

# **Kantian Perspectives**

*Melissa Seymour Fahmy, University of Georgia*

(forthcoming in the *Routledge Handbook on the Philosophy of Paternalism*)

## **1. Paternalism**

My aim in this chapter is to articulate the Kantian moral opposition to paternalism. Assessing the moral legitimacy of paternalism is complicated by the fact that there competing accounts of just what paternalism is. Some accounts of paternalism understand restricting another's freedom to be an essential component of paternalistic action (Kleinig 1984).<sup>1</sup> Others defend much broader accounts of paternalism arguing that being "liberty restrictive" is not a necessary feature of a paternalist act (Shiffrin 2000). Matters are less contentious when it comes to the paternalist's motive for acting. Most agree that in order for A's action toward B to count as paternalistic, A's motive must be concern for B's good. Of course, "one's good" can be understood in a variety of ways, and A's conception of B's good might be quite different from B's understanding of her own good. Finally, some contend that B's attitude matters as well, arguing that A's action counts as paternalist only if it is unwanted by B (De Marneffe 2006). Because I believe that Kantian opposition to paternalism is not limited to acts which are liberty restrictive, for the purposes of this chapter I shall adopt a slightly modified version of Seana Shiffrin's broad account of paternalism. Accordingly, an action by A toward B is paternalistic provided that (i) A intends her action to have an effect on B or on B's sphere of legitimate agency, (ii) A intends for her action to promote B's interests and this is A's sole or primary reason for acting and (iii) A's action involves the substitution of A's judgment or agency for B's "on the grounds that compared to B's judgment or agency regarding those interests, A regards her judgment or agency to be (or as likely to be), in some respect, superior to B's" (Shiffrin

2000: 218).<sup>2</sup> The following example is illustrative. B is a long-time smoker who desires to quit smoking, but is understandably finding it difficult to do so. A is a friend of B's and is aware of B's desire to quit smoking. If A elects to help her friend quit smoking by hiding, stealing, or destroying B's cigarettes without B's consent, then A acts paternalistically toward B.<sup>3</sup> Notably, A's behavior is paternalistic despite the fact that there is no fundamental disagreement between A and B regarding B's interests.

Ethicists have provided several distinct reasons for thinking that paternalism is, if not *necessarily* objectionable, at least *pro tanto* morally problematic. Some make the argument that paternalism is objectionable in virtue of the epistemic obstacles and limitations that even the best-intentioned paternalist faces. According to this line of thought, these epistemic obstacles make it unlikely that the paternalist will achieve her goal of promoting the good of another. An argument of this sort can be found in John Stuart Mill's *On Liberty*. There he contends that the strongest argument against public interference with purely self-regarding conduct is that the odds are great that the public will be swayed by its own preferences and judge poorly the good of others (Mill 1978: 81). Paternalism is objectionable in virtue of its likelihood to be self-defeating.

Others, like Joel Feinberg, argue that the objectionable feature of paternalism resides in the fact that liberty-limiting paternalistic interference constitutes an unjustifiable violation of the individual's sovereign right of self-determination. According to Feinberg, restricting individual liberty on strictly paternalist grounds is morally offensive "because it invades the realm of personal autonomy where each competent, responsible, adult human being should reign supreme" (Feinberg 1986: 25).

But perhaps the most common complaint made against paternalism is that such interference is *disrespectful*. As Shiffrin describes it,

Even when paternalist behavior does not violate a distinct, independent autonomy right, it still manifests an attitude of disrespect toward highly salient qualities of the autonomous agent. The essential motive behind a paternalist act evinces a failure to respect either the capacity of the agent to judge, the capacity of the agent to act, or the propriety of the agent's exerting control over a sphere that is legitimately her domain (Shiffrin 2000: 220).

Stephen Darwall similarly argues that the objectionable character of paternalism is “primarily a failure of respect, a failure to recognize the authority that persons have to demand, within certain limits, that they be allowed to make their own choices for themselves” (Darwall 2006: 268). And more recently, Michael Cholbi has identified the wrong-making feature of paternalism as “substituting others’ judgment concerning what a person’s good is for that person’s own judgment concerning her good, thereby failing to respect the individual as a locus of rational agency” (Cholbi 2013: 117). The respect objection has a broader reach than the previously considered objections insofar as it purports to give us a reason to reject paternalism even when paternalistic interference does not violate an autonomy right or restrict individual liberty and the paternalist in fact does know better how to secure the other’s good.<sup>4</sup>

