

Anti-paternalism and Invalidation of Reasons

Kalle Grill
Uppsala University

Abstract: I first provide an analysis of Joel Feinberg's anti-paternalism in terms of invalidation of reasons. Invalidation is the blocking of reasons from influencing the moral status of actions, in this case the blocking of personal good reasons from supporting liberty-limiting actions. Invalidation is shown to be distinct from moral side constraints and lexical ordering of values and reasons. I then go on to argue that anti-paternalism as invalidation is morally unreasonable on at least four grounds, none of which presuppose that people can be mistaken about their own good: First, the doctrine entails that we should sometimes allow people to unintentionally severely harm or kill themselves though we could easily stop them. Second, it entails that we should sometimes allow perfectly informed and rational people to risk the lives of themselves and others, though they are in perfect agreement with us on what reasons we have to stop them for their own good. Third, the doctrine leaves unexplained why we may benevolently coerce less competent but substantially autonomous people, such as young teens, but not adults. Last, it entails that there are peculiar jumps in justifiability between very similar actions. I conclude that as liberals we should reject anti-paternalism and focus our efforts on explicating important liberal values, thereby showing why liberty reasons sometimes override strong personal good reasons, though never by making them invalid.

Key words: anti-paternalism, invalidation, Joel Feinberg, liberty, reasons.

In his argument against paternalism, Joel Feinberg states at one point that the anti-paternalist "must argue that paternalistic reasons [...] are morally illegitimate or invalid reasons" (1986, 25-26). In this article, I will develop an account of what it might mean that a reason is invalid on moral grounds in this sense, or *invalidated*, as I will call it. I will then go on to present four arguments against invalidation in the context of anti-paternalistic doctrines, and by implication against invalidation in general.

I will be concerned with the principled or deontological understanding of anti-paternalism, which tells us to disregard certain reasons as a matter of moral requirement, rather than of practical expedience. There are often practical reasons to disregard certain reasons: Lack of time or information, risk of mistake, social coordination mandating a division of responsibility, and so on. Feinberg addresses his anti-paternalism to "an ideal legislator" and claims to be on "a quest not for useful policies but for valid principles." (1984, 4) I take it that it is on this ideal, general or abstract level that reasons can be invalidated, as opposed to excluded for practical reasons.¹ Though practical considerations and more principled arguments have as a rule been intertwined in the discussion on paternalism since John Stuart Mill set the standard with *On Liberty*, I find it preferable to keep them separated. We can fruitfully discuss paternalism either in the abstract or in some concrete

1] Joseph Raz' (1990 [1975]) account of "exclusionary reasons" as reasons not to act on other reasons is based on coordination problems and on practical constraints on deliberation. I take invalidation to be more purely normative. The ideas also differ in that it is a good thing to act in accordance with excluded reasons, though not for them, while invalidated reasons are more thoroughly emptied of normative significance.

situation or institutional setting. Anything in between is likely to fail both to account for the normative core of the problem and to provide action guidance.

Feinberg oscillates between understanding paternalism on the one hand as the counting of certain reasons as good and valid reasons for limiting liberty (as in the quote in the first paragraph of this section), and on the other as limitations of liberty with a certain (implicit) rationale. Though the latter understanding is dominant in contemporary discussion of paternalism, I believe the former is the most promising for the anti-paternalist. It is less ambitious to oppose certain reasons for limiting liberty, than to oppose certain actions or policies because they limit liberty and are supported by certain reasons.²

Feinberg's uncompromising defence of anti-paternalism on moral grounds is unusual. The standard approach to paternalism among contemporary authors is to assume anti-paternalism as a general rule and then propose exceptions to this rule. These exceptions may invoke hypothetical consent (e.g. Van De Veer 1986, 88), avoidance of great harm (e.g. Groarke 2002) or the common-sense reasonableness of preventing significant harm at limited cost to liberty (e.g. Beauchamp and Childress 2001, 185-91). There has been a lack, however, both of detailed analysis of principled anti-paternalism and of more comprehensive arguments against it. This may be due to the strong focus on practical and political circumstances and to the perceived need to decide which actions and policies are rightly called paternalistic. Recently, some authors have presented more head on critiques of anti-paternalism, most noteworthy Richard Arneson (2005) and Peter de Marneffe (2006). With this article, I hope to join that effort and take it one step further.

I will focus mainly on Feinberg's absolutist understanding of anti-paternalism according to which all paternalistic reasons are invalid. I will also, however, consider moderate versions according to which paternalistic reasons are discounted rather than invalidated, or where exceptions are made for certain types of paternalistic reasons. Throughout I discuss moral reasons. Reasons that are invalidated as moral reasons may possibly remain valid as reasons of some other sort (e.g. prudential).

Three of my four arguments are based on the fact that anti-paternalism necessarily only applies to sufficiently voluntary action. This, I argue, leads to wrong conclusions in some cases, to peculiar jumps in justifiability, and to an unwarranted disregard for the liberty of those whose actions are, according to the doctrine, insufficiently voluntary. The remaining argument is based on the fact that we can have decisive reasons to interfere with a person who is acting perfectly voluntarily, and she can accept these reasons fully as far as they concern her, yet anti-paternalism unreasonably entails that these reasons are invalid. All of my arguments concern the peculiar effects of invalidation per se. None of them depend on people being mistaken about their own good. Though I happen to believe

2] For an extensive defence of this view, see Grill 2007.

that people can be so mistaken, I will assume throughout, for the sake of argument, that we should accept people's views of their own good at face value.

