

Antipaternalism as a Filter on Reasons

Kalle Grill

Abstract [not included in printed chapter]: I first distinguish four types of objection to paternalism and argue that only one – the principled objection – amounts to a substantive and distinct normative doctrine. I then argue that this doctrine should be understood as preventing certain facts from playing the role of reasons they would otherwise play. I explain how this filter approach makes antipaternalism independent of several philosophical controversies: On the role reasons play, on what reasons there are, and on how reasons are related to values. I go on to contrast the filter approach with the competing and dominant action-focused approach, which understands objections to paternalism in terms of paternalistic action, behavior, law, policy and the like. Seana Shiffrin and Peter de Marneffe are singled out as prominent recent proponents of this approach. By engaging with their definitions of paternalism, I explain how the action-focused approach makes antipaternalism dependent on the sorting of actions into paternalistic and nonpaternalistic according to what reasons support them. Because one and the same action can be supported by many different reasons, and by different sorts of reasons, such sorting is very difficult. The upshot is that antipaternalism on the action-focused account fails to provide the precise normative implications of the filter approach that I favor.

4.1 Introduction

The charge of paternalism is a common objection to the actions of political and other authorities. Sometimes the charge is only that the authority has undervalued typical liberal values like freedom and autonomy relative to other values, such as physical, mental or financial wellbeing. Making this objection is consistent with accepting that in some cases, wellbeing outweighs freedom and autonomy and should be furthered at their expense. Other times, however, the charge of paternalism is more principled. The objection is not that wellbeing considerations are overstated, but that they are allowed to weigh in on the matter at all. This is the sort of antipaternalism that I will analyze in this article. My discussion and my proposals are meant to be helpful to the antipaternalist, and to anyone who wants to understand her. However, I should state at the outset that the antipaternalist position I describe and develop is not one I endorse.

I propose that principled antipaternalism entails that certain facts are prevented from playing the role of reasons they would otherwise play. Using an obvious metaphor, I call this the *filter approach* to antipaternalism: The potential reasons provided by some facts are filtered out and so do not play the role of reasons. Exactly which these facts are determines the precise normative content of antipaternalism. I take no stand on this issue.

My thesis is a conceptual thesis on the structure of antipaternalism and not a normative position.

The main competitor to the filter approach is the *action-focused approach* according to which antipaternalism entails that certain actions or policies are paternalistic and therefore impermissible or otherwise morally problematic. The action-focused approach is the dominant approach in academic discussions of paternalism [p. 48] and antipaternalism. A significant problem with this approach is that it requires the sorting of actions into paternalistic and nonpaternalistic. It has proven quite difficult to define the criteria for such sorting. On the filter approach, actions need not be sorted in this way: Antipaternalism simply demands that some reasons are filtered out, without categorizing actions as either paternalistic or not. The filter approach therefore provides more straightforward normative implications, and, arguably, just the sort of normative implications antipaternalists typically intend.

The filter approach is independent of more fundamental questions concerning what is valuable and what we have reason to do. The filtering can be motivated in several different ways, including by practical concerns with the ability of certain agents (typically government agents) to properly consider certain reasons, by more normative concerns with the appropriateness of these agents considering these reasons, independently of effects, and by stronger normative principles according to which certain facts simply do not provide reasons in certain situations.

Both the action-focused and the filter approach are approaches to *antipaternalism*. However, the action-focused approach is dependent on a definition of paternalistic action, or in other words of *paternalism*. I will therefore also discuss some proposed definitions of this related concept. However, I defend no thesis on the best understanding of paternalism.

This article is outlined as follows: In section 1, I specify the principled objection to paternalism that I am concerned with and contrast it with other objections. In section 2, I present the filter approach in further detail. In section 3, I contrast the filter approach with the dominant action-focused approach.

4.2 Antipaternalisms

I mentioned two variations on the paternalism charge in the first paragraph of the introduction, but in fact there are several. In this section, I will survey four and explain why my interest here is only with one of them. I will, for ease of presentation, throughout use “liberty” as a placeholder for various liberal values like freedom, autonomy, self-determination etc., and “limiting liberty” as a placeholder for the infringement, diminishing, disrespect etc. of some such value. I will also make a controversial assumption: That antipaternalism only applies to cases where the paternalist promotes the good of the very same person whose liberty she is limiting (or aims to do so). Seana Shiffrin rather famously rejects this assumption, arguing that it is paternalism (and so objectionable) to limit the liberty of some person also to make things better more generally, as long as the infringement concerns this person's sphere of legitimate control (2000, p. 216). I make my controversial assumption only for convenience. It will enable me to speak of the relevant reasons in a straightforward manner and thereby much

simplify my presentation. However, my analysis could be reformulated to accommodate Shiffrin's wider take on paternalism (given that her position can be made sufficiently clear). [p. 49]

First of the four variations of the paternalism charge is the one that I will be exclusively concerned with after this present section: The principled objection that consideration of a person's good should not even count in favor of limiting her liberty. I propose that this objection is a plausible interpretation of Mill's liberty principle as well as Feinberg's antipaternalism (or "soft paternalism"). While Feinberg's account is perhaps unique in its detailed specification, the objection figures frequently, though more or less explicitly, both in philosophical accounts of paternalism and in applied ethical contexts and in the public debate. The objection is typically formulated as a rejection of the balancing or weighing of different considerations, in favor of the priority of liberty over other concerns.¹

The second variation, mentioned in the introduction, is the common sense objection that considerations of a person's good are sometimes given too much weight in relation to respect for her liberty. In contrast to the first objection, this objection explicitly recognizes that liberty can be balanced against other good things for a person. The objection applies whenever some agent gets the balance wrong and so liberty is limited too lightly. Though determining the relative importance of liberty may be a very complex matter, the structure of the objection is straightforward – liberty is undervalued. The objection does not represent a normative principle or position distinct from the common-sense position that one should give each value its due.