In her recent book, *Against Autonomy*, Sarah Conly defends the use of coercive paternalism (e.g. legislation) to force people to act, or refrain from acting, according to their best interests. Conly’s argument is informed by research from behavioral economists and social psychologists, which concludes that human beings routinely exhibit certain cognitive biases and

reason poorly in predictable ways. The paternalism that Conly defends does not rely on an objective account of well-being or objectively valuable ends. Rather, the paternalism she defends is one that “helps people act according to their own values” (Conly 2013: 12). Conly argues for “intervention in cases where people’s choices of instrumental means are confused.” In such cases, action is constrained “only in order to get the person to do what he would want to do if he were fully informed and fully rational” (Conly 2013: 43).<sup>5</sup>

What intrigues me most about Conly’s defense of coercive paternalism is that it speaks directly to the respect objection. Conly suggests that paternalistic interference may be *more respectful* in virtue of the cognitive deficiencies that plague human rational agency. According to Conly, “Those who say we should respect autonomy by letting people hurt themselves irreparably do not, on my view, show as much respect for human value as they purport to” (Conly 2013: 1-2). I shall refer to this line of thought as *Conly’s Challenge*. We can frame Conly’s Challenge in the form of a question: Why think that interfering in the lives of others for the sake of promoting their welfare is *less respectful* than standing by and allowing individuals to make poor choices that undermine their well-being as they understand it?

Recent defenses of paternalism, like Conly’s, provide a welcomed opportunity for anti-paternalists to reconsider their opposition. As mentioned above, the value most often invoked in opposition to paternalism is respect, and this is certainly true for the standard Kantian opposition to paternalism.<sup>6</sup> We ought not treat adults in a paternalistic manner because, well, they are adults and such treatment would fail to provide them with the respect we are obliged to show them. By itself this claim is not particularly helpful or convincing, especially in light of Conly’s Challenge. Arguments defending the legitimacy of paternalism beckon us to step back and ask why. Why is

paternalistic interference disrespectful? What sort of respect are competent adults entitled to and how does paternalism fall short of this ideal?

My intention in this chapter is to explicate the Kantian grounds for the moral rejection of paternalism. I argue that a thorough account of the Kantian grounds for opposing paternalism involves three distinct yet related points. The first point concerns the Kantian moral grounds for acknowledging the decision-making authority of competent adults. In the following section I provide a Kantian answer the question *why think that paternalistic interference entails a failure to properly respect others?* The second point conjoins social and psychological facts about human moral development with a distinctly Kantian conception of the relative worth of particular ends. While happiness (well-being) and discretionary ends are by no means unimportant in Kant's theory, they are subordinate to the supreme good, virtue. In section three I argue that, even when we choose poorly, making our own decisions instrumentally serves our moral development. Insofar as paternalistic interference aims to promote the well-being of another (however this is understood) at the possible expense of undermining the development of her rational powers, the paternalist wrongly prioritizes well-being over self-development. The third and final point concerns the innate right to freedom, which is the foundation of Kant's political theory. Forms of paternalism that restrict an agent's freedom solely for the sake of promoting her welfare are doubly wrong: they are rights violations as well as failures of respect. What I aim to provide is an account of why Kantian normative theory is opposed to paternalism that goes deeper than the simple declaration that paternalism is a failure of respect. In the course of providing this account I aim to explain why a Kantian would not be persuaded by Conly's recent defense of paternalism. The Kantian grounds for rejecting paternalism reflect commitments to a particular conception of persons, as well as particular moral goods, such as

self-perfection, and the secondary importance of well-being. In this way, Kantian opposition to paternalism is distinct from a Millian or Feinbergian rejection of paternalism.

## **2. Kantian Opposition to Paternalism: The Value of Humanity**

My intention in this section is to explicate, from a Kantian perspective, the very common complaint that paternalism entails a failure of respect. Why think that paternalistic interference entails a failure of respect? Answering this question requires answering a pair of more fundamental questions about Kantian respect. First, what is the *source* of the Kantian duty to respect others? And second, *how* are we obligated to respect others? In other words, what does Kantian respect look like in practice? The answer to the first question informs the answer to the second and so we must begin there.