I. ANTI-PATERNALISM AS INVALIDATION

I take it for granted that if an action protects or promotes some person's good (more than alternative actions) that is normally a valid reason for that action. I will from now on call such reasons *personal good reasons* rather than paternalistic reasons, since they can be invoked for actions that do not limit liberty and since, strictly speaking, reasons as such cannot be paternalistic. Personal good reasons may concern such things as a person's health, prosperity, achievement, happiness, or long-term autonomy. I will not be concerned with what exactly is good for a person, but rather, as already stated, assume that her own view on this matter should be accepted. Neither will I make a distinction between protection and promotion of good – that is between preventing harm and providing benefit. If the reader thinks that such a distinction is warranted, she may read my argument as concerning harm-prevention.

Anti-paternalism is, on my understanding, a doctrine of invalidation. We may describe invalidation as the blocking of a reason from influencing the moral status of an action (forbidden, permissible, obligatory etc.) and so what we ought to do (not do, may do etc.). To influence an action's moral status is, I take it, to have weight on the scales, to figure among the factors that should be considered when forming an all things considered judgement (under ideal conditions). A reason may have influence in this sense even if it is ultimately overridden by other reasons. We could say with John Broome (2004) that the reason figures in a "weighing explanation" of an ought fact. However, while Broome says that deontic principles replace weighing explanations (making them merely potential), I propose that some deontic principles operate by regulating the weighing, for example by making certain reasons invalid, banishing them from the scales. The picture of reasons on the scales is of course metaphorical. I admit that I do not have a theory of how exactly to derive an all things considered judgment from a set of valid reasons. I can only say, with Broome, that such a judgment should be based only on consideration of the strengths of the relevant reasons (or more Broomian – ought facts are determined by the aggregated weights of the relevant reasons) (37-38).

Invalid reasons are normatively impotent and in that sense not reasons at all. This presents a terminological problem. On the one hand, we are very tempted to talk, like Feinberg, of invalid reasons. On the other hand, it is a widespread position that the term reason should be reserved for valid reasons, for reasons with influence (e.g. Scanlon 1998, 156; Kamm 2006, 237; Parfit Manuscript). Philosophers sometimes disagree on what we should do, or at least why we should do it, though they agree on what is important. For example, non-consequentialists typically agree with hedonistic utilitarians that the fact that an action will alleviate suffering is a relevant consideration in a sense that most other consequences are not, even if they hold that this consideration does not always

provide a reason for action. So for example, Thomas Scanlon (1998) in his critique of a morality based on wellbeing readily admits: “It would be absurd to deny that well-being is important” (141). Since in the present context the very object of investigation is whether apparent reasons are valid reasons, I find it preferable to use “reason” for all considerations that are relevant in this wide sense. A reason’s strength is the influence it would have if it were valid.³ I find that this terminology clarifies the role of deontic constraints on moral reasoning, for example the constraints of anti-paternalism, and so one important difference between consequentialism and non-consequentialism.

Principles or doctrines of invalidation can be placed within a larger family of influence-regulating doctrines – doctrines that strike a wedge between the strength of a reason and its influence. The most widely recognized form of morally based influence-regulation is side constraints (introduced by Robert Nozick (1974) in the context of rights), or, in other words, absolute reasons. An absolute reason entails that an action should or should not be done, irrespective of other reasons.⁴ This a priori-like property of a reason entails that other reasons have no influence – they should be disregarded. As a relationship between reasons, invalidation is distinct from and weaker than the relationship between absolute and non-absolute reasons, since it entails only that some reason(s) have no influence, leaving other reasons to potentially override the invalidating reason. In fact, absolute reasons can be defined as reasons that invalidate all other reasons, with invalidation the more basic concept.

Another form of morally based influence-regulation can be found by applying the idea of lexical ordering to types of reasons. Lexical ordering or priority was introduced in moral theory by John Rawls (1971) in application to principles of justice and has later been explored in application to values (e.g. Griffin 1986). It can similarly be applied to reasons. Lexical ordering allows reasons to be absolute in relation to reasons lower in the hierarchy while standing in a weighing relationship to reasons on the same level. Invalidation is in one sense stronger and in one sense weaker than lexical ordering.

Invalidation is weaker than lexical ordering in that it is not hierarchical. Reason of a type that is invalidated by another type can still override reasons of a third type, that can override reasons of the invalidating type. For example, assume that reasons of type H are invalidated by reasons of type L. Now there may be other reasons, say of type O, which can override reasons of type L, but which can in turn be overridden by reasons of type H.

