Maybe, but this calls for a (brief, under the circumstances) reflection on the nature of weakness of the will. At least since Plato’s Republic philosophers have tried to make sense of how it is that someone fails to do what he thinks he should, and many have adopted the language of weakness, akrasia, and incontinence, suggesting a lack of strength or control underlies such choices. What the massive literature on cognitive bias suggests, however, that in many cases to call a mistaken choice “weak” is to misdiagnose the mechanism in play. Often actions that are poor means to an agent’s ends aren’t so much a function of poor “will-power” whatever that may be, as failure of rational thought. Much of the discussion of bounded rationality, cognitive bias, and other failures of reason may be used to explain what previously has been chalked up to weakness. Robert Goodin’s classic discussion of smoking, for example, shows that smokers don’t simply give in to temptation, but rather are affected in their choices by cognitive biases (Goodin, 1989). Our “wills” become stronger when we correctly cognize and rationally respond to the role an action will play in bringing us to an undesired end.
Just as we as individuals may need help to avoid poor instrumental reasoning when it comes to health care, so may we need help when it comes to implementing our moral values. Our agency is not undercut but is rather displayed if we willingly embrace mechanisms that make it more likely that we will do what we think we ought to. The fact that we need help doesn’t mean we don’t really value the end in question; it just means that we sometimes make irrational choices in pursuing our goals, which at this point in time should come as no surprise to anyone.

Given this, incentivizing actions that make us able to act in accordance with our overall values don’t suggest that we are inferior moral agents, or that we shouldn’t be held as praiseworthy for doing what is necessary to achieve our moral goals. Signing up for a monthly donation to United Way incentivizes charitable giving, because it makes that easier. We don’t have to think about it, and we don’t have to motivate ourselves to do the right thing over and over again. It is, then, more effective in helping us reach our goals than would be leaving the choice up to us each month. Yet, it seems to be morally right to sign up for automatic donations to United Way, NPR, the World Wildlife Fund, or any such, even though it reduces the opportunities for repeated evaluation and choice. Assuring that we will do what we believe is right is a way of acting morally rightly, not something that diminishes the value of our actions. Making it easier to recycle seems morally unproblematic, even though it then takes less effort to get oneself to do the right thing: who would think it would be a good idea to make all such actions more difficult?

Can the same be said of coercive policies that do, by their nature, eliminate choice? If you are forced to take what you believe to be a morally right action, can it still be an action that exhibits praiseworthy moral character? It depends. At least on some conceptions of morality, this won’t be very different from a system of (mere) incentives. If the agent endorses the coercive regulation because she thinks it is the best way to achieve a valued goal, then again, that doesn’t seem to make her less of a moral agent. If, on the other hand, anyone is willing to say “let children starve if I (or someone like me) cannot save them through the pure force of individual will,” then that person will naturally deprecate good outcomes that are not reached through voluntary means. To me, at least, this seems to be a very odd stance to take. Indeed, the belief that the outcome is less important than the agency of the person trying to bring that outcome about seems contrary to good moral character rather than indicative of it. For most people, the point of right action is not to burnish their own claims to virtue but to see that the right thing is done.

2. Second, in terms of public policy, it may well be thought that there is much less uniformity about the moral goals of a populace than about the material goals. We are all biologically the same, and what conduces to discomfort in one generally conduces to it in another. No doubt we vary in how much we value health relative to other things, but there is at least a
base line agreement as to what good health consists in, and almost complete uniformity in believing that good physical health is at least one significant constituent of welfare. We may have no reason to expect the same agreement when it comes to moral goals, where there is no obvious empirical standard for what is good. If we force a particular moral action on people who don’t share that conception, it seems to bring us back to the dark side of paternalism, the rejected paternalism that tried to foist “objective” standards of welfare on those to whom such standards were merely oppressive.

Fortunately, there is a practical difference between measures intended to stop a particular activity and those intended to promote achievement of the agents’ moral goals. If we place a prohibitive tax on cigarettes, we need to do that uniformly across a country. A high tax in Maine without one in New Hampshire will simply cause many Mainers to drive to New Hampshire to buy their cigarettes, as they often do now for alcohol. Insofar as we want to initiate positive efforts to fulfill moral goals, though, we can allow people a variety of choices as to action. We could require that a certain percentage of income (above a threshold) be spent on charitable endeavors, but could allow people to choose where their money is spent, as some umbrella charitable organizations now do. They could focus exclusively on the needs of the American poor, or even more locally, if they think “charity begins at home.” Or, they could give to the neediest across the world, or to environmental causes, or to Médecins sans frontières, or to Planned Parenthood. In the end they would simply need to show that they had contributed something.

Granted, charitable giving won’t meet all moral goals, any more than prohibiting cigarettes and trans fats will meet all health goals. An individual might have a moral value that is idiosyncratic and personal, and that just isn’t the sort of thing that can be approached this way. Ideally, all policies would be crafted to fit each individual’s needs and goals, but we just don’t have the ability to do that. This is a problem for all legislation, though, and not a devastating problem for charitable aims in particular.