What is the *source* of the Kantian duty to respect others? In other words, what grounds this moral obligation? This question can be answered in a deceptively simple way. The source of the Kantian duty of respect is *humanity*. As Kant describes it,

Every human being has a legitimate claim to respect from his fellow human beings and is *in turn* bound to respect every other. Humanity itself is a dignity; for a human being cannot be used merely as a means by any human being (either by others or even by himself) but must always be used at the same time as an end (Kant 1996: 579; 6:462).<sup>7</sup>

*Humanity* is a deceptively simple answer to the question – *What is the source of the Kantian duty of respect* – because it matters very much what *humanity* refers to and yet Kant offers different descriptions of *humanity* throughout his work. For instance, in an early section of *The Doctrine*

of *Virtue*, Kant declares that “The capacity to set oneself an end – any end whatsoever – is what characterizes humanity (as distinguished from animality)” (Kant 1996: 522; 6:392). On the basis of this passage (as well as others) we might be tempted draw the conclusion that it is simply the capacity to set an end that entitles human beings to respect. But this would be a mistake. It is not simply our capacity to set an end that grounds our claim to respect, but our capacity for morality. Kant makes this explicit in a later passage in *The Doctrine of Virtue*:

In the system of nature, a human being (*homo phaenomenon, animal rationale*) is a being of slight importance and shares with the rest of the animals, as offspring of the earth, an ordinary value (*pretium vulgare*). Although a human being has, in his understanding, something more than they and can set himself ends, even this gives him only an *extrinsic* value of his usefulness (*pretium usus*)...But as a human being regarded as a *person*, that is, as the subject of morally practical reason, is exalted above any price; for as a person (*homo noumenon*) his is not to be valued merely as a means to the ends of others or even his own ends, but as an end in himself, that is, he possesses a *dignity* (an absolute inner worth) by which he exacts *respect* for himself from all other rational beings in the world...Humanity in his person is the object of the respect which he can demand from every other human being, but which he must also not forfeit (Kant 1996: 557; 6:434-5).

The view that human beings possess dignity in virtue of their capacity for morality and not merely their capacity to set an end is also expressed in the earlier *Groundwork*. There Kant contends that “morality is the condition under which alone a rational being can be an end in itself, since only through this is it possible to be a lawgiving member in the kingdom of ends.

Hence morality, and humanity insofar as it is capable of morality, is that which alone has dignity” (Kant 1996: 84; 4:435).<sup>8</sup> Note that it is our *capacity for morality* that has dignity, something distinct from our moral goodness.

To summarize: human beings - insofar as their rational nature is the source of the moral law and thus allows them to give the law to themselves - have dignity. To say that someone is a *person*, or that she possesses *dignity* is just to say that she ought always to be treated as an end in herself.<sup>9</sup>

Our answer to the first question has taken us some ways in answering the second question - *How* are we obligated to respect others? First and foremost, we are negatively obligated to avoid treating others *merely as means*. But we are also positively obligated to treat others as ends, that is, to treat them in ways that acknowledge their dignity (Kant 1996: 579; 6:462). Over the past several decades Kant scholars have devoted a great deal of thought and energy to the task of explicating what must be the case in order to draw the conclusion that someone has treated another *merely as a means* or has failed to treat her as *an end in herself*.<sup>10</sup> I cannot hope to do justice to either the subject or the literature here, yet something more needs to be said about these concepts if we are to arrive at an adequate appreciation of the Kantian grounds for rejecting paternalism.

Samuel Kerstein has recently proposed several sufficient conditions for treating another merely as a means. One of these sufficient conditions is the Actual Consent Account. According to the Actual Consent Account, an agent uses another merely as a means...

if it is reasonable for her to believe that either a.) the other was informed, before it occurred, of the agent’s intended use of him and that at the time voluntarily dissented from



it, or b.) the other was not or could not, before it occurred, be informed of her intended use of him. But if the other had been so informed, he would have voluntarily dissented from it, and c.) the other has not, prior to the agent's use of him, given his voluntary, informed consent to it or to a set of rules governing his and the agent's interaction, according to which her use of him is legitimate (Kerstein 2013: 118).<sup>11</sup>

It follows from the Actual Consent Account that many paradigmatic cases of wrongful interference – deception, coercion, manipulation, and force – are cases of using another merely as a means.<sup>12</sup> I find the Actual Consent Account of using another merely as a means to be quite plausible and, of the various accounts of using another merely as a means that Kerstein presents, the Actual Consent Account is the most relevant to the evaluation of paternalism.