3] I aim to follow the standard use of “strength” (sometimes “weight”) employed by e.g. Raz (1990 [1975]). Joshua Gert (2007) has argued convincingly that reasons have two distinct dimensions of strength – requiring strength and justifying strength. I will disregard this complication since my argument holds for both kinds of reasons. I will however avoid talk of the “balance of reasons” and stay with the more general “influence the moral status of actions”. When I say that one reason is stronger than another, this should always be understood as concerning the same kind of strength, typically requiring strength.

4] The term from e.g. Raz 1990 [1975], 27. Such reasons are sometimes called “decisive”, but “absolute” is preferable because “decisive” other times refer to what Raz calls “conclusive” reasons – reasons that are not overridden in a certain case (though they can be).

In fact, this is exactly what anti-paternalists typically claim. H reasons (to prevent harm to a person) are invalidated by L reasons (to respect the liberty of the same person), while O reasons (to prevent harm to another person) can override L reasons, in accordance with the harm principle. Further, H reasons can override O reasons since it can be right to prevent serious harm to one person instead of preventing less serious harm to another (when there are no liberty reasons either way). In contrast, reasons of a type that is lexically prior to reasons of a second type cannot be overridden by reasons of a type that can be overridden by reasons of the second type. For example, that Rawls' basic freedoms and liberties are lexically prior to concerns of distributive justice means that there can be no type of reason that override liberty reasons while being overridden by distributive justice reasons.

Invalidation is stronger than lexical priority in that it not only makes one reason dominate another, it makes the dominated reason completely impotent. One important point of lexical ordering is to capture cases in which reasons at one level of priority balance each other out, in which case the matter is decided by reasons on the next lower level, unless they also balance each other out, and so on. If a reason is invalidated by another reason, however, it has no influence and cannot decide anything.

My account of invalidation is an attempt to explicate a central aspect of the anti-paternalist position. A doctrine of invalidation should ideally be specified to a class of actions and a class of reasons. The doctrine then says that reasons in the relevant class do not influence the moral status of actions in the relevant class. This might have the further implication that being motivated by such a reason to perform such an action, or accepting it as justification for such an action, is inappropriate or condemnable in a way that warrants disapproval or punishment, perhaps because it manifests a bad attitude of some sort. However, the more basic idea is the normative invalidity of the reason. Whether failure to see this invalidity is immoral and if so what is the appropriate response to such immorality – these are secondary questions.⁵

II. INTERFERENCE AND VOLUNTARINESS

Anti-paternalism is, on my understanding, the influence-regulating doctrine that personal good reasons are invalid for a certain class of actions. Call these actions “problematic interferences.” On most accounts, actions in this class must be liberty-limiting in some sense. Anti-paternalism typically does not entail that personal good reasons are invalid for innocuous involvement in other people’s life, such as greeting them in the street or giving them small gifts.⁶ Specifying this criterion of problematic

5] This means that I disagree with accounts of paternalism which claim that a defining criterion of paternalism is the manifestation of a certain attitude. That is if our concept of paternalism is supposed to capture what liberals are traditionally opposed to.

6] Sunstein and Thaler (2003) argue for what they call “libertarian paternalism”, postulating that “a policy therefore counts as ‘paternalistic’ if it attempts to influence the choices of affected parties in a

interferences is difficult, but I will not discuss these difficulties here. Nor will I consider what criteria there are exactly. Instead, I will focus on the voluntariness criterion for inclusion in the class of problematic interferences, and assume that whatever other criteria there are, are fulfilled.

On any reasonable account, anti-paternalism protects only choices or actions that are *sufficiently voluntary* (Feinberg 1986, chapter 20). It is not problematic interference to restrain people in rage or panic, people heavily influenced by drugs, people with severe mental disorders, or small children (unless, when applicable, this is in conflict with their previous, voluntary choice). Importantly, voluntariness cannot be determined by the reasonableness of choices or actions (Feinberg 1986, particularly chapter 20, e.g. p. 133). The very point of anti-paternalism is to protect imperfectly reasonable choices.

For Feinberg, a person acts perfectly voluntarily if she is competent, there is no coercion or duress, no subtle manipulation, no ignorance or mistake and no distorting circumstances (such as excitement, strong emotion, or time pressure) (1986, 115). Of course, hardly any choice is perfectly voluntary. Feinberg explicitly proposes that a person should be protected from paternalism if her actions or choices are “voluntary enough” (chapter 20). Given that the other criteria for problematic interference are fulfilled, anti-paternalism entails that sufficient voluntariness functions as an invalidator of personal good reasons. In other words, liberty reasons invalidate personal good reasons on the condition that the person acts sufficiently voluntary.⁷

III. FIRST ARGUMENT – VOLUNTARY CHOICE CAN LEAD TO DISASTER

In this and the following sections, I will present a series of scenarios. The scenarios include assumptions on what reasons there are and their relative strength. I will not defend these assumptions. I take it for granted that the specifics of the cases can be adjusted so that these assumptions are reasonable. Remember that the way I use the term reason, reasons need not have influence. That a reason for an action has a certain strength means only that we can promote or protect some value by performing that action. If the reader

way that will make choosers better off.” (1162) A policy of providing information upon request because this makes for wise choices would thus be paternalistic. Though it is possible to oppose even such non-interfering policies I will focus on the more plausible because more restricted anti-paternalism that targets only liberty-limiting benevolence.