Sometimes agents are liable to systematically undervalue liberty, for example because of overconfidence in their ability to force people to improve their own lives. There are arguably many different values or goods for individuals and so many ways to get the balance between them wrong. When such mistakes are systematic, however, there may be reason to make them more salient by giving them names, such as "paternalism". There may even be reason to give the opposition to such mistakes a name, such as "antipaternalism". However, this name should not then be taken to refer to anything beyond this opposition.

The third variation of the paternalism charge, not mentioned in the introduction, is the objection to one person taking action towards another based on the first person's view of what is good or best for the second person. Mill famously argued that each person best knows her own interests, which would seem to tell against acting on one's ideas of what is good for others, whether or not liberty is at stake (1859, [p. 50] chapter

¹ Feinberg claims, for example, that "personal autonomy [...] is a moral trump card, not to be merely balanced with considerations of harm diminution in cases of conflict, but always and necessarily taking moral precedence over those considerations." (1986, p. 26) For a more recent example, Daniel Groll proposes that when it comes to benefitting a person, her will should be treated as "structurally decisive in determining what to do – it is meant to supplant the reason-giving force of other considerations not because it outweighs those other considerations but because it is meant to silence, or exclude, those other considerations from the practical deliberations of the subject of the demand" (2012, p. 701).

I should perhaps recognize that John Rawls is well known for his argument for the priority of liberty within the special context of distributive justice. Rawls, however, says very little about paternalism. He accepts paternalism since people's "capacity to act rationally for their good may fail, or be lacking altogether." (1999, p. 219) He does not discuss other, more problematic cases.

IV, 4th paragraph). Immanuel Kant in a characteristically more principled manner stated that when it is our duty to promote others' happiness "[i]t is for them to decide what they count as belonging to their happiness" (1991/1797, 388). Dan Brock (1988) considers this objection central to the very concept of paternalism:

[P]aternalistic interference involves the claim of one person to know better what is good for another person than that other person him- or herself does. It involves the substitution by the paternalistic interferer of his or her conception of what is good for another for that other's own conception of his or her good. (p. 559)

Gerald Dworkin (1983) too speaks of substitution in defining paternalism, but of "judgment" rather than of a conception of what is good. While it is clear in the context that Brock's "conception of what is good" refers to ideals about the good life, Dworkin's "judgment" seems to refer either to such judgments about ideals, or to more mundane judgments of how best to realize these ideals. These two possibilities are perhaps captured by his later distinction between strong and weak paternalism (2010).

When we substitute our judgment for someone else's, whether concerning what her good is or the best means of promoting it, this can lead us to limit their liberty. If it does, the principled objection applies and the common sense objection may apply depending on the balance of values. If our substitution of judgment does not lead us to limit someone's liberty, however, I do not see that there can be much of an objection. As Kant went on to say, "it is open to me to refuse them many things that *they* think will make them happy but that I do not, as long as they have no right to demand them from me" (1991/1797, 388). As Paul Guyer has argued, we should not think, nor take Kant to mean, that it is morally problematic that we make our own judgments about the best interests of other people (2014, 231-232).

For an illustration, assume that you judge that I am being overly friendly with you and should be less friendly, though I in no way impose on you. Perhaps you just believe that friendliness is a sign of weakness and therefore to be avoided. Suppose I know that this is your judgment but decide anyway to remain as friendly as ever, and that I do so partly or solely in order to get you to warm up and break your social isolation. This does not seem morally problematic and not the sort of thing liberals oppose. To the contrary, it seems a liberal ideal that I have a moral right to be as friendly as I choose, for whatever reason I find compelling, as long as I do not impose on others. Therefore, the third variation of the paternalism charge is not an independent objection. Only when a substitution of judgment leads to a limitation of liberty is it problematic, and so the problem lies with the limitation of liberty rather than with the substitution of judgment.

The fourth variation of the paternalism charge, also not mentioned in the introduction, is the objection to treating people in a condescending manner. This objection is often what people have in mind when they speak of "treating adults as if they were children" and the like (e.g. Szasz 1992, xiv, arguing against drug criminalization). Condescendence is an important part of Shiffrin's rich account of paternalism. Shiffrin claims that paternalism is always undertaken on the basis of a sort of disrespect toward the paternalized person's judgment or agency. This disrespectful [p. 51] attitude "is

central to accounting for why paternalism delivers a special sort of insult to competent, autonomous agents” (p. 220).

It is not quite clear what Shiffrin means by an attitude in this context. Sometimes it seems that she means that paternalist behavior, in virtue of its interfering nature, essentially embodies a condescending or disrespectful attitude, regardless of the actual psychological state of the paternalist (who could for example reasonably believe that the target is incompetent). So for example she states that paternalistic behavior “manifests an attitude of disrespect” (ibid.). This line of thought, however, does not amount to an independent objection. Rather, it is an argument for the principled objection: We should not count a person's good as a reason for limiting her liberty, because doing so embodies a condescending attitude towards her.

Other times, however, it seems that Shiffrin means that an actual disrespectful attitude is essential to paternalist behavior. For example, she claims that “the paternalist's attitude shows significant disrespect” (ibid.), as if there could be an otherwise similar interferer whose attitude did not show disrespect and who therefore was not a paternalist. It is this interpretation that amounts to an independent variation of the paternalism charge - an objection to having, or acting on, improper attitudes.²

To consider this objection, note first that displaying a condescending or disrespectful attitude is not particular to paternalism. People can be condescending in all sorts of situations, many of which are clearly not paternalistic. For example, we may ridicule an unsuccessful competitor, or reject the advice of some well-meaning acquaintance as useless without hearing it. This shows that we do not need the concept of paternalism, or antipaternalism, to explain the possible moral problems involved in being condescending and disrespectful, i.e. treating people as if they were less competent than they are.

