

Nudge and Shove

Author(s): Richard J. Arneson

Source: *Social Theory and Practice*, Vol. 41, No. 4, Special Issue: Preference, Choice, and (Libertarian) Paternalism (October 2015), pp. 668-691

Published by: Florida State University Department of Philosophy

Stable URL: <http://www.jstor.org/stable/24575754>

Accessed: 03-04-2018 22:59 UTC

---

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://about.jstor.org/terms>



JSTOR

*Florida State University Department of Philosophy* is collaborating with JSTOR to digitize, preserve and extend access to *Social Theory and Practice*

## Nudge and Shove

Richard J. Arneson

**Abstract:** This essay reexamines the idea of paternalism and the basis for finding it objectionable in light of recent writings on “libertarian paternalism.” Suggestion: to qualify as paternalistic, an interference that restricts someone’s liberty or interferes with her choice-making with the aim of helping the individual must be contrary to that very individual’s will. A framework for determining the justifiability of paternalistic action is proposed, under the assumption that the individual has a personal prerogative, up to a point, to engage in less than maximally beneficial action. Beyond that point, the content of the will of the individual disposed against interference can extinguish the presumptive wrongness of paternalism.

**Keywords:** paternalism; libertarian paternalism; nudge; the personal prerogative; content of will

Innovative work by Cass Sunstein and Richard Thaler has reinvigorated discussion of paternalism.<sup>1</sup> They propose that without restricting people’s liberty for their own good, we can induce people to choose courses of action that will lead them toward outcomes that are better for them, by their own lights, than the actions they would have chosen absent our intervention. We can do this by exploiting widespread defects in human decision-making, by setting up choices for some other people in ways that will trigger decisions by them that are not well considered, but that will lead to an outcome that is good for them.<sup>2</sup> This is what these authors call “nudging.” Nudging interventions do not restrict the individual’s

---

<sup>1</sup>Richard H. Thaler and Cass R. Sunstein, *Nudge: Improving Decisions about Health, Wealth, and Happiness*, revised and expanded ed. (London: Penguin, 2009); Cass R. Sunstein, *Why Nudge? The Politics of Libertarian Paternalism* (New Haven: Yale University Press, 2014).

<sup>2</sup>On common and pervasive defects in human decision-making, see, for example: Daniel Kahneman and Paul Slovic (eds.), *Judgment under Uncertainty: Heuristics and Biases* (Cambridge: Cambridge University Press, 1982); Daniel Kahneman and Amos Tversky, *Choices, Values, and Frames* (Cambridge: Cambridge University Press, 2000); and Daniel Kahneman, *Thinking—Fast and Slow* (New York: Farrar, Straus, and Giroux, 2011). The defects in human decision-making are to some extent not simply defects, but shortcuts in thinking that tend to yield pretty good results in certain familiar contexts. Nor is it the case that when one’s fully deliberative thinking process is triggered, one’s thinking always becomes more rational, much less fully rational.

choice set, except perhaps in trivial ways, such as by making it a tiny bit more difficult to attend to the presence of a gooey desert option and reach for it and put it on one's plate. Sunstein and Thaler attach the label "libertarian paternalism" to their program. A social planner following the libertarian paternalist agenda seeks to identify and implement public policies that will manipulate people's tendencies to irrational decision-making to their own advantage, by setting up situations in which the irrational decision-making tendencies will predictably prompt an individual to choose options that bring about greater satisfaction of her own preferences, weighted by their importance as the individual herself subjectively ranks them.

Responding to the Sunstein and Thaler way of thinking, we confront the old questions, what is paternalism, what if anything is wrong with paternalism, and how do we draw the line between paternalism that is morally acceptable all things considered and unacceptable paternalism. The issue is not merely what is the correct usage of a certain familiar label. We seek to clarify the relevant concepts involved in disputes about paternalism and to discover, in this region of action and policy, what we ought to find acceptable and unacceptable *pro tanto* and all things considered, whether or not you choose to call "paternalistic" what we end up classifying as unacceptable.

Sections 1 and 2 of this essay consider how it would be useful to conceive of paternalism. The upshot is that paternalism is restriction of an individual's liberty or manipulation of his choice-making, against that individual's will, motivated by the aim of benefiting the individual. Nudges of individual choice-making as characterized by Sunstein and Thaler are manipulative, so when they are against the will of the targeted individuals and intended to benefit them, they qualify as paternalistic.

Section 3 highlights tensions among the normative strands woven together in Sunstein and Thaler libertarian paternalism. Section 4 contrasts two polar extreme positions regarding the moral wrongness of paternalism as here characterized, welfarist consequentialism and semi-libertarianism. The essay then suggests a framework in which welfarist consequentialism could be modified so that a range of moderate positions regarding the wrongness of paternalism becomes available. The individual in this framework will have a moral right against paternalism that is limited by the limits of a personal prerogative and by a duty to self. If an individual's self-affecting action in fact is beyond the protection of the personal prerogative, then either the agent's underlying will is not against interference even if his occurrent will is, or the agent's under-

lying will does oppose interference, but in that case his will is defective.<sup>3</sup> Either way, the fact that the paternalistic interference is against the will of the person who suffers the interference ceases to be a reason against interference.

Section 5 considers an objection to the effect that on the view proposed, the agent's actual will concerning interference drops out of the picture as irrelevant to the permissibility of paternalism. Section 6 considers an objection to the effect that the view proposed has unacceptable implications regarding issues beyond paternalism, so is unlikely to give acceptable guidance regarding paternalism. Section 7 is a concluding summary.

### 1. Is So-Called "Libertarian Paternalism" Aptly Named?

Suppose a person's behavior over time seems to reveal preference instability. On Tuesday, I resolve to exercise more, restrict my calorie intake, and lose weight. I buy an exercise bike and salad makings. During the rest of the week, I watch TV a lot, eat calorie-laden fast food, exercise not at all. The lettuce rots in my refrigerator. Next week a similar cycle repeats itself. I throw the TV in the trash. Interpreting my behavior as inconsistent is tricky. Perhaps my behavior reveals a steady preference for pretending to go on a diet and lose weight. Perhaps my preferences shift over time, and in this case, why accord a privileged status to the pro-diet desires that move me to action at some times rather than to the anti-diet desires that move me at other times? The fact that my today self is in conflict with my next day's self does not automatically justify the positing of a temporally extended me that contains all of these time slices and that is bound by norms of rationality to exhibit behavior that is interpretable as the maximization of a coherent set of preferences.