7] Interference with insufficiently voluntary choice is sometimes called “soft paternalism” but is typically not opposed by anti-paternalists. Richard Arneson (2005) claims that soft anti-paternalists, accepting soft but not hard paternalism, occupy an unstable position between full acceptance of paternalism and hard anti-paternalists, who allegedly do not think that lack of voluntariness makes interference any more justified (266-68). However, even hard anti-paternalists must embrace some of the components of Feinberg’s concept of voluntariness, minimally competence and lack of duress – infants and people at gunpoint may certainly be stopped from harming themselves. The distinction between hard and soft anti-paternalism, though useful, is ultimately a matter of degree. There is no remotely reasonable end point position to occupy. (One could of course change terminology by claiming that interference with non-voluntary action is not paternalism at all, but this changes nothing of substance, see further below in the section “VI. Third Argument – Liberty for all”.)

finds that there is no possible specification under which my scenarios make sense, this must be because she has a radically different view on what has value.

Voluntariness often has an indirect effect on the moral status of actions, since a high degree of voluntariness normally makes for good decisions, which normally makes for good consequences. When things are not normal, however, a person's very voluntary choice can lead to catastrophic consequences. Consider:

The Bridge. A person tries to cross a bridge but we have a chance to stop her. We know that the person wants to live and is well informed about the general condition of the bridge, is acting in character, is calm and collected, attentive, mature and intelligent, under no duress or pressure, etc. Stopping her would interfere with her liberty to move around freely, which is a strong reason against doing so. However, being equally well informed about the condition of the bridge, and having considered its durability more thoroughly, we firmly believe that, appearances to the contrary, the bridge is unsafe. Stopping the person would therefore most probably prevent her from falling to certain death, which is a much stronger reason for doing so.

We know that the person in *The Bridge* scores high on the standard aspects of voluntariness. Presumably, she just happens to be wrong, as very able people sometimes are. Perhaps she did not bother enough to analyze the available information, perhaps she miscalculated this one time and so reached the wrong conclusion. Whatever the cause, her very voluntary choice happens to be very bad for her. Anti-paternalism seems to imply that our reason to stop the person is invalid and so we should not stop her (unless there are other reasons to do so). That is the wrong conclusion – it is morally unreasonable. We should stop the person, because otherwise she will most probably die.

In his original bridge case, from which *The Bridge* is adapted, Mill argues that the person may be coercively turned back “without any real infringement of his liberty; for liberty consist in doing what one desires, and he does not desire to fall into the river.” (1991 [1859], 107) Even if it is true that the person does not want to fall into the river, this does not mean that he does not want to cross the bridge (for a thorough argument, see Day 1970). We certainly desire some things though we do not desire their consequences (indeed we may do so even when we anticipate the consequences).

Crossing the bridge is most probably in conflict with the person's goals and values. In that sense her behaviour rests on a mistake. This mistake is small in terms of information processing, but great in terms of consequences. Voluntariness could be defined as efficient goal-satisfaction, in which case the person in *The Bridge* probably acts non-voluntarily. However, on Feinberg's concept of voluntariness, ignorance of the consequences of one's action is only one of many voluntariness-reducing factors and comes in degrees. That is why I can claim that the choice to cross in *The bridge* is very voluntary, and that is why, as

we will see in the next section, Feinberg introduces a variable standard for how voluntary is sufficiently voluntary.⁸

In general, very small imperfections in people's decision-making process can make for disastrous decisions. If the consequences of an imperfect decision are bad enough and if these consequences can be avoided without losing anything of comparable value, they should be avoided. We can save the person in *The Bridge* because, in this particular case, we understand the consequences of her actions better than she does and we have an opportunity to intervene. With all the complexities of modern life and with all the expert knowledge and the sophisticated forms of intervention available, there are many such cases.

It may seem that we can save anti-paternalism by insisting that no good can be brought about by interfering with voluntary choice. This may be a conceptual argument, claiming that what a person voluntarily chooses is always, by definition, good for her. Or the argument may be empirical, claiming that because of our great ability to make good choices for ourselves, and our great inability to help others, the outcome of sufficiently voluntary actions can never be improved by intervention. Both assumptions seem incredible, raising the standards of sufficiently voluntary action to inhuman levels and so leaving most actual choices and actions outside of the doctrine's domain. That is unless we accept a very scattered or even inconsistent idea of the good (in the conceptual case) or a very pessimistic view of human benevolence (in the empirical case).

Importantly, if it were nonetheless true that interference with sufficiently voluntary choice could do no good, anti-paternalism of the form under investigation would be saved only at the price of redundancy. If no good can come from interference, we have no reason to interfere. Then there are no reasons for anti-paternalism to invalidate, and so the doctrine has no application.

A popular moderation of anti-paternalism makes exceptions for personal good reasons that concern autonomy (e.g. Dworkin 1972; Kleinig 1983). On such a moderation, anti-paternalists could plausibly defend stopping in *The Bridge* as autonomy-preserving (if autonomy is preserved by saving a life). Most any harm can diminish autonomy in some way to some extent and so this moderation can entail that some personal good reasons are always valid, making for a substantially weakened doctrine. However, given a more narrow conception of autonomy, cases can be constructed that are analogous to *The Bridge* but where the harm does not diminish autonomy (financial ruin or disfigurement or horrific but temporary pain may be substituted for falling to certain death).