We are now in a position to answer the question articulated at the beginning of this section: Why think that paternalistic interference entails a failure of respect? Paternalism is a failure of respect insofar as it entails treating a competent adult merely as a means to her own welfare, interests, or even her own ends. In order to see this more clearly, let us return to the example introduced at the beginning of the chapter. B is a long-time smoker who desires to quit smoking, but is finding it difficult to do so. B's friend A elects to help B quit smoking by hiding, stealing, and destroying B's cigarettes whenever she can do so without B's knowledge. A's behavior counts as paternalistic insofar as it (i) has an effect on B and B's sphere of legitimate agency, (ii) aims at promoting B's interests, and (iii) involves A substituting her judgment and agency for B's on the grounds that A regards her judgment and agency to be superior to B's in virtue of the fact that A does not suffer from nicotine addiction.

Features (i) and (iii), which allow us to identify A's action as paternalist, also make it the case that A's behavior likely satisfies the conditions for treating another merely as a means according to Kerstein's Actual Consent Account. In the example, A interferes with B's property without B's knowledge. Why does A not inform B of her plan to help B quit smoking? The most plausible explanation seems to be that A believes that if B knew what she was doing – hiding, stealing, and destroying B's cigarettes – B would object to this form of "help" and would attempt to thwart A's unwelcomed interference by not leaving cigarettes in plain sight. Thus it looks like A satisfies condition (b) of the Actual Consent Account. The second conjunct, (c), stipulates that in order for A to treat B merely as a means it must be the case that B "has not, prior to [A's] use of him, given his voluntary, informed consent to it or to a set of rules governing his and the agent's interaction, according to which her use of him is legitimate". A's behavior in the example also satisfies this condition. Things would be much different had A explicitly offered to hide, steal, or destroy B's cigarettes and B had accepted A's offer. A's actions would fail to satisfy the conditions for treating B merely as a means, and, notably, A's actions would also fail to satisfy the criteria for being paternalist. If B consents to A's way of acting, then A does not substitute her judgment or agency for B's. Thus there appears to be a very close connection between the third feature of our modified version of Shiffrin's account of paternalism and Kerstein's Actual Consent Account of treating another merely as a means.

Certainly, not every case of treating another merely as a means is a case of paternalism. In most cases, we use others merely as a means for selfish purposes and this might lead us to think that the selfishness is an essential wrong-making feature of our behavior. But this is erroneous. Selfish motivation is not a necessary condition for mere means treatment. We can be beneficently motivated and still treat another merely as means. This explains why even modest

paternalism – paternalist interference that does not entail substituting the paternalist’s perception of the other’s good for her own – will still count as disrespectful from the Kantian perspective.

As noted earlier, the coercive paternalism that Conly defends is modest. As she explains it,

The paternalism I promote here is not a paternalism about ultimate ends; that is, I do not argue that there are objectively good ends, or objectively rational ends, or ends objectively valuable in any way, which everyone should be made to pursue. I am arguing for intervention in cases where people’s choices of instrumental means are confused, in a way that means they will not achieve their ultimate ends (Conly 2013: 43).

The paternalist regulations that Conly defends “are designed to help us reach our own goals” and “constrain action only in order to get the person to do what he would want to do if he were fully informed and fully rational” (Conly 2013: 12, 43).<sup>13</sup>

But even this modest form of paternalism fails to address the core Kantian complaint. The core objection is not that paternalistic interference fails to respect the agent’s conception of her good, or her choice of ends, or even the her rational capacities. Rather, Kantian opposition to paternalism is grounded in the claim that the paternalist fails to respect the *status* of the other as an end in herself – and being an end in oneself is not dependent on being perfectly rational.<sup>14</sup> As Kant puts it, “a human being regarded as a person, that is, as the subject of morally practical reason, is exalted above any price; for as a person (*homo noumenon*) he is not to be valued merely as a means to the ends of others *or even his own ends*” (Kant 1996: 557; 6:434-5, emphasis mine). The authority to select not only one’s ends but also the means to these ends (provided neither is immoral) comes with the status of being an end in oneself. This is why

attempts to justify paternalism by appealing to our imperfect rationality will not be persuasive to the Kantian. For our claim to respect from others was never dependent on the supposition that we are perfectly rational beings. Thus, at least where competent adults are concerned, paternalistic interference remains objectionable from the Kantian perspective even when the paternalist knows better what the agent really wants and how best to secure it.<sup>15</sup>

Conly challenges the anti-paternalist by asking why it is more respectful to allow another to choose poorly rather than interfering when our interference will most likely yield a better result with regard to her own understanding of her interests. The Kantian answer to this question is that it is more respectful to acknowledge another's authority simply because she is *entitled* to this recognition in virtue of being a fully mature – albeit imperfectly rational - moral agent. Giving others the respect to which they are entitled does not preclude concern for their welfare. It merely precludes promoting the welfare of others in ways that attempt to circumvent their agency. This removes a few options from the table (e.g. manipulation and deception), but it hardly ties our hands.