In this essay, I shall assume that we can bypass these worries and simply concede that many of us have preferences to save more for our retirement than we actually do save, follow a lifestyle that would increase our life expectancy to a greater extent than does the lifestyle we actually follow, avoid happiness-dampening quarrels and bickering with colleagues, friends, and family members more than we succeed in doing, and so on. To get this result, consider the sum of the person's basic (non-instrumental) desires at each time of her life concerning what shall hap-

---

<sup>3</sup> A warning to the reader: this essay relies on the idea of an agent's will, but does not explicate the notion or defend any particular construal of it. Here I rely on the common-sense idea. But interpretation of it is needed. Here this essay contains a blank that needs to be filled in. For relevant philosophical discussion, see Gary Watson, "The Work of the Will," reprinted in his *Agency and Answerability* (Oxford: Oxford University Press, 2004), pp. 123-57.

pen at times of her life, each desire ranked at each time by her own sense of its importance. Subtract the desires that are conditional on their own continuance and that do not persist. (An example would be the desire to eat cake on one's fiftieth birthday provided one still wants that at age 50.) Subtract desires with respect to some state of affairs that are supplanted by desires concerning that same state of affairs that are better considered, and subtract desires that replace similarly competing desires and that are less well considered than those they replace. I shall assume that this exercise carried out for many of us would yield results that accord with Sunstein and Thaler's hunches that cognitive and desire-driven flaws in our decision-making induce choices that do not effectively bring about the satisfaction of our overall preferences.<sup>4</sup>

Libertarian paternalistic intervention in a person's life is so called because it does not significantly restrict the liberty of the targeted person and because it aims to advance that very person's good. Some commentators object to the label on the ground that these conditions do not qualify interventions as instances of "paternalism" if that term retains its standard meaning. One common view is that paternalism is restriction of a person's liberty against her will for her own good.<sup>5</sup> Sunstein and Thaler nudges do not fit this characterization. Aiming to understand their view in a sympathetic spirit, we should probably interpret them as implicitly committed to an alternative view of paternalism: nudges, at least when they are deliberately undertaken as such, aim to manipulate the decision process of the targeted individual, that is, they exploit the person's disposition to irrational choice-making to induce a choice that the intervener desires. Being manipulative in this way, and also aiming to be to the advantage of the targeted individual, they merit classification as paternalistic.<sup>6</sup>

We can distinguish two sorts of manipulation by one agent in another's decision. In one sort, the manipulator distorts the decision process of the target agent in the sense that the manipulation displaces a more rational decision process that the target agent would have carried out in the absence of the manipulation. In another sort, the manipulation does not distort the decision process in the sense of rendering it less rationally deliberative. Manipulated, I choose A by a rationally defective decision process; left unmanipulated, I would have chosen A or some other option by an equally rationally defective decision process. Manipulations of

---

<sup>4</sup>In this paragraph I do not intend to affirm desire satisfactionism as the correct account of individual well-being (welfare, self-interest). I am exploring the views of Sunstein and Thaler.

<sup>5</sup>For a classic argument for widening the conception of paternalism, see Bernard Gert and Charles Culver, "Paternalistic Behavior," *Philosophy & Public Affairs* 6 (1976), 45-57.

<sup>6</sup>Sunstein further examines the idea of paternalism and the degree, if any, to which we should be opposed to it in *Why Nudge?*

either sort might have the effect of bringing it about that the targeted agent's choice of action is more rationally justified—is more responsive to the reasons there are that are relevant to the choice—or less rationally justified. These effects might be intended or unintended by the manipulator.

Another distinction here is between (1) a manipulative intervention by one person in another's decision-making, intended to achieve some result, when there was an equally effective means available to bring about that result that would not have been manipulative but rather would have involved only rational persuasion, and (2) a manipulative intervention for which an attempt at rational persuasion would not be an effective substitute.

Consider next the prominent view that paternalism is restriction of a person's liberty or interference with her autonomy, against her will, for her own good.<sup>7</sup> Here the idea of a restriction or interference being against one's will needs clarification. One clarification is that a person can commit her will for the future. On this view, if a person makes a contractual agreement, or just a voluntary agreement, to accept the imposition of a restriction on her liberty, subsequent restriction of her freedom for her own good according to the terms of the agreement may not necessarily qualify as paternalistic. If I voluntarily agree to enter a locked rehabilitation facility for a month that will deny me access to alcohol, restricting my liberty to leave the facility during the month in residence or to arrange for a delivery of alcohol to my room is not standardly considered to be restriction against my will, since I have precommitted my will in this respect by agreement.<sup>8</sup>

Another class of restrictions or interferences imposed for the agent's own good would be welcome to the agent if and when she becomes aware of the fact that she is being imposed upon in this way, because the restriction solves a collective action problem. Suppose a law is passed that prohibits bakers from working more than 40 hours per week without

---

<sup>7</sup>See Gerald Dworkin, "Paternalism," in *The Stanford Encyclopedia of Philosophy* (Summer 2014 ed.), ed. Edward N. Zalta; <http://plato.stanford.edu/archives/sum2014/entries/paternalism/>. Dworkin's formulation is that "X acts paternalistically toward Y by doing (omitting) Z" just in case "1. Z (or its omission) interferes with the liberty or autonomy of Y. 2. X does so without the consent of Y. [and] 3. X does so only because X believes Z will improve the welfare of Y." The formulation in the text stipulates that a paternalistic act is one that is done against the will of the targeted agent. An act to which the targeted agent does not consent, but which is not against his will, would count as paternalistic according to Dworkin but not according to the less inclusive stipulation in the text.

<sup>8</sup>The view in the text may require qualification. Sometimes restriction of a person's liberty or manipulation of her choices, against her present will, for her own good, can qualify as paternalistic and presumptively wrong even if the agent's past contrary will in this matter has been embodied in a contractual commitment. Getting clear about this matter is beyond the scope of this essay.

receiving overtime pay, the law aims to benefit bakers, I am a baker, and I welcome the restriction. If it were in my power to retract or sustain the law by my will, I would sustain it. In this case, there is restriction of an agent's liberty for her own good, but the restriction is not against the agent's will and hence fails to be paternalistic.

In a classic case of nudging intervention, a cafeteria food service positions healthy food at eye level and at the front of the display, so that customers are more likely to choose healthy food rather than more calorific snacks. In this case, there is interference in the agent's decision-making, but if the interference is not against the agent's will, it once again fails to be paternalistic. If the unhealthy food were directly in my field of vision, hence salient, I would desire to purchase it and would do so, but if I am glad, or at least not sad, that a trick is played on me that makes healthy food salient and triggers a desire to purchase it, which leads me to do so, the interference is not against my will.

However, the description given so far might be claimed to be insufficient to classify the set-up of the cafeteria food display as non-paternalistic. First, a tiny restriction of liberty is still a restriction, strictly speaking. Second, we might widen the conception of paternalism under review by relaxing the "against his will" clause. On this wider view, if a person is subjected to restriction or interference for his own good, and his will is not (overtly or latently) definitely in favor of what is being done to him, the interference counts as paternalistic.