8] Danny Scoccia (2008, 358) has argued that voluntariness on Feinberg's account is about successfully furthering one's own values, or acting as one would have done if one were perfectly informed, rational and capable. If this is true, Feinberg's anti-paternalism is not about letting people lead their own lives, but rather about helping them promote their values, through either interference or non-interference, depending only on what is most efficient. I think this interpretation fails to appreciate Feinberg's strong commitment to personal sovereignty.

Another form of moderation is to accept all personal good reasons as partially valid, though discounted by some factor (e.g. Groarke 2002, 219). Such versions will entail that we should stop the person in *The Bridge* if the reason to save her is strong enough and the discount factor large enough that the reason to respect her liberty is overridden. However, in similar cases, where the difference between the strength of the reasons for and against stopping is smaller, such versions will entail the morally wrong conclusion. The larger the factor (the smaller the discounting), the fewer such cases, and the less convincing. The most convincing cases must be constructed to match the exact discount factor. The larger the factor, the weaker the doctrine, and so, I think, the more reasonable.

In conclusion, any non-redundant anti-paternalism of substance will lead to morally wrong conclusions in some cases. Therefore, anti-paternalism should be rejected.

IV. FEINBERG'S COMPROMISE

According to Feinberg, what is a sufficiently voluntary choice varies with “the nature of the circumstances, the interests at stake, and the moral or legal purpose to be served.” (1986, 117) In particular, the threshold depends on the severity of the risks involved – the higher the risks, the higher the threshold – and on whether the risks include irrevocable harm (118-121). This is a very reasonable move to make for any anti-paternalist who, like Feinberg, is concerned not simply to “prevent people from acting with low degrees of voluntariness”, but rather to “prevent people from suffering harm that they have not truly chosen to suffer.” (119)

Tying the threshold for sufficiently voluntary action to risk amounts to a compromise between unmodified anti-paternalism and consequentialism. High risks normally entails strong reasons for interference. On Feinberg's account, personal good reasons do not support problematic interferences, but they do influence whether or not an action is a problematic interference. In this roundabout way, personal good reasons can override liberty reasons. However, the central structure of the doctrine remains intact: If an action is voluntary enough, personal good reasons are invalid.

The compromise makes anti-paternalism more flexible and so more reasonable. The anti-paternalist can now hold that we should stop the very voluntary person in *The Bridge* because she takes a very great risk, while we should let people who are less informed and rational go about their lives, taking smaller risks. However, the basic argument from *The Bridge* remains in force. The tiniest deficits in voluntariness can still make for disastrous consequences. Feinberg must accept that disaster is morally irrelevant as long as the degree of voluntariness is high enough to match the risk. Furthermore, high degrees of voluntariness need not correspond to great liberty interests. The compromise does not change the fact that high degrees of voluntariness invalidates very strong personal good reasons even when the voluntary choice is trivial from the point of view of liberty. This, I think, is unreasonable.

Severe risk should perhaps be accepted for the sake of important liberties, but not simply because choices are very voluntary. This point can be illustrated by comparing a well-planned philosophical suicide with a five-party game of Russian roulette. Feinberg's account entails that since the suicide is five times more risky than the Russian roulette, we should accept a much lower degree of voluntariness for the roulette. Given that the circumstances are similar in other respects, this seems absurd. What should count against stopping these activities is not so much the degree of voluntariness of the agents involved, but rather the sort of liberty at stake and its value.

V. SECOND ARGUMENT – JUSTIFIED INTERFERENCE WITH PERFECTLY VOLUNTARY ACTIONS

In *The Bridge*, we assumed that the person failed to appreciate the risks involved. Even if she scores very high on voluntariness, there are imperfections, which happen to be very important. Stopping presumably furthers her good and so does not conflict with her hypothetical, enlightened self-interest. An anti-paternalist impressed with the first argument above may restrict her doctrine to perfectly voluntary action. However, consider this case:

The Stunt. A person tries to perform a spectacular stunt but we have a chance to stop her. We know that the person is acting perfectly voluntarily. Stopping her would interfere with her liberty, which is a strong reason against doing so. Stopping her would also eliminate a substantial risk to her health, which is a relatively weak but non-negligible reason for doing so. In addition, stopping her would eliminate a small but real risk of harm to an innocent and non-consenting passer-by. This is a strong reason for doing so, but weaker than the reason against. In the aggregate, however, the two reasons for stopping are stronger than the one reason against.

Remember that the specifics must be filled in to make the assumptions reasonable. For example, the exact risk to the passer-by can be adjusted to match the assumed strength of the reason to protect her (we certainly accept some risk to innocent and non-consenting passer-bys even from very superficial activities such as driving around for fun).