More to the point, many liberals feel that benevolent limitations of liberty are morally problematic whether or not they are accompanied with a condescending attitude. Benevolent *and condescending* limitations of liberty may be especially problematic, but so are benevolent *and unfair* limitations of liberty, and so on. Condescendence is a separate problem; what is particular to paternalism is the limitation of liberty.

To sum up, the charge of paternalism can, on closer inspection, amount to one of several different objections, or indeed to more than one. However, the objection from substitution of judgment and one interpretation of the objection from condescendence both presuppose an objection from the limitation of liberty. Another [p. 52] interpretation of the objection from condescendence does not capture the heart of the matter but is a distinct concern with attitudes. Only the two objections from the limitation of liberty are independent and hit the right target. These two objections presuppose different views on

² Other authors with a similar take on paternalism are no clearer than Shiffrin in this regard. For example, Jonathan Quong argues that the essence of paternalism is that action is “motivated by a negative judgment about the ability of others to run their own lives.” (2010, p. 74). Paternalism is wrong, Quong claims, “because of the way it denies someone’s moral status as a free and equal citizen.” (Ibid.) It is not clear whether Quong thinks that being motivated by a negative judgment is a mental state distinct from simply judging oneself to have superior knowledge or ability in the particular case, or whether this (possibly true) judgment of superiority, regardless of other mental states, embodies a “negative judgment” that denies others' their moral status. Only the first understanding could be the basis of an independent objection.

the role of liberty. The common sense view assumes liberty to be one value among others, to be weighed against them. This view does not constitute a distinct normative view but is simply the view that each value should be given its due. In contrast, the principled view does constitute a distinct normative view, or rather a family of views with a shared structure. It is this structure that I will now go on to analyze.

4.3. Filtering reasons

My thesis is that principled antipaternalism is best understood as a filter that prevents certain facts from playing the role of reasons (for certain actions). Given my designation of ‘limiting liberty’ as a placeholder for various allegedly illiberal impositions, and given my assumption that the relevant rationales have to do with the good of the person imposed upon, this can be put in somewhat more specific terms: The antipaternalist filter prevents the fact that the limitation of some person's liberty causes the promotion or protection of her good from playing the role of a reason for such limitation. With due care, this approach can also be applied to many-person cases, and so to public policy.³

It may seem an extreme position to hold that facts that play the role of reasons in other contexts are entirely blocked from doing so. A more moderate position would be to admit that they play the role of reasons, only with reduced strength. I believe the typical antipaternalist position is that reasons are filtered out entirely and so this is the view I will discuss. However, the filter approach is entirely consistent with the filter being only partial, perhaps reducing the strength of reasons according to some complicated formula.

What I provide here is a generic or structural account of antipaternalism. Because it is quite general, I call it an “approach” and reserve the term “account” for more specific characterizations. The approach is non-committal in several respects. For starters, it is an open question whether the filtered-out facts are prevented from playing the role of reasons because they are not reasons, or because they are reasons but nevertheless should not play that role in the particular context specified by the details of the doctrine. It is also, relatedly, an open question what role exactly reasons normally play and so what role the filtered-out reasons are prevented from [p. 53] playing. It is precisely in order to keep the approach non-committal in these ways that I use the phrase “play the role of reasons”.

Without going into any detail, here is a simple map of the roles reasons can play: Reasons can play a role either in reasoning processes or in objective determinations of normative status. The latter include determining the permissibility of actions, the legitimacy of policies and the standing and authority of agents. Some believe these normative statuses are independent of anyone’s reasoning about them, yet are dependent on (objective) reasons. Reasoning processes, in turn, include both more volitional processes of forming intentions or motives, whether individually or collectively, and more evaluative processes that aim for knowledge concerning the normative statuses just

³ The formula cannot be applied directly. Strictly speaking, it is not correct to say that the filter prevents the fact that the limitation of *several* persons' liberty causes the promotion or protection of *their* good from playing the role of a reason for such limitation. This is because an action that limits the liberty of several people may promote the good of each one only or partly by limiting the liberty of the others. Antipaternalism does not apply to such liberty-limitation. For a thorough treatment of many-people cases, see Grill 2007, pp. 453-455.

described, including the permissibility of actions and policies, whether one's own or those of others, and whether future, present or past. The antipaternalist filter can block all of these roles, or some of them.⁴

To see that the basic notion of a filter on reasons is quite straightforward, note that having some reason to prefer some alternative does not imply that one should consider this reason in one's practical deliberation. There are often obvious pragmatic reasons for why one should not, including lack of time and ability, as well as social coordination resulting in the distribution of deliberative tasks.

If antipaternalism is specified towards practical deliberation and if its normative underpinning is pragmatic, the doctrine may be simply a useful guide to deliberation, a way of avoiding mistake, and wasted time and effort. A bit more interestingly, antipaternalism directed at some sort of reasoning can have a more thoroughly consequentialist underpinning, being based on a combination of pragmatic considerations regarding this kind of reasoning and a commitment to the furthering of some value or values. Even more interestingly, antipaternalism can be a deontological principle, directly preventing facts from playing the role of reasons, whether in reasoning or more objectively, for nonconsequentialist moral reasons. Whatever the normative underpinning, antipaternalism adds to (other) pragmatic guides to deliberation a further constraint – a recommendation or requirement that certain facts not play the role of reasons.