On the first point, we might just insist that for restriction of liberty to be paternalistic, the restriction must be nontrivial.<sup>9</sup> The second point should prompt resistance. There is little to be gained by arguing about definitions of terms, but it will be useful to understand the notion of *paternalism* in such a way that it will make sense to hold that an act that is paternalistic is pro tanto or at least presumptively objectionable. At least, paternalism directed at mentally competent adults not suffering from debilitating mental illness is presumptively objectionable. Paternalism directed at children and other incompetent agents might be perfectly acceptable, but paternalism directed at competent adults is treating them as though they were children and certain ways of doing this fail to be adequately respectful of the status of the adults as rational agents.

After this brief digression, we return to the question of whether it is a good idea to define *paternalism* in such a way that interfering with a person's liberty for her own good in a way that her will does not oppose, but does not actively support, should be counted as paternalistic. Consider this example. Hiking along a trail, a stranger gives me a shove, to prevent

---

<sup>9</sup>We might adopt a scalar notion of paternalism: the more restrictive of the liberty of the person targeted for helpful interference, the more paternalistic the interference.

my stumbling on a tree root and falling. Had I not been shoved, I would have suffered injury. If I have no definite will regarding encounters of this sort, I see nothing pro tanto or even presumptively objectionable in the stranger's conduct. So it is better to stick with the requirement that paternalistic imposition is against the will of the imposed-upon individual.

Let's recapitulate the discussion to this point. We seek to specify a notion of paternalism that will render plausible the claim that paternalistic action and policy are pro tanto morally wrong. The suggestion is that paternalism is restriction of a person's liberty or manipulation of her choices, against her will, for her own good.

If paternalism understood in this way is wrong, the idea is not that certain restrictions and manipulations ought not to be done, but that they ought not to be done for certain reasons. If I am beating up the neighbors, you decide that this course of action is likely to be detrimental to me, for this reason you force me to stop carrying out this assault, and I do not want to be treated this way, what you are doing is paternalistic and pro tanto wrong. This judgment leaves it entirely possible that forcing me to stop my assault in order to protect the victims of my assault would be perfectly fine and even mandatory.

The upshot is as follows: if nudges involve action or policy directed at an agent that aims to exploit his decision-making deficiencies for his own good, they are manipulative, and so are interferences with the agent's autonomy, and so can be paternalistic. Whether they are so depends on whether or not they occur against the will of the targeted agent.

## 2. Negative Judgments of Competence

Some suggest that the conception of paternalism outlined here is too narrow. Amending a proposal advanced by Seana Shiffrin, Jonathan Quong defines *paternalism* as an act by a person A directed at a choice problem that another person B faces and aimed at advancing B's welfare or interests, the act being motivated by a negative judgment about even a fully informed B's ability "to make the right decision or manage the particular situation in a way that will effectively advance" B's own welfare or interests.<sup>10</sup> Paternalism construed in this way involves overriding or pre-

---

<sup>10</sup>Jonathan Quong, *Liberalism without Perfection* (Oxford: Oxford University Press, 2011), p. 80. See also Seana Valentine Shiffrin, "Paternalism, Unconscionability Doctrine, and Accommodation," *Philosophy & Public Affairs* 29 (2000): 205-50, p. 218. Shiffrin suggests regarding as paternalistic any interaction with another intended to help that is motivated by a negative judgment about his competence and a substitution of judgment regarding an action that is within his legitimate sphere of authority. This essay does not adopt, and sets aside without comment, this suggested widening of the idea of paternalism.



empting the judgment of the person one is trying to help about how to manage her own life. Simply offering reasons in an attempt to persuade the person to accept your judgment does not qualify as paternalistic. Shiffrin gives this example: You ask someone you know to help you with a project. The person is under no duty to help but would help except that she judges that you would be better off working through the project on your own, and declines to help on this basis. Declining to help for this reason is paternalistic.

The trouble with the Shiffrin-Quong idea of paternalism is that there is nothing *prima facie* objectionable about believing that you know better than another individual what is good for him or what he should do in order to advance his interests effectively. Nor is there anything objectionable *per se* about acting toward the person in ways that only make sense on the assumption that you believe you know better than the person does herself in the just-mentioned respects. It is inappropriate to believe you know better than some other individual when you do not, and you can be culpable for such beliefs if you should have known better or should have investigated further before making up your mind on this point. It is wrong to insult another person gratuitously by calling attention to his alleged inadequacies, as you perceive them, to no purpose. But suppose you have knowledge that my vacation plans are poorly reasoned and deficient. Carrying through my plans will be bad for me, maybe even dangerous. Declining to cooperate with me or make deals with me that you would to be willing to undertake except that you would be facilitating a futile enterprise would not be even presumptively wrong. I have no right that you become complicit in futility. It is not generally the case that if it is otherwise permissible for you to do X, doing X becomes impermissible if your doing it would be motivated by the thought that another is making a mistake and that your refraining from doing X will make things worse for him.

So far I am merely denying what Shiffrin and Quong assert. The reader may object that no progress is being made. In response, I would suggest that if we consider cases in which one acts with a view to helping someone on the basis of a negative judgment about the person's decision-making competence or capacity to execute whatever decision is made in the situation, and we take care to describe the cases so that no other wrong-making features accompany them, we simply will not share their negative verdict concerning the cases. I am not saying that Shiffrin or Quong misread the examples, but as readers of their work, we might. Finally, I note that so far as I can see, they offer no reason beyond an appeal to intuition to support their position. If we do not share their intuition about the examples they describe, there is no further case to rebut.

Can more be said? One might appeal to the general idea that when a

variety of reasons point in different directions regarding what to do, one should take all into account and be willing to be moved to action by their overall balance. Resolving to pay no heed to a class of reasons is irrational, unless one has good reason to suspect one is likely to respond to them so irrationally that they are better ignored in practice.

Another possibility would be that if in practice we will be acting on uncertain judgment in any event, and the risk of being moved by one type of reason if we proved to be mistaken would be expectably worse than the risk of not being moved by that type of reason if we were mistaken in resolving to ignore it, we should resolve to ignore it. But in the absence of some specific argument to the conclusion that acting on a mistaken negative judgment about another's competence is a moral catastrophe, it would just tilt the scales unfairly against paternalism to resolve not to act on the basis of such a claim to superior judgment. But being the target of a mistaken judgment about one's competence is an everyday humdrum occurrence, not in itself a moral catastrophe.