The Stunt illustrates the important and, in the paternalism debate, neglected fact that there are normally several reasons for and against any particular action. That some of these reasons are strong in no way excludes the possibility that a weak reason is decisive (tips the balance). Anti-paternalism clearly implies that we should not stop the stunt (unless there are other reasons to do so), since the personal good reason is invalid. That is the wrong conclusion. We should stop the person because of the risks to herself and to the passer-by.

It might seem surprising that we can have overriding reason to stop a perfectly voluntary action. The person obviously considers the reasons for performing the stunt

stronger than those against. Since, by assumption, she also has a correct view of her own good, we cannot question her reasons for performing the stunt, nor her self-interested reasons against. However, we can disagree about her reason to avoid risking the health of the innocent passer-by. She may have little concern for the wellbeing of some passing stranger, while we (correctly) judge that this consideration decides the matter against performing the stunt.

By assumption, there must be agreement on the person's self-interested reasons for and against acting. More importantly, however, the person may very well agree with us concerning those of our reasons for stopping her that concern her. Though she values her own health higher than other people's, she may still agree with us on the exact strength of our reason to stop her (weak but non-negligible). She may also agree that this reason is valid. It might well be that she opposes our intervention only because she thinks that we exaggerate the strength of our reason to protect the passer-by. Anti-paternalism therefore implies that we should disregard as invalid a reason that concerns the good of a person, even though she is herself in perfect (and perfectly voluntary) agreement with us concerning the strength and validity of this reason.⁹

It could perhaps be argued that we have no reason to eliminate voluntarily assumed risks to start with. This is not a view about invalidation then but about what has value. It is an unreasonable view. People may regret having to choose between two risky options and still make a choice, voluntarily. If the circumstances are harsh, the risks may be great. If we can do something to lower these risks, without losing anything of comparable value, we should do so. If you voluntarily choose to risk your life driving to work by route A over risking your life by driving to work by route B, this in no way implies that I (working with city planning) have no reason to try to decrease the risks involved in driving either route.

As for moderations, substituting discounting for invalidation can entail that we should stop the stunt. If protecting the person from herself is a valid but discounted reason for stopping, it may, together with the reason to protect the passer-by, override the liberty reason against. However, there will be similar cases, with the relative strengths of the three reasons adjusted to match the discount factor, where such moderate anti-paternalism entails that we should not stop the stunt. As in *The Bridge*, the larger the discount factor, the weaker the doctrine, and the less convincing the counter-examples.

By modifying *The Stunt* we can see how moderation by discounting is in one way sharply distinct from absolute anti-paternalism. Assume that the liberty reason and the protection of the passer-by reason are equally strong. Now the case is a sort of moral dilemma if the personal good reason is invalid. However, as long as this reason has some influence, however small, we should stop the stunt. Any discounting version, regardless of the discount factor, will give the same answer. Absolutist anti-paternalists, on the other

9] This could be avoided if perfect voluntariness entailed correct judgment of the relative strength of all reasons. This, however, seems an unwarranted conflation of concepts and would only mean that anti-paternalism is redundant for perfectly voluntary action.

hand, will not like the idea that the personal good of the stunt artist should decide the case if favour of intervention.

In conclusion, we should sometimes interfere with people in order to protect or promote their good, as well as the good of others. We should do so even if the action we interfere with is perfectly voluntary and we accept people's view of their good at face value. We should do so because we have overriding reasons to, both reasons concerning the good of others, and reasons concerning the good of the person. Anti-paternalism entails that we should disregard or discount the latter reasons, even if the person in question is in informed and perfect agreement on the strength and validity of these reasons. In effect, anti-paternalism tells us to let people harm themselves in ways they would not if they were not mistaken about their reasons to prevent harms to others. This is peculiar and arguably in conflict with liberal fundamentals. Therefore, anti-paternalism should be rejected.¹⁰

VI. THIRD ARGUMENT – LIBERTY FOR ALL

The first two arguments both accuse anti-paternalism of unacceptable disregard for important personal good reasons. Somewhat paradoxically, anti-paternalism may indirectly lead to disregard also for important liberty reasons. The doctrine only protects sufficiently voluntary choices. I propose that liberty is important not only for the most informed and capable decision-makers, but also for the demented, for minors, for the ignorant, and for people under time pressure. It is not as if the value of controlling (to some extent) one's own life kicks in only at a certain degree of voluntariness. Perhaps there is some level under which people cannot choose for themselves or cannot appreciate self-determination. Liberty, however, has value for people that are well above this level but that we should nonetheless sometimes coerce in their own interest, for example people in their lower teens. With this, anti-paternalists should agree. Why then does not interference with the liberty of young teens activate the doctrine?

Three answers are possible for the anti-paternalist: First, she can insist that the liberty of young teens (and generally people acting under the sufficient degree of voluntariness) is of another type than the liberty of informed and rational adults and so does not activate the doctrine. This distinction in value is mysterious. The anti-paternalist can say that the important value is not liberty but autonomy, and that young teens are not fully autonomous. They may not be, but neither are many adults, and young teens are surely autonomous to some extent. There is no difference in kind between the self-determination of more and less capable decision-makers.