### **3. Kantian Opposition to Paternalism: Self-Perfection & the Subordinate Value Well-Being (Happiness)**

The previous section located the principal wrong-making feature of paternalism, at least from a Kantian perspective, in what paternalistic interference *does*. Paternalistic interference is disrespectful insofar as it treats a competent agent as a mere means to her own interests, welfare, or ends. Some authors believe that paternalism can be made acceptable provided that the paternalist does not substitute her judgment of another's interests for the agent's own, that she acts to secure what the agent wants or would want if she were not such a flawed thinker. In the

previous section I explained why this qualification on paternalistic interference would not satisfy the fundamental Kantian complaint. In this section, I want to focus on the paternalist's *motivation* for acting, namely a beneficent interest in promoting another's good. My suspicion is that those, like Conly, who offer a defense of paternalistic interference, operate with a conception of another's interests that is not compatible with fundamental Kantian ideas about the subordinate value of individual happiness or well-being. This adds an additional layer to the Kantian rejection of paternalism and distinguishes it from others of its kind.

Despite the fact that Kant understands the human desire for happiness to be "a powerful counterweight to all commands of duty" (Kant 1996: 59; 4:405), he does not recommend that we attempt to purge ourselves of this desire. This would be futile. The desire for happiness is an inextricable part of the human condition. Kant could not be less ambiguous about this. In the *Critique of Practical Reason* he proclaims that "To be happy is necessarily the demand of every rational but finite being and therefore an unavoidable determining ground of its faculty of desire" (Kant 1996: 159; 5:25) and in a later text that "it is unavoidable for human nature to wish for and seek happiness" (Kant 1996: 519; 6:387). For better or worse, human beings are creatures that need happiness. This fact about our nature has implications for Kant's moral theory. Kant declares that others' happiness is an end that is also a duty and that one's own happiness might also be a duty if it weren't for the fact that it is self-contradictory to say that we are constrained to adopt an end we necessarily and unavoidably have (Kant 1996: 517; 6:385-6).

Kant clearly understands happiness to be important; however, it is not the most important end for a human being. The subordinate value of individual happiness is perhaps best observed in Kant's doctrine of the highest good. According to Kant, the highest good is a composite good made up of virtue and happiness. Kant is critical of the Stoics and the Epicureans for what he

perceives to be their failure to recognize that happiness and virtue are “extremely heterogeneous concepts” (Kant 1996: 229; 5:111). Virtue makes one worthy of happiness, but not necessarily happy. Kant explains,

Now inasmuch as virtue and happiness together constitute possession of the highest good in a person, and happiness distributed in exact proportion to morality (as the worth of a person and his worthiness to be happy) constitute the *highest good* of a possible world, the latter means the whole, the complete good, in which, however, virtue as the condition is always the supreme good, since it has no further condition above it, whereas happiness is something that, though always pleasant to the possessor of it, is not of itself absolutely and in all respects good but always presupposes morally lawful conduct as its condition (Kant 1996: 229; 5:110-111).

Virtue is unconditionally good, the *supreme* good, but it is not the *complete* good for human beings. Human beings need happiness, so happiness must be part of the highest good; however, its goodness is always conditioned by virtue.<sup>16</sup> The subordinate value of happiness is a theme that we find throughout Kant’s ethical writings. For instance, in the well-known opening of *Groundwork I* Kant declares that “It is impossible to think of anything at all in the world, or indeed even beyond it, that could be considered good without limitation except a good will...Power, riches, honor, even health and that complete well-being and satisfaction with one’s condition called happiness, produce boldness and thereby often arrogance as well unless a good will is present” (Kant 1996: 49; 4:393).

Thus we see that the relationship between happiness and virtue is complicated. Human beings need happiness and we unavoidably pursue our own happiness; however, in the absence of a good will, the value of happiness plummets. Our desire for happiness will at times tempt us to transgress the moral law, and this provides us with moral reason “attend to one’s happiness” (Kant 1996: 214; 5:93).<sup>17</sup> But the relationship between happiness and virtue is more complicated still. We get the Kantian picture wrong if we think that the desire for happiness is simply some unfortunate human liability that we should wish to be without. Barbara Herman has observed that there is a more symbiotic relationship between the pursuit of happiness and self-development. She explains,

Because effective agency is not like getting one’s adult teeth, it will not just happen with time and food, a moral theory that prizes the value of rational agency has to be especially sensitive to its social and material conditions as it goes about the business of parceling out goods...the vehicle that drives the development of human rational agency is the natural interest we have in our own happiness (Herman 2002: 241).