One concern might be that refraining from paternalism based on negative judgment of competence is connected, in contingent and non-contingent ways, to achieving and sustaining the ideal of a society of equals, in which none is markedly inferior in status, rank, or power. In this connection, Jeremy Waldron comments that whereas he would not object to a society in which Thaler-Sunstein nudges occasionally occur, he would strenuously object to a nudge society, in which the manipulation of some by others is pervasive in social life.<sup>11</sup> Waldron might especially be reacting against the prospect of a society in which there is a distinct class of nudging experts and a separate class of subordinate nudged nonexperts. Interpreting and assessing the ideal of a society of equals would be a large task, one beyond the scope of this essay. I simply note that in a democratic society, paternalism proceeding from claims of superior competence need not impede the achievement of a society of equals and might facilitate it. For one thing, people might switch roles from the one who carries out paternalism to the one who is its object, so no significant deviation from a society of equals develops over the long run (although it remains true that some might object to the envisaged trading-places egalitarianism). Another consideration is that sometimes paternalism proceeding from claims of superior competence, when the claims are correct, is necessary to prevent misfortunes from falling on the targeted individuals that would themselves undermine the maintenance of a society of equals. And one should anyway be aware of the possibility that the concern to maintain a society of equals might be overblown and might proceed from lines of

---

<sup>11</sup>Jeremy Waldron, "It's All for Your Own Good," *New York Review of Books*, 9 October 2014.

thought that we should reject for reasons similar to those that should lead us to balk at any sort of strict anti-paternalism.

### 3. Libertarian and Semi-Libertarian Paternalism

An interesting feature of Sunstein and Thaler's libertarian paternalism is its odd nod to libertarianism. Without much discussion, they propose that the aim of government policy in this domain is to help individuals live better by their own lights. The paternalism they favor "tries to influence choice in a way that will make choosers better off, *as judged by themselves*."<sup>12</sup> A Sunstein-Thaler policy is successful in improving your life if and only if it increases the degree to which (you judge that) your own preferences about your life are satisfied.<sup>13</sup> The background thought is perhaps that the individual is sovereign over her good in the sense that the standard for assessing the degree to which her life goes well is set by her own subjective preferences, not by any external standard.

Of course this idea of personal sovereignty is miles apart from the libertarian ideal of personal sovereignty. This latter ideal holds that the individual who is not violating the rights of others should be left free to act as she voluntarily chooses even if her free and voluntary choices entirely shred her long-term prospects of preference satisfaction.

Thaler and Sunstein are not embracing the libertarian ideal of personal sovereignty. Their proposal is to enact judicious paternalism that is liberty-preserving.

Libertarian paternalism as characterized by Sunstein and Thaler reveals itself to be an unstable normative position. The problem is not that whether or not what they take to qualify their view as libertarian chimes in with other characterizations of the doctrine. The problem is why the agent contemplating paternalistic intervention ought to eschew restriction of liberty in this enterprise and confine her interventions to manipulative ones. (Restrictions of liberty limit the options among which individuals choose, or change the economic incentives individuals have to choose one or another option.) Restricting people's liberty will sometimes influ-

<sup>12</sup>Thaler and Sunstein, *Nudge*, p. 5 (emphasis in original).

<sup>13</sup>The reader will notice that these formulations slide around between different ideas. One idea is that you are better off, over the course of your life, the more your preferences are satisfied. A different idea is that you are better off, over the course of your life, the more it is the case that you judge that your preferences are satisfied. Yet a different idea is life satisfaction: your life goes better for you, the more your life as a whole satisfies your own standards for a life that goes well. Yet another idea is subjective life satisfaction: you are better off, the greater the extent to which you judge that your life as a whole satisfies your own standards for a life that goes well. L.W. Sumner defends a qualified life satisfaction view in his *Welfare, Happiness, and Ethics* (Oxford: Oxford University Press, 1996).

ence choices in a way that will make the choosers better off, as judged by themselves. The background idea to which they appeal suggests more paternalism, or anyway a wider range of paternalism, than they endorse.

Restriction of liberty might seem more onerous than subtle manipulation of choice even when care is taken to make the existence of the manipulation transparent to the agent whose choice process is being nudged. But this does not seem to be necessarily so. To take a nonpaternalistic example, the massive restriction of liberty involved in traffic safety laws that are regarded as broadly sensible fades into the background of ordinary life and is not ordinarily perceived as more than a minor inconvenience.

Another source of instability is taking the existence of pervasive and pretty much unavoidable defects in ordinary human decision-making as revealed by behavioral decision theory as a basis for welfare-improving intervention while insisting that justifiably paternalistic intervention must improve people's lives according to their own assessment or by the measure of lifetime preference satisfaction. After all, the cognitive limitations that beset human choice-making and that open the opportunity for interventions that nudge people toward better choices no doubt also play a mischievous role in the processes by which values and aims and basic preferences are formed. Human frailty impedes our attempts to choose effective means to our goals and execute our choices, and human frailty likewise impedes our attempts to identify and embrace worthwhile and truly fulfilling goals. This point holds of value-forming and choice-making below the threshold of substantial voluntariness but also above it.

The libertarianism from which the Thaler and Sunstein position needs to be sharply distinguished is not the full political doctrine of Lockean libertarianism, but a narrower view that we might call *semi-libertarianism*.<sup>14</sup> In this view, each individual adult person who is not crazy and has normal cognitive powers has full sovereignty over herself, full self-ownership within the constraints set by the moral duties that she owes to other persons, which may include positive duties to aid. Within these limits, each person should be free to live as she chooses without suffering coercive restriction of her liberty at the hands of others intending to benefit her. In this moral vision, each person at each moment of time is the rightful trustee of her future interests.<sup>15</sup> If the individual now has the capacities of the

<sup>14</sup>Robert Nozick, *Anarchy, State, and Utopia* (New York: Basic Books, 1974).

<sup>15</sup>In the late twentieth century, the most able defender of semi-libertarianism was Joel Feinberg. He elaborated the position specifically as part of his comprehensive vision of the moral limits of the use of the criminal law sanction. See Joel Feinberg, *The Moral Limits of the Criminal Law, vol. 3: Harm to Self* (Oxford: Oxford University Press, 1986). For criticism, see Richard J. Arneson, "Joel Feinberg and the Justification of Hard Paternalism," *Legal Theory* 11 (2005): 259-84.

normal competent person and is acting in a substantially voluntary way, she gets to decide how much weight to give to the interests and desires she will predictably come to have in future.

For the semi-libertarian, forcing a person who is committed to a choice made in a way that qualifies as substantially voluntary to engage in further deliberation about this choice or to confront information concerning it, in order to advance the person's welfare, would be wrongfully paternalistic. In this view, John Stuart Mill's example of the man who is blocked from venturing onto an unsafe bridge is underdescribed in his account of it. If the person's will is to proceed without considering further information about the bridge's safety, forcing the person to consider further information restricts his liberty against his will for his own good. (If stopping the person and imposing bridge condition information on him provides him an opportunity he welcomes, the intervention is not paternalistic.)<sup>16</sup> Moreover, we could expand semi-libertarian paternalism so that it also rules out manipulative interference in a person's decision process against her will as well as restriction of liberty.