I noted that antipaternalism on the filter approach is non-committal regarding whether or not the facts that do not play the role of reasons are reasons at all. This means that the doctrine does not presuppose any particular view on what kinds of reasons there are. Nor, of course, does it presuppose any particular view on how reasons are related to values. This flexibility means that antipaternalism is neutral concerning what is the correct moral theory on this fundamental level. It is a module that can be included in various moral and political theories.

That the filter approach avoids integration with controversial views on what has value, what reasons there are, and what role reasons play, is a great advantage. Antipaternalist doctrines can be and are in fact supported on very different grounds, [p. 54] and are usually intended not to depend on controversial positions on the nature of reasons, of justification, or of reasoning. When antipaternalism is separated out from other commitments, it can more easily be appraised.

The filter approach will hopefully strike many as quite intuitive, perhaps even obvious. It is, however, in opposition to almost all conceptual treatments of paternalism and antipaternalism currently on offer.

Joel Feinberg may at points seem to tend towards a filter approach, but in fact he never formulates it as an alternative. Feinberg's commitment to the action-focused approach is obscured by his ten liberty-limiting principles. These are formulated in terms of which reasons they sanction as good reasons, rather than in terms of which actions they sanction. Indeed, Feinberg defines his liberal position as the position that only two

⁴ To the extent that a filter account prevents the consideration of certain facts, it is in a sense focused on action, namely on the mental action of considering facts. This does not, however, amount to an overlap with the action-focused approach, which is concerned with physical actions (and omissions), policies, laws etc.

of these principles – the harm principle and the offence principle – are valid principles (e.g. 1984, pp. 14-15).⁵ This directly implies a sort of filter on reasons, as Feinberg explicitly concludes: “Paternalistic and moralistic considerations, when introduced as support for penal legislation, have no weight at all.”⁶ (Ibid., p. 15) However, Feinberg does not stop at this point, but takes this conclusion to be but one step in his further argument. Rather than taking the liberty-limiting principles to operate directly on reasons, Feinberg takes them to determine which prohibitions are and which are not paternalistic. Later authors have followed Feinberg both in emphasizing the crucial role of reasons, and in nevertheless assuming that antipaternalism must target actions, policies or laws.

Douglas Husak has provided one of very few exceptions to the dominance of the action-focused approach.⁷ Husak offers a filter account of antipaternalism in his proposal that

a theory about the conditions under which paternalism is justified [...] might constrain the set of considerations to which legislators are allowed to appeal in their deliberations about whether to support or oppose a given piece of legislation.⁸ (2003, p. 391-2) [p. 55]

Husak goes on to more or less reject the action-focused approach, proposing that while reasons can be paternalistic, laws cannot. However, Husak then notes, correctly, that philosophers often talk of paternalistic laws. Apparently contradicting his own proposal, he claims that this talk is neither confused, nor a mistake (p. 390).⁹

I propose that it is indeed a mistake to suppose that the objection to paternalism can be fruitfully interpreted in terms of the rejection of certain laws or actions. I will argue this point at length in the following section.

⁵ For Feinberg, the harm principle is the principle that preventing harm to other persons than the actor is always a good reason for prohibition given that there are no better ways of preventing this harm.

⁶ Towards the end of the four volumes, Feinberg retracts this claim and in fact states to the contrary that these considerations are “always relevant” (1990, p. 322). This could be understood as a complete abandonment of principled antipaternalism, along with principled antimoralism. However, Feinberg insists that the retraction is not that consequential. He still thinks that the considerations are “hardly ever” good reasons and “perhaps never” decisive (p. 323). Especially in the case of paternalism, Feinberg reaffirms his earlier stance that liberty, in the form of personal sovereignty, is a “trump” that “cannot be put on the interest-balancing scales at all” (p. 322). Tenable or not, his position, even here, is that there is a principled difference between different kinds of reasons, though perhaps a principle with some exceptions.

⁷ One earlier example is C.L. Ten (1980), e.g. p. 40: “There are certain reasons for intervention in the conduct of individuals which must always be ruled out as irrelevant”.

⁸ Husak refers to Waldron’s “Legislation and moral neutrality” (1989), where Waldron proposes a very similar interpretation of neutralism in politics. The debate on neutrality is in general more explicit and more consistent regarding the role of reasons than is the paternalism debate. Still, the filter approach may be useful for interpreting neutrality too. Rawls’ idea of public reason, restricted in content, is a sort of filtering device (e.g. 1997, p. 776). Authors on neutrality such as Larmore (1987), de Marneffe (2010) and most explicitly Wall (1998) talk of neutrality as a constraint or restraint on some sort of reasons.

⁹ Husak recapitulates these points in a more recent contribution on penal paternalism (2013, p. 40-41).

I have myself previously (Grill 2007) argued that paternalism is essentially about action-reason compounds and so that neither actions nor laws can be paternalistic. I have also previously (Grill 2010) argued that antipaternalism is committed to what I then called the invalidation of reasons, which is a form of normatively based filtering. In relation to my earlier treatments, the filter approach is more general and more clearly positioned in relation to general moral theory and practical reasoning.

4.4 Filter vs. action-focus

Principled antipaternalism is the doctrine that a person's good should not count in favor of limiting her liberty. The standard interpretation of this doctrine is action-focused, taking it to be an objection to actions (including government or organizational actions) that limit some person's liberty and that are supported by reasons provided by the protection or promotion of this person's good.¹⁰ Paternalism is typically defined as the performance of such actions. Sometimes it is further assumed or argued that antipaternalism only targets actions that are motivated by a view of some person's good that differs from her own view, or that it only targets actions performed with a condescending attitude. However, I argued against these understandings of antipaternalism above and now disregard them.