3. What if the person has terrible values? Is the paternalist condemned to promote them?

No. Paternalists are not merely paternalists. Paternalists generally want to make people better off—that is the point of paternalist policies. Whether it is a question of material interests or moral interests, the paternalist will consider the costs to the rest of society in promoting, or failing to promote, individuals’ achievement of their goals. If a person has an idea of material welfare that consists of his having more money than everyone else, no matter what it takes, a paternalist may obviously consider that that is a goal it is too costly to society to implement. Similarly, if an individual Nazi had as a moral goal—that is, a goal he himself considers morally valuable—such as eliminating Jewry from the face of the earth, no paternalist would feel bound to promote it. On the contrary, the paternalist would consider the goals of others (survival;
happy lives; equal treatment) and take whatever steps necessary to prevent the racist from achieving his goals.

4. Is a paternalist justification for promoting conscientious action necessary? It may seem that a paternalist justification for pushing people to act on their moral values is unnecessary. After all, we’ve always allowed interference for the sake of third parties. Mill was the great opponent of paternalism, but in *On Liberty* his Harm Principle claimed that interference for the sake of third parties was the only interference that could be justified—it was interfering with an individual for that individual’s own sake that, on his account, was impermissible (Mill, 2003). Interference for the sake of others’ interests is fine. Given this, it may be argued that looking for a paternalist justification is redundant, at least in terms of practical action: it will make no difference to when interference is permissible.

In fact, however, a paternalist justification for promoting morally good action will change the way we approach support for morally good action in at least two ways.

First, it provides justification for interference. It is true that we already feel no compunction about stopping you from acting contrary to your own values if that means preventing you from hitting someone in a burst of temper. That, however, is a fairly limited range of permissible interference. While Mill himself said that we might interfere to prevent harming the interests of others, he said that that was true where those interests might be considered as rights (Mill, 2003). This meant that we might be prevented from actions that would violate rights (blows to the head) and in some cases obligated to act in a way that other citizens may demand of us (such as serving on a jury.) His permissible realm of interference clearly does not include all those things which we believe we ought morally to do. (Mill, of course, is a utilitarian. Whether these constraints on permissible government interference are consistent with the utilitarian criterion of right action is a large question, not to be debated here).

These limitations are not surprising, because this way of justifying interference frames the justification in terms of a conflict of interests. I want x, you want y, and can’t both have what we want. In such a case, we weigh the interests or rights at hand and determine whose has priority. The loser gives up something: his interest in x (hitting you in the head, avoiding jury duty) is not met. If, when we require something of the loser, we are acting against his interests, we need a great deal of justification, and are likely to limit our “takings” as much as possible. A more paternalistic approach, however, would in many cases see that there is no real conflict between two parties, and thus be readier to intervene. If you need charity, and giving to charity is actually consistent with my goals, then in getting me to give to charity the government creates no deficits: both parties concerned are achieving their goals. Thus, interference is much more easily justified.
Second, if we allow government action to promote actions that accord with what people see as their duty, we may eliminate one of the primary reasons people fail to do individually what they think generally should be done. A person may care a great deal about the environment, for example, but recognize quite rationally that in some cases what she herself does makes very little difference. In such a case it becomes virtually impossible for her to act effectively in service of her goals, which naturally may diminish her interest in acting in support of her goal, especially if that entails some sacrifice to her material interests. That depends both on the goal and on the person, of course: a person may value natural areas that are free of trash and contribute materially to that by collecting trash or getting her municipality to institute trash collection. Some broader areas, though, remain out of range of any given individual’s efforts, like preventing global warming, and whether or not a person will amend her actions in light of her overall values depends. Some people will nonetheless act in the way that they know would be effective if everyone did it even when they know that everyone won’t do it, but other people won’t feel motivated to do what they know effects no good.

If a government regulation is enacted to further such broad goals, however, the game is changed: it is the nature of effective regulations that they insure that most people act in accordance with them, so a government regulation geared to minimize global warming will release the individual from the fear that her act is an isolated one that has no effect. If the regulation is one that, if conformed to, will have an effect on climate change, then the individual sees a point to her action. Such regulations then, again increase the individual’s incentive to act in the way that promotes a person’s values, and the fact that such actions are now effective provides justification for the government requiring it in any given case.

Given this, considering a paternalist rationale for promoting other-regarding moral values means that we can extend the sorts of actions we may expect of people without infringing on their liberty. We would be helping them to do what they really want.

4 Conclusion

While poor moral choices are not always made for the same reasons as poor prudential choices, paternalist intervention can often be justified in moral choices just as it can be justified in prudential choices. Some may see this is akin to a reductio: they will argue that paternalism must be wrong if accepting paternalist intervention when it comes to material interests also entails the permissibility of such interventions when it comes to moral values. However, the opposite is true. While there are those who doubt that people really fail to do what they want when it comes to their material interests, many more
are willing to accept that we sometimes fail to act in accordance with our moral values. The most reasonable conclusion is that we are flawed agents when it comes to advancing both our material and our moral interests, and that we will be more truly ourselves if we accept the help we need in living according to our values. While many fear even the most anodyne paternalistic interference on the grounds that it might justify more interference, I think such extensions arise simply from the recognition of what we are, and that is nothing to be deplored.

References