Although Sunstein and Thaler give a deferential nod to semi-libertarianism, there is no reason to believe they are committed to it. The position that paternalistic interference is justifiable if and only if it increases the overall lifetime preference satisfaction of the individual who is subject to the interference (provided the interference does not wrongfully or excessively impose costs on others and is not overly burdensome to the interfering agent) conflicts with semi-libertarianism as just described. If Sunstein and Thaler hold the lifetime preference satisfaction view of justifiability, their position is not libertarian at all, not even semi-libertarian. Consider the individual who with full information chooses in a substantially voluntary way to pursue her present lesser preference satisfaction at the cost of her future greater preference satisfaction. The satisfaction of preferences she will have in future, right now, leaves her cold. Sunstein and Thaler should be willing to countenance nudging interferences in such cases; the advocate of semi-libertarianism will oppose them, at least if they qualify as manipulative interferences and so paternalistic.

#### 4. A Moderate Framework on Justifiable Paternalism

At the opposite extreme from the semi-libertarian position on the limits of acceptable paternalism lie utilitarianism and more broadly welfarist

---

<sup>16</sup>See the discussion of Mill's unsafe bridge example in Sunstein, *Why Nudge?* pp. 106-7. See also David Brink, *Mill's Progressive Principles* (Oxford: Oxford University Press, 2013), for discussion of Mill on paternalism.

consequentialism. The welfarist consequentialist holds that one morally ought always to choose the act among the valuable alternatives that maximizes the morally correct function of overall human well-being; for the utilitarian, the correct formula is that one should maximize the aggregate sum of individual well-being. On any such view, when one's actions can affect only oneself, the same principle determines right conduct, so in those situations one morally ought to choose the act among the alternatives that would maximize one's own well-being. Anything else one might do instead is morally wrong. In principle, for the welfarist consequentialist, there is just as much reason to interfere paternalistically in people's conduct, restricting their freedom and interfering with their autonomy for their own good, as there is to interfere in people's conduct that is wrongful by virtue of failing to give due consideration to the interests of persons other than the acting agent and those who voluntarily consent to share the effects of what he does.

Between these extremes, there will be a range of middle-of-the-road positions concerning justifiable paternalism. Some will relax semi-libertarianism by raising the standards for what qualifies as substantially voluntary choice. The prohibition of paternalism is exceptionless, but the paternalism that is forbidden is hard paternalism that interferes with choice that meets the standard of being sufficiently voluntary and thus is deemed to be proceeding from a will of the agent's that is truly hers. Another possibility would be to introduce a sliding scale, so that the constraint against paternalism becomes firmer, the more voluntary the choice of the agent that is subject to interference for his own good. The more voluntary the choice, the greater must be the loss to the agent that paternalism prevents, to be justified.<sup>17</sup>

What follows is a sketch of a framework in which to place intermediate positions between the polar extremes of semi-libertarianism and welfarist consequentialism.<sup>18</sup> Assume first that morality includes a signif-

---

<sup>17</sup>This position is defended in Feinberg, *Harm to Self*, chap. 20.

<sup>18</sup>I mean to suggest a framework in which those who reject both welfarist consequentialism and Feinberg-style semi-libertarianism as too extreme might consider how to decide the limits of acceptable paternalism. I do not here argue against either of the extreme positions. (For the record, my own inclinations are toward welfarist consequentialism.) I am exploring reasons for moderates to eschew rigid anti-paternalism, at least in principle. Further clarification: this essay suggests conditions in which paternalism would be morally permissible and does not discuss the extent to which these conditions do or do not obtain. Considerations against paternalism raised by such writers as J.S. Mill in *On Liberty* and Edward L. Glaeser in "Paternalism and Psychology," *University of Chicago Law Review* 73 (2006): 133-56, lie outside the subject matter of this essay. I simply accept the behavioral science evidence that individual choices tend to be defective in ways that bring it about that in principle, another agent, who happens to be insulated from the defects in this instance, might intervene to improve the outcome in terms of the good or

icant beneficence component, a standing duty to choose actions that bring about the impartially assessed best outcome one can reach within the limits and exceptions to be described. Assume that there may be moral rights that people have that constrain what it is permissible for us to do to them even in the service of the general welfare (the maximizing impartial consequentialist goal). The individual is bound to respect these constraints, unless the consequences of not violating them would be very bad. Besides moral constraints that restrict people's freedom to act as they choose, there are also moral options that expand their morally protected freedom.

Here we can appropriate and amend a suggestion made by Samuel Scheffler.<sup>19</sup> He proposes that each person has a personal prerogative, to some degree, to act so as to advance projects and interests that she favors, even though these actions bring about less than the best outcome that she might instead have brought about. We can think of the personal prerogative as constrained by the moral rights that people have, which also constrain the permissible pursuit of the best reachable outcome on any occasion of choice. To get the issue of paternalism in focus, let us restrict attention to cases in which the agent's actions affect only his own interests, and suppose that the impartially best outcome in this range of cases just is the outcome that maximally advances the agent's own interests. We can think of the personal prerogative as allowing one, to an extent, to do what one wants, even if this results in an outcome that does not maximally advance one's own interests. More generally, the personal prerogative permits one to act in any way one chooses within moral constraints so long as the shortfall between the impartially assessed value of the best outcome that one might have brought about and the impartially assessed value of one's actual course of action does not exceed a threshold.

This essay does not delve into the issue of how best to conceive of the structure of the prerogative.<sup>20</sup> For our purposes, we just need the ideas that there may be a personal prerogative to bring about less than best and that the permission it confers to bring about less than best sets some limit

---

well-being of the defectively choosing individuals.

<sup>19</sup>Samuel Scheffler, *The Rejection of Consequentialism*, revised ed. (Oxford: Oxford University Press, 1982), chap. 2.

<sup>20</sup>One complication should be noted. Whether or not a self-affecting choice of action is protected by the Scheffler prerogative might be deemed to depend not only on the shortfall between the best one might have done for oneself and the actual outcome, but also on how costly and difficult it would be to make and carry out a choice that would be substantially better for oneself. If so, the moderate should allow that there might be justifiable paternalistic interference of action that is within the range of the Scheffler prerogative, when the shortfall is large and the course of action, if subjected to restriction, would not be costly or difficult for the agent to avoid. In this case, the fact that action runs afoul of the prerogative and thus violates a duty to self would suffice to open the door to justifiable paternalism, but would not be necessary for it.

on the badness of the outcome one is allowed to choose.