The normative debate on the limits of benevolent limitation of liberty is intertwined with a parallel debate on the concept of paternalism, which is partly independent of normative concerns. This wider conceptual paternalism debate is peculiar because, on the one hand, it largely relies on linguistic intuitions about what cases are properly called paternalism, while, on the other hand, it seems to engage authors because they are interested in normative issues to do with antipaternalism. This is peculiar because it seems obvious that the normatively most plausible version of antipaternalism need not target the linguistically most accurate characterization of paternalism. [p. 56]

Having acknowledged these reservations about the relevance of the wider conceptual debate, I will in the remainder of this section consider and criticize definitions of paternalism in terms of actions, first in general and then in the form of two accounts that are relatively recent and relatively plausible – those of Seana Shiffrin (2000) and Peter de Marneffe (2006). Since the action-focused approach to antipaternalism is dependent on the identification of paternalistic actions, a critique of the project of defining paternalism in terms of actions is also a critique of the action-focused approach to antipaternalism.

On the filter approach, antipaternalism is normatively straightforward. I observed above that reasons may either objectively determine normative status or may figure in reasoning processes. If facts are prevented from playing a role in objectively determining normative status, they simply do not weigh in on the matter, they do not affect the balance of reasons. If facts are instead prevented from playing a role in some reasoning process, antipaternalism is directly action-guiding in prohibiting certain well-defined

¹⁰ Important adherents to the action-focused approach include Dworkin 1972; Gert & Culver 1976; Arneson 1980; Kleinig 1983; VanDeVeer 1986; Archard 1990; Shiffrin 2000; De Marneffe 2006; Dworkin 2010. All of these authors focus their conceptual concern on paternalism rather than on antipaternalism, but their contributions are motivated by the observation that paternalism is, or is allegedly, morally wrong.

(mental) actions. The exact scope of the doctrine determines to what status determinations or what reasoning processes it applies exactly, but whatever the scope it is quite clear what it means to abide by the doctrine when it does apply.

For example, if the doctrine applies to reasoning about public policy, then the fact that some policy furthers the good of a person by limiting her liberty does not play the role of a reason for that policy. It is either a practical mistake or a moral failure to be persuaded by such facts to, for example, vote for the policy or to enact it or to abstain from revoking it. If the doctrine applies to reasoning by physicians about the treatment of their patients, then the fact that some treatment furthers the good of a patient by limiting her liberty (e.g. because it is coercive or manipulative) does not play the role of a reason for that treatment. It is either a practical mistake or a moral failure to be persuaded by such facts to, for example, provide the treatment or urge colleagues to provide it.¹¹

In contrast, the normative implications of a policy's being paternalistic on the action-focused approach are quite unclear. The action-focused antipaternalist must explain how we should respond to the fact that an action or policy is paternalistic. The most typical explanation is probably that paternalistic actions should not be performed, and policies not enacted. This, however, leaves many questions open, [p. 57] such as whether or not paternalistic actions and policies should be prevented by third parties, and if they should be revoked once enacted. To some extent, these problems are shared by any rule formulated in terms of actions. That we should not lie and not murder does not entail that we should prevent lies and murders (and it could plausibly be argued that we should prevent murders but not necessarily lies). However, that it is shared does not make the problem less acute.¹² Furthermore, the problem is arguably greater for antipaternalism than for many other moral prohibitions because there is no obvious net harm to anyone from paternalism. Indeed, there is a net benefit, or at least an intended net benefit. This makes it less obvious how to respond to the categorization of some action as paternalistic. Moreover, as I will now go on to discuss, the concept of paternalism is particularly tangled up in reasons, to the extent that it is less a type of action than the combination of a type of action with a type of rationale.

The dominant approach to defining paternalism is quite preoccupied with reasons. Actions and policies are deemed paternalistic in large part depending on what

¹¹ Like most other moral doctrines, antipaternalism, on the filter approach, has no obvious implications for how transgressions should be evaluated. It may be tempting to call "paternalism" failures to abide by antipaternalism directed at deliberation. However, such failures can be trivial in the sense that they do not affect the outcome of deliberation. It is not clear to me whether such trivial failures should be called "paternalism" or what would be gained by doing so.

It may seem extreme to hold that furthering people's good should not even be an operative reason to limit their liberty, in public policy and perhaps especially in the context of medical care. However, it should be noted that the *preferences* of citizens, residents or patients can still provide operative reasons. The principally antipaternalist physician would presumably treat patients in accordance with their preferences and not in accordance with their best interest (and if these are identical, her operative reasons would still be based on patient preference as preference, not as constituent of the good).

¹² Amartya Sen attempts to deal with this problem by emphasizing our imperfect duty to aid (e.g. 2012, p. 96, but also in 2009, chapter 17). As noted by Frances Kamm, however, Sen's duty to aid "seems merely to ask us to think and act appropriately about important matters" (2011, p. 94). This is not very helpful.

reasons there are for them.¹³ There are two main obstacles to specifying this condition on paternalistic actions: 1) There are different sorts of reasons, including motivational and justificatory. 2) Actions most often have mixed or multiple rationales.¹⁴ Both problems are particularly acute in the political realm, where rationales are more thoroughly considered and are often collective, or are aggregates of many individual rationales, and so are quite diverse. While both the filter and the action-focused approach must identify which *reasons* antipaternalism applies to, the action-focused approach must in addition specify for any mix of reasons for an action whether or not this mix makes the *action* paternalistic. It is the necessity of this further conceptual work that is the main weakness of the action-focused approach.