Our desire for happiness drives the development of human rational agency, in part, because happiness is such a nebulous end. We desire to achieve a state of lasting satisfaction with our condition, and yet we cannot know what would bring this about. Kant laments that “it is a misfortune that the concept of happiness is such an indeterminate concept that, although every human being wishes to attain this, he can still never say determinately and consistently with himself what he really wishes and wills” (Kant 1996: 70; 4:418).

The desire for happiness drives the development of rational agency because pursuing happiness requires working out a conception of happiness. Pursuing happiness also requires coming to terms with the fact that not all of our desires can be satisfied. We must deliberate about which desires to take up as ends and which to reject. When we are left to make our own decisions - decisions about ends as well as means - we experience the consequences of our good and poor choices. We learn from our mistakes. We develop judgment. Moreover, we learn that we can delay gratification and that we can say no to our inclinations. Self-mastery, if not the goal of the pursuit of happiness, is at least a desirable byproduct. Kant describes the vice of avarice as “restricting one’s own enjoyment of the means of good living so narrowly as to leave one’s own true needs unsatisfied” (Kant 1996: 555; 6:432). Avarice is not merely *imprudent*, according to Kant; it is a *vice* insofar as it entails a “slavish subjection of oneself to the goods that contribute to happiness” (Kant 1996: 557; 6:434). In short, the self-command we gain in pursuing happiness is instrumental to our self-perfection as moral agents.

The danger of paternalism is that it risks interfering with the development of self-command and self-perfection by not allowing agents to reap the benefits of making their own choices. Thus, even if the paternalist succeeds at promoting the welfare of another, even welfare as the agent herself understands it, the paternalist errs in prioritizing well-being over self-development.

Conly considers this objection in *Against Autonomy*. She observes that some argue “if we aren’t given the opportunity to make our own mistakes, we will fail to accrue the self-reliance and critical expertise that adverse experience – that is, making mistakes and suffering the ensuing results – gives us, and will be lesser people as a result” (Conly 2013: 27). Conly’s response to this objection is not very persuasive. On the one hand, she maintains that “in many



cases we learn too late”. Too late for what? Her examples suggest that we might learn too late to secure our interests. “When you’re diagnosed with lung cancer, it’s too late to learn that your cognitive bias towards anchoring is something you shouldn’t trust. When you’re forced to retire and find that you don’t have enough to live on, it’s too late to learn that your tendency toward irrational time discounting has really done you harm” (Conly 2013: 27). Fair enough.

Sometimes we learn too late to secure particular ends or even fundamental interests. But for the Kantian, these interests are not the supreme good. Even if learning from our mistakes comes late in life or at a high cost to our well-being, it may well be worth it if these mistakes make us better persons. The Kantian position is simply that well-being is neither the only thing that matters nor the most important thing.

More pessimistically, while Conly concedes that the objection has “some intuitive power, because of course we do sometimes learn from our mistakes,” she insists that “lots of people don’t learn from their mistakes” (Conly 2013: 27). It is no doubt true that many of us have made the same mistake an embarrassing number of times. But it does not follow that we *cannot* learn from our mistakes or come to have a better understanding of how to navigate our own cognitive biases. In addition to being disrespectful, bypassing another’s agency, as paternalism does, deprives the other of the opportunity to develop their rational agency.

Thus, while paternalistic interference may succeed in promoting the individual’s well-being, it typically does so at the cost of imposing an obstacle to self-development and self-command. From the Kantian perspective, this is simply not a good trade off.

#### **4. Kantian Opposition to Paternalism: The Innate Right to Freedom**

The final point regarding the Kantian grounds for opposing paternalism is relevant only to liberty restricting forms of paternalism. Depriving a competent adult of her freedom, even when one is beneficently motivated, is contrary to our juridical duty not to wrong others (Kant 1996: 392; 6:236). Kantian juridical duties, or duties of right, are grounded in what Kant believes is our sole innate right, the right to freedom understood as “independence from being constrained by another’s choice insofar as it can coexist with the freedom of every other in accordance with a universal law” (Kant 1996: 393; 6:237). Presumably, the only warranted restriction on individual freedom would be those restrictions which are necessary to protect the freedom of all citizens. Thus “liberty restrictive” paternalism would constitute a rights violation in addition to being an ethical failure to show others the respect to which they are entitled.