This framework immediately yields the result that insofar as the acts that persons might choose will mainly have welfare effects on their own lives, the prerogative sets limits on the degree to which one is permitted to pursue actions that will lead to welfare deficiencies for oneself compared to other courses of action one might instead have pursued. In effect, the framework, coupled with the claim that there is a personal prerogative and that it is limited in extent, imposes among other things a moral duty to self, a requirement to be prudent rather than imprudent in one's life choices (unless one's imprudence brings about improvements in the lives of others that outweigh the losses incurred by oneself). The framework in effect imposes on each person a vague moral duty to make something valuable for oneself and others of the opportunity one has to lead a human life that (except in rare tragic circumstances) is replete with chances for getting good and doing good.<sup>21</sup>

Let us stipulate also that if one's self-harming action does not exceed the allowable shortfall set by the prerogative but would bring one's overall lifetime well-being level sufficiently far below some minimal acceptable level, without achieving compensating benefits for others, this global shortfall might also be deemed to render one's act in violation of the duty to self.

Though vague as so far characterized, the duty to self has significant content. For example, it surely rules out suicide in all cases in which continued life would provide significant overall benefit for oneself without also imposing countervailing losses on others. The duty to self rules out gross neglect and reckless unconcern for one's future welfare. The prerogative with its duty to self also sets limits on the degree to which one is permitted to be self-abnegating, acting to bring about small aggregate gains for others at the expense of greater losses that fall on oneself. There are limits to morally permissible imprudence.

The idea of a duty to oneself may seem pointless, because one can always consent to one's own action that conflicts with the duty, eliminating the conflict. But this is contestable. Each of us might be bound by a nonwaivable duty to self to make something good of one's life. Saying this assumes that one can have a directed duty to oneself, a duty owed to oneself. In the normal case of a directed duty, the person to whom the

---

<sup>21</sup>The duty to self as stated implies a duty in part to make something good of one's life by doing good for others. In further discussion, I assume we are dealing with situations in which the duty to self binds the agent because the action she is disposed to choose would violate the duty by harming herself or failing to benefit herself as she might, or by being self-abnegating (helping others in ways that impose excessive costs on oneself so that there is net loss as assessed from the correct impartial perspective).



duty is owed has the prerogative of waiving the right to its performance. I have promised to come to your party, but you say "Don't bother," and I am released. On this familiar picture, the idea of having a directed duty to oneself can seem at the borderline of incoherence.

But there are cases and cases. I might have a directed duty to protect a child from harm, but we might deny that the child can release me from the duty by telling me "Don't bother." We might regard the duty to oneself as a duty to oneself regarded impersonally as one person among others, whose good we are unconditionally bound to promote, up to a point. Viewing my duty to self in this way, I cannot release myself from this duty and then without wronging myself sail off to hell in a handbasket, if I am so minded.

In fact, this duty automatically arises from assumptions about moral requirements that many will be disposed to accept: (1) against consequentialism, we affirm that there are (a) moral constraints that limit the permissible pursuit of good outcomes and (b) moral options not to choose one of the actions that would lead to the very best reachable outcome; (2) against libertarianism, we affirm that morality includes a serious enforceable beneficence requirement that requires us to promote good outcomes; (3) against semi-libertarianism, we affirm that morality's beneficence requirement is impartial and requires that each agent deciding what to do should seek to achieve gains and avoid losses for herself as one among those who might be benefited or harmed by her action. Assumptions (1)-(3) together imply that one has a duty of beneficence that requires concern for self as well as concern for others. (The duty to self affirmed here may appear to go beyond (1)-(3) in requiring each person to have a special asymmetrical concern for herself beyond her general duty to others. Do I have a duty toward everyone to help each person make something valuable of her own life by doing good for self or others? Do I have a duty to help prevent anyone's life from falling too far below some minimal acceptable well-being level? This issue looks to be delicate. In this essay I set it aside.)

It does not obviously follow from the fact that there are limits on permissible imprudence that paternalistic restriction of individual freedom and interference with autonomy are morally acceptable. Maybe people should not be imprudent, but if they are, others should still not interfere.

It is plausible to think that the nerve of anti-paternalism is a conviction that at least when the individual's conduct only affects herself, it is wrong to interfere with her conduct against the individual's will to act as she chooses and not suffer interference. If this is so, the framework proposed in this section suggests that the agent's will, set against interference, can be defective in character, in a way that reduces or eliminates

the moral presumption against interference.

Speaking of a defective will can be misleading, as this suggests the familiar idea that defects in the process of an agent's decision-making that lead her to choose and pursue a course of action can open the door to justifiable paternalism (or to interference that would qualify as paternalistic but for the lack of voluntariness). This is not the defect in question here. Let the agent's decision-making be as considered and careful and in all respects as voluntary as you like. If the agent decides to pursue a course of action that is self-harming to an extent that is beyond the range that is protected by the Scheffler prerogative, the agent in choosing exhibits a defective quality of will. What the agent is aiming at is excessively subpar. I would add that in such a case, the agent violates a duty to self to act in ways that bring about consequences for herself that are at or above a threshold level of goodness, this level being set as a function of the best outcome for herself that she might instead have brought about (we are assuming that the agent is choosing among options that will have no impact on others). In this case, the content of the agent's will is defective in a way that opens the door to acceptable paternalism, according to the suggestion I am making.

Suppose the law forbids the recreational use of certain dangerous drugs, and those who support the legislation do so on the ground that it will benefit (most) people whose freedom the law restricts even though those restricted are of a different opinion. I might oppose the law in its application to my own choices on the ground that if left free to choose, my drug-taking would be judicious and on the whole beneficial for me. I might recognize that this conviction of mine might be correct or incorrect. Regarding my current will, one can ask, would I still want not to be subject to interference, if it turned out that my current conviction is incorrect?

Suppose the only person I might harm by what I do is myself. But I want to live my life making choices not constrained by meddling interference in my life paternalistically motivated. I want to live free from paternalistic constraint whether or not my life choices made freely would result in my being better off than I would be if subjected to some such constraint. In part I might regard making my own choices free from constraint as itself enhancing the quality of my life, independently of the quality of the outcomes for me that those choices bring about. But I am willing to countenance the possibility that even taking into account the value of free choice, it still might be the case that all things considered, my life would go better, and expectably better, under a regime of some paternalism than under a regime of no paternalism. Nonetheless, for myself, I prefer to live free from paternalism even at some cost to my welfare.

This stance does not look to be on its face perverse or unreasonable. If restriction of my liberty for my own good is contrary to my will, and my will has the content just described, there does look to be significant ground for holding that such restriction is morally objectionable.

But look a bit further. Let's distinguish two possibilities. One is that my life, unburdened by the paternalism I oppose, would go somewhat worse for me, but not greatly worse, and my life overall would still be of acceptable quality. In this case, we might take the firm disposition of the person's will against paternalism as strong reason to find it objectionable policy. Suppose instead that my life will either go very substantially worse, or will fall below some minimal acceptable quality, or both together, if those contemplating paternalistic restriction of my liberty refrain from imposing on me. In other words, the choice I would make conflicts with the duty to self set by the personal prerogative.