De Marneffe (2006) aptly captures some of the problems caused by 1 and 2 in his critique of Shiffrin's (2000) definition of paternalism. Shiffrin provides her definition in the context of a defense of the unconscionability doctrine in contract law. De Marneffe points out that Shiffrin's definition is explicitly based on motives rather [p. 58] than on justificatory reasons.¹⁵ However, he notes, her aim and strategy is to defend policies from the charge of paternalism by providing alternative *justifications* for them. This defense is problematic because motives and justifications are different sorts of reasons. As De Marneffe goes on to note, policies can be motivated by certain reasons even though other reasons, which happen not to motivate, would provide an adequate justification. Conversely, policies need not be motivated by certain reasons even though these reasons provide the only adequate justification. Therefore, providing alternative justifications does not directly affect the status of a policy as paternalistic or nonpaternalistic on a motivational account of paternalism. (De Marneffe 2006, p. 71)

On the action-focused account, Shiffrin's project seems somewhat incoherent. She condemns benevolently motivated limitation of liberty as deeply insulting, yet she accepts just such limitation in the case of unconscionability since there are good justificatory reasons for this doctrine, unrelated to benevolence (these reasons are roughly to avoid being complicit to exploitation). She does not discuss how these proper justificatory reasons should be balanced against the deep insult that presumably remains as long as the benevolent motive remains. Furthermore, it is far from obvious how these justificatory reasons could remove or even mitigate the insult.

On the filter approach, in contrast, Shiffrin's different positions can be consistently accommodated. Translated to this approach, Shiffrin seems to hold, first, that

¹³ Some examples: VanDeVeer (1986, p. 22) says that behavior towards a person S is paternalistic only if it has "the primary or sole aim of promoting a benefit for S", Archard (1990, p. 36) says that behavior by P towards Q is paternalistic only if "P's belief that this behaviour promotes Q's good is the main reason for P's behaviour", De Marneffe (2006, pp. 73-74) says that a policy is paternalistic towards A only if "the government has this policy only because those in the relevant political process believe or once believed that this policy will benefit A in some way", and Dworkin (2010) says that X acts paternalistically towards Y by doing Z only if "X does so just because Z will improve the welfare of Y".

¹⁴ Feinberg clearly identifies both of these problems (1986, pp. 16-23). So does Husak, see esp. pp. 390-91. Feinberg uses "multiple" and "mixed" interchangeably. I will for the most part stick with "multiple". It should not be assumed that multiple rationales are a set of rationales that are individually sufficient (to motivate, to justify, etc.) - rationales may be of any strength or weight.

¹⁵ Other actions-focused accounts that are in this way motivational include Kleinig 1983; VanDeVeer 1986; Archard 1990; Husak 2003; Dworkin 2010.

in considering unconscionability, the fact that the weak party is benefitted does not play the role of a reason for the doctrine. She also seems to hold that certain other facts do provide good reasons for the doctrine. This is quite clear and consistent. The only remaining question is whether or not the fact that the promoter of unconscionability is motivated by the benefit to the weak party plays the role of a reason against such promotion. To hold that it does amounts to a sort of extreme antipaternalism that goes beyond the filtering out of reasons.

The problems that de Marneffe identifies in Shiffrin's account are related to the interplay of motivational and justificatory reasons. However, after his critique of Shiffrin, de Marneffe goes on to argue against a justificational account of paternalistic policy as well. The problem with these accounts is, de Marneffe notes, that it is hard to see how exactly justifications would sort actions into paternalistic and nonpaternalistic.

Before I look closer at this problem, I must briefly note a complication in de Marneffe's presentation. In describing the problem with deciding which actions are paternalist by looking at their justifications, and indeed throughout the article, de Marneffe talks about "paternalistic reasons" (and this is not a term he picks up from Shiffrin, who does not use it). He specifies this at one point as reasons "that cite some benefit to A that A does not want" (p. 72). This is unduly narrow, however, since it may presumably be paternalism to provide A with a benefit he does want, if this is done in a manner which limits his liberty (and that he does not want). More generally, it is not obvious, and is indeed unintuitive, that reasons can be paternalistic on the action-focused approach to paternalism. However, for ease of presentation [p. 59] and to avoid introducing new terms, I will use the term 'paternalistic reason' as a placeholder for whatever reasons are such that they make an action paternalistic by being reasons for that action. This may be de Marneffe's implicit intention as well (cf. p. 74, footnote 18).

De Marneffe's considers and rejects three justificational accounts of paternalism. The first is that a policy is paternalistic if and only if it can be justified only by counting paternalistic reasons in its favor. Since de Marneffe interprets 'justified only by' to imply 'justified by', all paternalistic policies are justified on this account, which is obviously not acceptable. The second account is that a policy is paternalistic if and only if it cannot be fully justified unless paternalistic reasons are counted in its favor. Given that 'unless' has the same truth conditions as 'or', all unjustified policies are paternalistic on this account, which is also unacceptable. The third account combines the first two to say that a policy is paternalistic if and only if it i) cannot be fully justified without counting paternalistic reasons in its favor, and ii) would be fully justified if paternalistic reasons counted in its favor. On this account, a policy for which there are only paternalistic reasons but which is unjustified (since these reason are not sufficient to outweigh the reasons against it) is not paternalistic. This is also unacceptable.