In his exposition of the duty of beneficence in the *Doctrine of Virtue* Kant raises the following casuistical question: “is not the injustice of depriving someone of his freedom something so contrary to duty of right as such that one who willingly consents to submit to this condition, counting on his master’s beneficence, commits the greatest rejection of his own humanity, and that the master’s utmost concern for him would really not be beneficence at all?” (Kant 1996: 572; 6:454). Though Kant does not provide an explicit answer to this question, ‘yes’ appears to be the most plausible Kantian response. Consenting to be paternalistically looked after is “the greatest rejection of [our] own humanity”. In this case, the benefactor’s conduct would not satisfy the conditions of Kerstein’s Actual Consent Account of treating another merely as means, in virtue of the benefactor’s having secured the other’s consent. But this leaves open the possibility, as Kant suggests, that consenting to this arrangement is tantamount to treating oneself merely as means.

## 5. Conclusion

It is frustrating to watch others engage in imprudent and self-destructive behaviors and enraging to observe entire industries devoted to encouraging this. It is understandably tempting to want to use the blunt tool that is paternalistic legislation to suppress such conduct. But it is not clear that this is even a pragmatic approach. Take one of Conly's favored examples, smoking. The educational campaign to highlight the health consequences of smoking has been largely successful at reducing the percentage of the population that smokes. "Self-reported adult smoking peaked in 1954 at 45%, and remained at 40% or more through the early 1970s, but has since gradually declined. The average rate of smoking across the decades fell from 40% in the 1970s to 32% in the 1980s, 26% in the 1990s, and 24% since 2000" (Saad 2008). In 2014 a mere 16.8% of American adults reported smoking (CDC 2015). Conly attributes smoking to cognitive biases "which we all share," but this explanation is undermined by the fact that there is a substantial correlation between smoking and income and education level. 43% of adults whose highest academic achievement is a GED smoke, and as educational achievement increases, the likelihood of being a smoker decreases dramatically (CDC 2015). If the cognitive biases explanation were correct, then we should expect to see no substantial difference in rates of smoking across demographics – but we do. The poor and less educated are more likely to smoke.

The cognitive biases explanation is too simplistic. We need to ask why individuals voluntarily engage in self-destructive behavior. Engaging with the rational agency of others will likely produce a better outcome than attempting to force others to conform to what we think they would want to do if they were "fully informed and fully rational". Finally, it is worth noting that when we engage with the agency of others, we avoid many of the epistemic challenges that

plague both modest and not-so-modest versions of paternalism, such as knowing what is truly in another's best interest, knowing another's ends, distinguishing another's "ultimate ends" from her merely sub-ultimate ends, and knowing how best to achieve particular ends. Thus, in addition to being more respectful, the Kantian approach may be more efficacious.

## Notes

---

<sup>1</sup> See Kleinig (1984). Joel Feinberg understands *legal* paternalism in this way. See Feinberg (1986).

<sup>2</sup> Shiffrin's own account is broader than the one I adopt here. Her account recognizes actions solely and directly concerned with a third party's welfare as paternalist, as well as some actions not motivated by welfare concerns at all. Shiffrin acknowledges that a full account of paternalism will depend on some account of an individual's "sphere of legitimate agency." However, I think we can proceed with merely a rough intuitive sense of what this sphere likely includes.

<sup>3</sup> This example comes from Shiffrin (2000), p. 215.

<sup>4</sup> Shiffrin provides examples of paternalism that are not violations of autonomy. See Shiffrin (2000), pp. 213-218.

<sup>5</sup> See also Conly (2013) p. 36.

<sup>6</sup> See Baron and Fahmy (2009), p. 213.

<sup>7</sup> All volume and page numbers refer to the Prussian Academy edition of Kant's *Gesammelte Schriften*.

<sup>8</sup> See also Kant (1996), p. 85 "Autonomy is therefore the ground of the dignity of human nature and of every rational nature" (4:436).

---

<sup>9</sup> See Sensen (2011) for a more thorough exposition of Kant’s use of the concept of dignity, as well as Kerstein (2013).

<sup>10</sup> See Hill (1980, 2003), O’Neill (1985), Korsgaard (1986), Kerstein (2011, 2013), and Parfit (2011).

<sup>11</sup> According to Kerstein, “an agent uses another (or, equivalently, uses another as a means) if and only if she intentionally does something to or with (some aspect of) the other in order to realize her end, and she intends the presence or participation of (some aspect of) the other to contribute to the end’s realization” (Kerstein 2013: 58).