Now consider my present will, opposed to paternalism, and ask me whether my being disposed against being subjected to such treatment is conditional on its being the case that being free from paternalism does not bring it about that my life goes substantially worse than it would if paternalistically fenced in or would fall below some minimal acceptable quality if not paternalistically fenced in. If the answer is "No," my quality of will now begins to look deficient. My opposition to paternalism in this case is arguably cavalier in showing insufficient concern for my own welfare.

If paternalistic interference in people's lives (restricting their liberty or manipulating their choice-making or both) is contrary to the determined, unconditional will of people who believe their unimpeded actions are within the scope of the personal prerogative but who would continue to be opposed to interference even if their action is beyond its scope and is in violation of the duty to self, the reason against paternalism stemming from the opposition of their will is lessened or eliminated by virtue of the defective content of their will. On the other hand, if paternalistic interference is contrary to the conditional will of people who believe their unimpeded actions are within the scope of the personal prerogative but would not continue their opposition in the event their belief is false, their underlying or deeper will concerning this latter scenario arguably is not against interference after all.

The philosopher who points to the evil of forcing or manipulating people against their will as the basis for anti-paternalism then apparently faces a dilemma in case her anti-paternalism is directed at situations in which the paternalistic action or policy is otherwise justifiable but for this opposition of will. Suppose the proposed paternalistic action is net beneficial, achieves more benefit overall than nonpaternalistic alterna-

tives, does not wrongfully harm or neglect third parties, and targets courses of action that would be unprotected by the personal prerogative. But the action is against the will of the intended beneficiaries. Then either the will of these people is unconditional as described above, hence defective, hence unavailable as a reason against interference, or it is conditional as described above, in which case the interference being proposed would not be against the underlying will of the individuals targeted for interference. In neither case do we get justified opposition to paternalism. This will be so even if the paternalism under review is interfering with substantially voluntary choices and actions, and even if the paternalism is directed at helping people achieve goals that they themselves do not embrace or at any rate do not regard as sufficiently worthwhile to warrant efforts to achieve in their situation.

The dilemma just sketched is only apparent, however. This emerges once one notes that the person who is choosing a self-harming act that he believes is not excessively self-harming, to the point that it would run afoul of the Scheffler prerogative, and whose will is set against paternalistic interference, may simply not have any definite will now as to whether he would oppose interference in the event that his belief that his choice lies within the range protected by the Scheffler prerogative is mistaken.

One possibility here is that he is not consciously aware of whether he would oppose interference if he were to become aware that his belief that his choice is not excessively self-harming is false, but still he is disposed now to choose one way or another in that circumstance. If the individual has a disposition of this sort, then his dispositional will is either defective in content or not, and the dilemma sketched two paragraphs back is reinstated.

Suppose instead that the individual we are describing just has no disposition one way or another regarding the hypothetical. Nothing in his personality inclines him to choose one way or another in that case. If the hypothetical situation arose, some slight feature of the situation would trigger him to go one way or another or remain indifferent. We should assimilate this case to the one in which the person is disposed to continue to resist interference even in the event that the excessively self-harming character of his behavior that is attracting interference is known to him. In both of these cases, the person lacks a will to abide by the Scheffler prerogative duty to self, and in this respect, the content of the person's will is defective.

The skeptical reader may suppose that I am appealing to a "real self" doctrine that Isaiah Berlin once and for all exposed as "monstrous impersonation" decades ago in his essay "Two Concepts of Liberty." Berlin wrote:

It is one thing to say that I know what is good for X, while he himself does not; and even to ignore his wishes for its—and his—sake; and a very different one to say that he has *eo ipso* chosen it; not indeed consciously, not as he seems in everyday life, but in his role as a rational self which his empirical self may not know—the “real” self which discerns the good, and cannot help choosing it once it is revealed.<sup>22</sup>

The position defended here is not that the person who is being considered as a target for paternalistic interference may be ascribed a rational preference for his true good independently of any empirical evidence for this ascription. The claim is that he might or might not have an underlying disposition of will to accept interference that is a proportionate response to his violation of a duty to self, in the event, contrary to what he now believes, that his behavior really does constitute a such a violation. If he does have this disposition of will, there is something in his empirical self that is now actually so disposed, whether or not we can in practice detect it. The claim then is that if he does have this empirical disposition, then paternalism does not run contrary to it, and this suffices to show that the paternalism (if otherwise unobjectionable) is justified in this case. The further claim is that if he lacks such a disposition of will, this itself shows his will to be defective in a way that also eliminates the moral objection to paternalism directed at his excessively self-harming conduct.

The account presented so far seems vulnerable to objection. Even if the reader accepts the claims made to this point, all that has been asserted is that when a person's self-harming conduct is excessively self-harming, it violates a moral duty to self, and that when this is so, if the person's current will is against paternalistic interference and she is not disposed to give up her will to interference were she to come to understand that her behavior is excessively self-harming, her current disposition of will is defective. So what? It does not follow that it is not wrong to interfere with the person's liberty to go forward and act wrongly in this way. For it is surely not generally true that if a person is behaving in a way that is morally wrong, it is morally acceptable for anyone to block her from doing it.

On the general point: let us accept that a person might have a moral right to do X, even if exercising her moral right to do X on this occasion would be morally wrong, and that at least on some occasions of this sort it would be wrong for others to interfere with the person's doing X. In the case at hand, we are trying to determine the limits of the moral right against paternalistic interference. The suggestion is that it is extinguished at the limit of the Scheffler prerogative, when we are considering situations in which only the agent will be affected by the choices she might make.

---

<sup>22</sup>Isaiah Berlin, *Four Essays on Liberty* (Oxford: Oxford University Press, 1969), p. 133.

If we consider the question of whether we may permissibly interfere with another's conduct when paternalism is not at issue, the fact that the act we are proposing to block would be morally wrong and would violate a moral duty owed to a third party goes a long way toward opening the door to justified interference. Suppose that you are considering stopping me from being mean to my brother. If I am withholding candy from him that he wants, but that I have no duty to give to him, that is a significant consideration that should inhibit you from interfering. If on the other hand I have a moral duty to give him the candy (suppose he is diabetic and needs a blood sugar boost), matters are different. I have no right against interference. I am inclined to think that unless there are pressing moral considerations extraneous to the situation that bear on what we should do, in this situation if you can interfere with my wrongful plan of action, securing the candy for my brother, that is morally permissible to do unless the costs of interference are disproportionate to what is at stake, as would be the case, for example, if your interference would result in my death and nothing close to that bad is in store for my brother.