De Marneffe's response to the identified difficulties is to propose that a definition of paternalism, at least for Shiffrin's purposes of reconciliation, incorporate both a motive and a justification component. His preferred version of such a hybrid definition is:

A government policy is paternalistic toward A if and only if (a) it limits A's choices by deterring A from choosing to perform an action or by making it more difficult for A to perform it; (b) A prefers A's own situation when A's choices are

not limited in this way; (c) the government has this policy only because those in the relevant political process believe or once believed that this policy will benefit A in some way; and (d) this policy cannot be fully justified without counting its benefits to A in its favor. (pp. 73-74)

Among all action-focused definitions of paternalism that I am aware of, de Marneffe's best captures the reason aspect of paternalistic action. However, the definition still has peculiar implications. First, the definition inherits a problem from the three rejected accounts: If a policy targets a group of x persons and if it is motivated and justified by its effects on any set of $x-1$ members of the group, then it is not paternalistic towards anyone, by virtue of both (c) and (d). For no member of the group is it true that the government has the policy only because of the belief that the policy will benefit her, since the belief that it will benefit everyone else is sufficient to motivate the policy. For no member of the group is it true that the policy cannot be justified without counting its benefits to her in its favor, since the benefits to everyone else are sufficient to justify the policy. This implication is obviously undesirable, as it means that almost no policies are paternalistic. The definition can be reformulated to avoid this implication, but this would make it even more complex.¹⁶ [p. 60]

Other undesirable implications of de Marneffe's definition may be even harder to avoid. The definition is arguably both too wide and too narrow. It is too wide in that it implies that a policy is paternalistic towards A even if it limits A's choices (in a way that A prefers they were not limited) merely as a side effect of saving A from unwanted harm. For example, a policy may regulate the use of explosives and though A may prefer, *ceteris paribus*, to be free to choose whether or not to use explosives himself, this may be a fairly trivial preference in relation to his strong preference for general regulation. Still, the regulation is paternalistic towards A, given that benefits to A are necessary for its motivation and justification (or some weaker condition of this sort reformulated in light of the problem pointed out in the previous paragraph). Antipaternalists would not, I take it, be principally opposed to this policy on A's behalf.

De Marneffe's definition is too narrow in that it implies that a policy is *not* paternalistic towards a group even if it limits their choices against their will, promotes their good, is generally endorsed for that reason, is in fact unjustified because of its oppressive character, and was enacted with nonpaternalistic motives (perhaps a long time ago). Antipaternalists, I take it, would be opposed to this policy and would want their doctrine to condemn it.

Recall that there are two problems with action-focused definitions of paternalism: There are different sorts of reasons and actions most often have multiple rationales. De Marneffe attempts to deal with the first problem by including both motives and justifications in his definition. However, as shown by the 'overly narrow' objection, there are other sorts of relevant reasons, such as the reasons for which others than policy-

¹⁶ The problem cannot be avoided by simply tweaking de Marneffe's definition so that the relevant question is whether the policy can be expected to benefit each member of some group to which A belongs, as this would imply, implausibly, that a policy is not paternalistic toward A even if it limits her choices, she does not want it, and its benefit to her is a necessary part motive and part justification, as long as the policy also affects some other member, who is not (expected to be) benefitted.

makers endorse a policy. De Marneffe attempts to deal with the second problem by making reasons of both kinds necessary reasons rather than for example the only or main reasons, as on other proposed definitions.¹⁷ However, as show by the ‘overly wide’ objection, the ways in which a policy can limit someone’s choice and benefit her are more diverse than the definition can handle.

One might respond to the identified problems by constructing an even more intricate definition of paternalism, and undoubtedly someone will. The project of defining paternalistic action invites creative counter-examples, further specification and modification, further counter-examples, and so on. However, it is not clear that this method, fruitful in other contexts, is helpful in this case, since the filter approach avoids the identified problems, without defining paternalistic action at all.

De Marneffe’s explicit aim in the article is to refute antipaternalism, or what he calls “the general presumption against paternalism” (p. 69). In the latter two thirds of his article, he rather convincingly does so by arguing that neither paternalistic motives nor paternalistic justifications are inherently problematic. In the introduction, de Marneffe states that if “a general principle of antipaternalism is valid, then we should evaluate ... policies by evaluating whether or not there is sufficient [p. 61] *non*paternalistic reason for them.” (p. 69) This brief statement takes us very close to the filter approach - policies should be evaluated without regard to paternalistic reasons. Just as I have done, de Marneffe leaves it open to specification exactly which reasons are paternalistic and which are nonpaternalistic.

The characterization of antipaternalism that can be extracted from De Marneffe’s brief statement is unnecessarily restricted to evaluation, since there are other kinds of reason-taking contexts. It is also needlessly restricted to the issue of whether or not there are *sufficient* reasons for a policy. We may also be interested in for example whether there are stronger nonpaternalistic reasons for some policy than for an alternative policy. Still, this brief characterization is perhaps sufficient for de Marneffe’s normative investigation and I do not see how the following seven pages, where de Marneffe labors with a definition of paternalistic policy, does anything to advance either our understanding of antipaternalism or his case against this doctrine.

De Marneffe’s and Shiffrin’s contributions to the conceptual paternalism debate are, I believe, the state of the art in this area. However, they do not provide definitions of paternalism that enable us to employ this concept fruitfully in normative contexts. Their definitions cannot be used to capture that which antipaternalism is opposed to. De Marneffe’s two-line characterization of antipaternalism is more helpful than his seven-page definition of paternalism. None of this proves that the action-focused approach to antipaternalism is deficient through and through, but it is a strong indication. Considering also the general problems faced by any action-focused definition – the variety of sorts of reasons and the multiple rationales for most actions – the prospects for the action-focused approach are slim. Especially so since there is a ready alternative – the filter approach.

¹⁷ For example John Gray (1983) claiming that paternalism is “to coerce an individual solely in his own interest” (p. 90) and Archard (1990) objecting to this condition and proposing instead that promoting the good of the person is the “main reason” (p. 38).