Second, the anti-paternalist can lower the threshold and claim that the liberty of young teens (etc.) should never be limited in their own interest, and in fact that their well-

[10] De Marneffe (2006, 77-79) argues against anti-paternalism that since the government can justifiably and without insult substitute its judgement for an individual's for the sake of others, it can do so for her own sake too. My argument, in contrast, does not presuppose that there is no morally relevant distinction between first and third party interests in this context.

being is not even a valid reason for limiting their liberty. We have seen the problems such a view entails even for capable decision-makers. Even those who stubbornly bite the bullet and let the people in *The Bridge* and *The Stunt* kill or harm themselves, and others, cannot reasonably accept such passivity in relation to young teens.

Third, the anti-paternalist can say that for young teens, the value of liberty is appropriately reflected in the strength of liberty reasons, and so there is no need for invalidation. This is the most reasonable answer. But if this is the anti-paternalist's answer, it is entirely unclear why things should be any different for adults. Interference with more capable decision-makers generally yields smaller rewards, since there is less room for improvement. It may also be that interference with more capable decision-makers has a greater cost in terms of liberty, because more liberty (or autonomy) is sacrificed in some sense. However, none of this indicates that personal good reasons should be invalid.

The argument does not depend on the arbitrariness of any threshold between more and less capable decision-makers. Nonetheless, it may seem easier to make a distinction in value (as in the first answer above) if there is a neutral or obvious level at which to draw the line. Anti-paternalists often do make a sharp distinction between competents and incompetents, between the healthy and the (mentally) ill, between adults and children. These categories, however, depend on underlying physical properties, which vary by degree. It may be thought that legal status provides a less arbitrary basis for a threshold. This is not so. Once bestowed, legal status may admittedly make a normative difference. It is perhaps worse to limit the liberty of a person of age, because this frustrates legitimate expectations not to be so treated that are induced by the legal system. However, such legal circumstances can only reinforce an underlying moral principle, which must be spelled out in terms of non-legal, concrete physical or psychological properties of persons. It would be hopelessly vacuous to argue that the people that we must protect from benevolent interference are those that have been granted a legal right to be so protected. A moral principle like anti-paternalism should help us decide upon such matters as when people should reach lawful age, and so cannot itself depend on the answers to such questions.

In conclusion, anti-paternalism fails to address the issue of whether or not to interfere with less than sufficiently voluntary choices made by rather autonomous people such as young teens. It makes no sense that the liberty of people who make more voluntary choices should trump other concerns, while the liberty of people who make somewhat less voluntary choices has no special moral status whatsoever. Therefore, anti-paternalism should be rejected.

VII. FOURTH ARGUMENT – JUMPS IN JUSTIFIABILITY

I will now present an argument that draws on the fact that voluntariness is entirely a matter of degree (as noted by Feinberg (1986, 104)). Arneson (2005) correctly observes that Feinberg's anti-paternalism is committed to "the enormous overriding importance of the line between self-harming choice that is not quite voluntary enough and choice

that just passes the threshold of being voluntary enough.” (268) Sometimes lines have to be drawn and sometimes much depends on whether or not a threshold is reached. For example, it is very reasonable that if a bridge is safe enough we will walk over it, while if it is not safe enough we will take a long detour and lose valuable time. This is reasonable even if the threshold is somewhat arbitrary and small differences may shift the status of the bridge from safe enough to not safe enough. For another example, if we want to categorize actions into justified and unjustified actions, there will arguably be a grey area of actions difficult to fit into either category. However, in the present case we are considering two theoretical perspectives, anti-paternalism or no anti-paternalism, only one of which demands that we draw a line at all. If drawing a line leads to unreasonable consequences, this speaks against the perspective that demands that we do so.

The problem with a threshold, however, is not only its unreasonably great importance in determining what we ought to do. The threshold does not determine directly the moral status of actions, but rather whether certain reasons have influence on this status. A valid reason makes an action more (or less) justified. Anti-paternalism implies, therefore, that an action that affects a person whose choice or action is just below the threshold of sufficient voluntariness may be much more justified than an otherwise similar action that affects a person whose choice or action is just over this threshold. At the threshold the justifiability of the action takes a ‘jump’.¹¹ The line gives rise to a gap, and a very large gap if the reason is a strong one. Consider:

The Suicide. A person tries to kill herself but we have a chance to stop her. Stopping her would interfere with her liberty, which is a strong reason against doing so. Stopping her would also save her life, which is a much stronger reason for doing so.

Anti-paternalism implies that whether or not we should stop the suicide depends on the person’s degree of voluntariness (unless there are other reasons to do so). If the voluntariness is under the threshold, we have much stronger reason to stop it than not, and so stopping is clearly and with good margin justified. If, on the other hand, the voluntariness is over the threshold, we have only, as far as valid reasons go, a strong reason not to stop it, and so stopping is clearly and with good margin unjustified. I propose that this jump in justifiability is unacceptable. In fact I doubt that we can bring ourselves to even comprehend it, at least when it depends on infinitesimal differences – one further tiny piece of information, or a tiny improvement in decision-making capacity, will imply that an action that would have been overwhelmingly unjustified becomes overwhelmingly justified. Moderation by discounting would make the jumps smaller but not affect the heart of the problem. Moderation by exception might make this particular example

[1] It may be thought that such jumps are common in the law, since it is much more justified to punish a person who has barely committed a crime than one who has not. However, I propose that this is due entirely to practical considerations such as the efficiency and transparency of the legal system. Disregarding such considerations, bare crimes are not much less justified than almost crimes.

irrelevant (if suicide diminishes autonomy for example), but there will be analogous examples with the same force.