Now turn to paternalism. According to the framework being proposed, my self-harming action that exceeds the harmfulness that is allowed under the Scheffler prerogative is morally wrong and violates a nonwaivable directed duty to myself. The suggestion I am advancing is that when this is so, then even if my will is firmly disposed against paternalistic interference, that will would have a morally defective character, and not generate in itself any reason against paternalistic interference. Say there is a moral presumption against paternalism, construed as restriction of an agent's freedom or manipulation of her choice, against her will, for her own good. When the agent's will is defective as just described, the moral presumption against paternalism is extinguished.

I do not in this essay attempt to argue for the framework proposed. Nor for that matter do I argue that the welfarist consequentialist treatment of paternalism should be rejected. That's as may be. My suggestion is that if you identify the presumptive wrongness of paternalism with acting against the targeted individual's will in a matter that concerns only herself, you should be open to considering the moral character of the agent's will as setting the line where presumptive wrongness materializes.

## 5. Debunking the Normative Significance of the Will?

The reader might be thinking, what is sauce for the goose should be sauce for the gander. The claim being advanced is that the fact that paternalistic interference with choices one might make would be opposed to one's will that is defective by virtue of failing to accord with the duty

to self suffices to eliminate the pro tanto moral wrongness of that interference. If this claim is accepted, then by the same token, the fact that some interference in one's choice-making and conduct, motivated by concern for one's welfare, would not be contrary to one's will does not in itself establish that the interference is pro tanto morally permissible. One's will here may be defective; perhaps one ought to oppose the interference. If so, the fact that one's actual will is not opposed may be of little or no normative significance.

On the view being proposed in this essay, interference with a person's freedom to act motivated by concern for her welfare is not paternalistic unless the interference is contrary to her will. Such interference does not qualify as paternalistic, whether or not it is morally permissible.

On the view here proposed, what the agent actually wants and does not want and is disposed to accept and oppose in the way of interference in her choices that affect only herself does play a role in determining when paternalistic action is justified, albeit a moderate role. If the agent's act is self-harming, but within the limits of the Scheffler prerogative, restriction or manipulation of her liberty for her own good would not qualify as paternalistic if she is not opposed to such interference. So here the orientation of the agent's actual will does matter. If in these scenarios of moderately self-harming conduct the agent's will is opposed to interference, then interference would be paternalistic, and presumptively wrongful. Beyond the range of the Scheffler prerogative, the orientation of the agent's will matters not so much.

In this essay, I have not been considering in what ways the character of the process by which an agent's will is formed might affect the justifiability of paternalism. Suppose my act is only moderately self-harming, but my decision to act was made when I was drunk, and poorly informed about the consequences of my choices, and in an angry or depressed mood. Do these sorts of facts that reduce the voluntariness of an agent's choice lower the moral bar against paternalistic interference, and if so, how, why, and to what extent? This essay does not consider these issues.

## **6. Implications Regarding the Assessment of Innocently Erroneous Conscience?**

The suggested analysis of the impact of the quality of will of an agent who might be subjected to interference in order to dissuade him from a chosen course of action on the moral permissibility of such interference appeals to the possibility of conflict between actual will and morally required will that can obtain in situations in which paternalism is not the issue. Suppose morality requires an individual to participate in waging a

just war, but the individual is conscientiously opposed. She has non-culpably formed the belief that participation in this war would be morally wrong. Suppose it is urged on this basis that we ought not to restrict the person's liberty to refrain from participating or manipulate her choice-making to induce her to choose to participate. We ought to respect the person's liberty of conscience. Here one could press the same dilemma I have urged on those who oppose interference in individual choice-making on paternalistic grounds when the conduct being interfered with runs afoul of the duty to self. Either the conscientious agent has a will now that is disposed to accept the requirement to participate in the war if it were to turn out that her judgment that participation would be immoral is erroneous, or she does not have such a will. In either case, the opposition of her will does not plausibly ground a moral requirement not to interfere with her conscientious choice. Given that participation in this war is indeed morally mandatory, either her underlying will is not deeply opposed to interference or her underlying will is morally defective in a way that undermines the claim to liberty of conscience. A critic of the position tentatively espoused in this essay might hold that surely we ought not to interfere in the conscientious choice-making and conscientious action of agents, at least in many circumstances, so the analysis developed in this essay that implies a contrary verdict must be wrong. Hence its application to paternalism is also suspect.

In reply: I agree that the analysis developed here and applied to paternalism applies more broadly. The analysis of its implications regarding the duty to respect conscience must await another essay.<sup>23</sup> I would simply point out that the possibility that the analysis might unsettle some people's views regarding the duty to facilitate people's acting without interference on their conscience even when it is erroneous is not in itself a black mark against the analysis. Perhaps some people's views in this matter are wrong.

## 7. Conclusion

Recent writings on "libertarian paternalism" by Sunstein and Thaler should encourage us to see manipulation as sometimes paternalistic, along with restriction of an agent's liberty, when motivated by the agent's own good. Whether manipulation or restriction of liberty done for an agent's good are paternalistic depends on whether they are against the agent's will.

---

<sup>23</sup>For a brief foray into this issue, see Richard J. Arneson, "Against Freedom of Conscience," *San Diego Law Review* 47 (2010): 1015-40.



This essay has resisted recent suggestions to construe paternalism more broadly as actions intended to benefit an agent in a situation and motivated by negative judgment about her competence in that situation. The resistance is based on the thought that action so described is not plausibly regarded as presumptively wrong. This thought is especially compelling in the case of interactions among friends, family, and associates, but holds also in the case of state action, or so I would hold. (I have not argued for this last claim.)

If what makes paternalism presumptively morally impermissible is that it is opposed by the agent's will, to identify the limits of the presumption we need to engage in moral assessment of the agent's will that is opposed to interference in self-harming acts. In this essay, I have proposed a simple framework in which to conduct this exercise. Its main building blocks are the ideas of the Scheffler prerogative and of a moral duty to self. The upshot is that when the agent's will against paternalism is defective in character, the moral presumption against paternalism—if such there be—is extinguished. (I put in the qualifier “if such there be” because this essay has had nothing to say against a pro-paternalist welfare consequentialism.)

This framework allows that paternalism directed against substantially voluntary actions and aiming to help an agent achieve goals she does not herself embrace might be morally permissible. The framework also allows that interference that takes the form of coercive restriction of an agent's liberty for her own good might sometimes be permissible. The line between coercive shoves and gentler nudges that is important in the libertarian framework of Thaler and Sunstein does not show up as in itself normatively important in the framework here proposed. Paternalistic nudges may be permissible, and perhaps sometimes even morally required. Likewise for paternalistic shoves.<sup>24</sup>

*Department of Philosophy, University of California, San Diego*  
rarneson@ucsd.edu

---

<sup>24</sup>For instructive criticism of a draft of this essay, I am grateful to Kalle Grill, Danny Scoccia, and an anonymous referee for this journal.