4.5 Conclusion

Principled antipaternalism is the only objection to paternalism that has substantial independent normative thrust. It should be understood as demanding that certain facts do not play the role of reasons. This is the filter approach. Typically, the facts are of the form that some person will benefit from having her liberty limited or her autonomy infringed. The approach, however, is structural and conceptual, independent of which facts are targeted exactly. The approach is also independent of what reasons there are more generally and what role reasons in general play.

In the conceptual debate on paternalism, it is widely assumed that what needs defining is paternalistic action, conduct, behavior, policy or law. Relatedly, antipaternalism is typically understood in terms of resistance to these things. This dominant action-focused approach leads to intricate and unnecessary problems, as illustrated by Seana Shiffrin's and Peter De Marneffe's discussion of normative issues to do with paternalism. The failure of the action-focused approach to capture the proper role of reasons should lead us to favor the filter approach. There may be [p. 62] no escaping talk of paternalistic actions and paternalistic policies as shorthand in less stringent contexts, but these concepts are ill suited for careful normative investigations of the moral problems that allegedly surround paternalism.

The filter approach makes the most of the traditional liberal opposition to benevolent interference. The strongest antipaternalist position is reached by abstaining from sorting actions into paternalistic and nonpaternalistic, in favor of designing a plausible filter between facts and operative reasons.

References

- Archard, David. 1990. Paternalism Defined. *Analysis* 50(1): 36-42.
- Arneson, Richard. 1980. Mill versus Paternalism. *Ethics* 90: 470-89.
- Brock, Dan. 1988. Paternalism and Autonomy. *Ethics* 98(3): 550-565.
- De Marneffe, Peter. 2006. Avoiding Paternalism, *Philosophy and Public Affairs* 34(1): 68-94.
- De Marneffe, Peter. 2010. *Liberalism and Prostitution*. Oxford: Oxford University Press.
- Dworkin, Gerald. 1972. Paternalism. *The Monist* 56: 64–84.
- Dworkin, Gerald. 1983. Some Second Thoughts. In Rolf Sartorius (ed.), *Paternalism*. Minneapolis: University of Minnesota Press, pp. 105-11.
- Dworkin, Gerald. 2010. Paternalism. Zalta, E.N. (ed.) *The Stanford Encyclopedia of Philosophy (Summer 2010 Edition)*., URL = <http://plato.stanford.edu/archives/sum2010/entries/paternalism/>.
- Feinberg, Joel. 1984. *Harm to Others*. Oxford: Oxford University Press.
- Feinberg, Joel. 1986. *Harm to Self*. Oxford: Oxford University Press.
- Feinberg, Joel. 1990. *Harmless Wrongdoing*. Oxford: Oxford University Press.
- Gert, Bernard and Culver, Charles M. 1976. Paternalistic behavior. *Philosophy and Public Affairs* 6(1): 45-57.
- Guyer, Paul. 2014/2006. *Kant*. Second edition. New York: Routledge.
- Gray, John. 1983. *Mill On Liberty: A Defence*. London: Routledge & Kegan Paul.

- Grill, Kalle. 2007. The Normative Core of Paternalism. *Res Publica* 13: 441–458.
- Grill, Kalle. 2010. Anti-paternalism and Invalidation of Reasons. *Public Reason* 2: 3–20.
- Groll, Daniel. 2012. Paternalism, Respect, and the Will. *Ethics* 122(4): 692–720.
- Husak, Douglas. 2003. Legal Paternalism. In Hugh LaFollette (ed.) *Oxford Handbook of Practical Ethics*. Oxford: Oxford University Press, pp. 387-412.
- Husak, Douglas. 2013. Penal Paternalism. In Christian Coons and Michael Weber (eds.) *Paternalism: Theory and Practice*. Cambridge: Cambridge University Press, pp. 39-55.
- Kamm, Frances. 2011. Sen on Justice and Rights: A Review Essay. *Philosophy and Public Affairs* 39(1): 82-104.
- Kant, Immanuel. 1991/1797. *The Metaphysics of Morals*. Trans. Mary J. Gregor. Cambridge: Cambridge University Press.
- Kleinig, John. 1983. *Paternalism*. Manchester: Manchester University Press.
- Larmore, Charles. 1987. *Patterns of moral complexity*. Cambridge: Cambridge University Press.
- Mill, J.S. 1859. *On Liberty*. (Many editions)
- Quong, Jonathan. 2010. *Liberalism without Perfection*. Oxford: Oxford University Press.
- Rawls, John. The Idea of Public Reason Revisited. *University of Chicago Law Review* 64: 765-807.
- Rawls, John. 1999. *A Theory of Justice*. 2d ed. Cambridge, MA: Harvard University Press.
- Blame*. Cambridge, MA: Belknap Press of Harvard University Press.
- Sen, Amartya. 2009. *The Idea of Justice*. Cambridge, MA: Harvard University Press.
- Sen, Amartya. 2012. The Global Reach of Human Rights. *Journal of Applied Philosophy* 29(2): 91-100. [p. 63]
- Shiffrin, S. 2000. Paternalism, Unconscionability Doctrine, and Accommodation. *Philosophy and Public Affairs* 29(3), 205-250.
- Szasz, Thomas. 1992. *Our right to drugs: the case for a free market*. Syracuse NY: Syracuse University Press.
- Ten, C.L. 1980. *Mill on Liberty*. Oxford: Clarendon Press.
- Van de Veer, Donald. 1986. *Paternalistic Intervention*. Princeton: Princeton University Press.
- Waldron, Jeremy. 1989. Legislation and Moral Neutrality. In R. Goodin and A. Reeve (eds.) *Liberal Neutrality*. London: Routledge.
- Wall, Steven. 1998. *Liberalism, Perfectionism and Restraint*. Cambridge: Cambridge University Press.