<sup>12</sup> Kerstein makes the cases that using another merely as a means is *pro tanto* wrong rather than always or necessarily wrong.

<sup>13</sup> Conly defends coercive paternalism by appealing to cognitive biases which undermine our ability to select appropriate means to achieve our goals. However, the same research suggests that we are just as bad at selecting ends as we are at selecting means. Thus Conly’s defense of a modest paternalism appears vulnerable to the charge of arbitrariness.

<sup>14</sup> I maintain that there is a fundamental difference between respecting someone’s capacities or judgment and respecting her authority or propriety even if authority is dependent on the possession of particular capacities.

<sup>15</sup> Michael Cholbi has recently defended a position he calls *Kantian Paternalism* (KP). According to KP, interference with another is justified when “we stand a reasonable chance of preventing that agent from performing actions she chose due to *distorted reasoning* and which would result in that agent’s rationally chosen ends not being as fully realized as they would have been had she so acted.” Cholbi understands “distorted reasoning” to mean reasoning that is

---

influenced by mental disorders and/or cognitive defects. See Cholbi (2013). I take Cholbi's position to be consistent with my own account of the Kantian opposition to paternalism.

<sup>16</sup> Note that Kant regards a *product of autonomy* (virtue) as the supreme or unconditioned good, not autonomy itself.

<sup>17</sup> See also Kant (1996) pp. 519-520; 6:388.

## References

- Baron, M. and Fahmy, M. S. (2009) "Beneficence and Other Duties of Love in The Metaphysics of Morals, in T. Hill (ed.), *The Blackwell Guide to Kant's Ethics*, West Sussex: Wiley-Blackwell.
- Centers for Disease Control and Prevention (2015) "Current Cigarette Smoking Among Adults – United States 2005-2014," *Morbidity and Mortality Weekly Report* 64, pp. 1233-1240.
- Cholbi, M. (2013) "Kantian Paternalism and Suicide Intervention," in C. Coons and M. Weber (eds.), *Paternalism: Theory and Practice*, Cambridge: Cambridge University Press, pp. 115-133.
- Conly, S. (2013) *Against Autonomy: Justifying Coercive Paternalism*, Cambridge: Cambridge University Press.
- Darwall, S. (2006) "The Value of Autonomy and Autonomy of the Will," *Ethics* 116, pp. 263-284.
- De Marneffe, P. (2006) "Avoiding Paternalism," *Philosophy and Public Affairs* 34, pp. 68-94.
- Feinberg, J. (1986) *Harm to Self: The Moral Limits of Criminal Law*, Oxford: Oxford University Press.
- Herman, B. (2002) "The Scope of Moral Requirement," *Philosophy and Public Affairs* 30, pp. 227-256.
- Hill, T. E., Jr. (1980) "Humanity as an End in Itself," *Ethics* 91, pp. 84-90.
- (2003) "Treating Criminals as Ends in Themselves," *Jarbuch für Recht und Ethik* 11, pp. 17-36.
- Kant, I. (1996) *Practical Philosophy*, ed. and trans. M. Gregor, Cambridge: Cambridge University Press.
- Kerstein, S. (2011) "Treating Consenting Adults as Mere Means," in M. Timmons (ed.), *Oxford Studies in Normative Ethics* vol. 1, Oxford: Oxford University Press.

- (2013) *How to Treat Persons*, Oxford: Oxford University Press.
- Kleinig, J. (1984) *Paternalism*, Totowa, N.J.: Rowan & Allenheld.
- Korsgaard, C. (1986) "The Right to Lie: Kant on Dealing with Evil," *Philosophy and Public Affairs* 15, pp. 325-349.
- Mill, J.S. (1978) *On Liberty*, Indianapolis: Hackett.
- O'Neill, O. (1985) "Between Consenting Adults," *Philosophy and Public Affairs* 14, pp. 252-277.
- Parfit, D. (2011) *On What Matters*, Oxford: Oxford University Press.
- Saad, Lydia (2008) "U.S. Smoking Rate Still Coming Down," [www.gallop.com/poll/109048/us-smoking-rate-still-coming-down.aspx](http://www.gallop.com/poll/109048/us-smoking-rate-still-coming-down.aspx) (accessed 02/01/2016).
- Sensen, O. (2011) *Kant and Human Dignity*, Berlin: Walter de Gruyter.
- Shiffrin, S. V. (2000) "Paternalism, Unconscionability Doctrine, and Accommodation," *Philosophy and Public Affairs* 29, pp. 205-250.