On our assumption that people's views of their good should be taken at face value, it may seem that we cannot have a strong reason to stop the suicide, since the person herself has judged that her life should end. However, the person might agree with us on the value of her survival, but still want to kill herself for some higher purpose, that we find lacks value. Just like in *The Stunt*, there may be no disagreement concerning the value of the person's survival and liberty, but only concerning the value of some other thing, in this case something for which she wants to sacrifice her life, such as her family honour, the independence of Tibet, or the glory of God (assuming these things are not part of her good).¹² It may also be that the person agrees with us that she has stronger reasons to keep living, but feels compelled to end her life, perhaps out of despair. In such cases, of course, she is not acting perfectly voluntarily.

It may be suggested that anti-paternalism does not in fact draw a sharp line between problematic interference and other actions, since there is an area in between in which we simply do not know, or where it is genuinely indeterminate, whether an action is or is not an interference. If the indeterminacy is epistemic, the doctrine is none the more reasonable for it, only more difficult to apply. If the indeterminacy is real (ontological), however, the threshold is turned into a hole – it is really sometimes indeterminate whether or not personal good reasons give valid support to an action. Consequently, some answers to moral questions are replaced by no answers. The cases discussed in previous sections (*The Bridge*, *The Stunt*) involve people of very high and even perfect voluntariness, and so are hardly indeterminate. For the *Suicide*, however, the sudden jumps in justifiability could be replaced with a twilight zone of indeterminacy. Stopping would be overwhelmingly justified on one side of this zone, overwhelmingly unjustified on the other side, and indeterminate in between. We avoid jumps by giving up comprehensiveness. To my mind, this is no improvement.

Another and better strategy for avoiding a sharp threshold is to reformulate anti-paternalism as follows: Personal good reasons for an action are valid only *to the extent* that the action is not a problematic interference. In other words, the influence of a reason for an action is some function of two variables – strength and the degree to which the action is a problematic interference (which in turn depends partly on the degree to which it is voluntary). This modification replaces the absolute moral ban on paternalism with a sliding scale of gradual acceptance.¹³ The resulting partial invalidity must be distinguished from relative weakness. If the partial invalidation only means that personal good reasons

12] Another possibility is that there is perfect agreement on the strength of the reasons involved and that these reasons determine that the person should try to kill herself and that we should try to stop her. Such agreement may entail that stopping is not liberty-limiting and so not contrary to anti-paternalism. Under perfect voluntariness, this is perhaps only possible in group cases where there are coordination problems (see Arneson (1980, 471) on enforcing a general preference through prohibition).

13] Technically, gradual acceptance is equivalent to discounting with a variable discount factor.

are not very strong relative to other reasons, then paternalism is in fact fully accepted over the whole range of partial invalidation.

Just like indeterminacy, partial invalidity would not affect the arguments in the previous sections, since the voluntariness in those cases is very high or maximal, and so the validity of the personal good reasons very low. Jumps in justifiability, however, would be replaced with a sliding scale, and so stopping in *The Suicide* would be more justified the lower the degree of voluntariness. I therefore recommend this moderation to anti-paternalists, though it comes at the price of increased complexity.

In conclusion, absolute anti-paternalism gives rise to peculiar jumps in justifiability such that one action may be overwhelmingly justified while another, very similar action, is overwhelmingly unjustified. This is a strange moral landscape. Indeterminacy does not make the doctrine any more appealing. Reformulation of anti-paternalism as gradual invalidation of personal good reasons avoids the jumps, at the price of greater acceptance of personal good reasons and greater complexity. Anti-paternalism should therefore be rejected, or, possibly, modified.

VIII. CONCLUSION

I have presented four arguments against anti-paternalism understood as a doctrine of invalidation, regulating the influence of reasons on the moral status of actions. The analysis of anti-paternalism in terms of invalidation hopefully has some merit independently of the arguments.

The fourth argument shows that in order to avoid peculiar jumps in justifiability, anti-paternalism must be reformulated as a doctrine of partial and gradual invalidation. This weakening of the doctrine will alienate many traditional anti-paternalists who are principally opposed to counting personal good reasons as valid reasons for interferences with voluntary action. The weakening, furthermore, does not make the doctrine any more resilient to the other three arguments.

The third argument shows that anti-paternalism leaves us at a loss when considering interferences with insufficiently voluntary choice. Presumably such interferences should be evaluated by simply weighing reasons for and against. If that is so, however, it is unclear why sufficiently voluntary choice should be treated any differently.

The first and second arguments show that anti-paternalism sometimes entails morally wrong conclusions. In particular, the first argument shows that the doctrine entails that we should allow very voluntary choice to cause severe harm to self, even if nothing much is at stake in terms of liberal values. The second argument shows that the doctrine entails that we should let perfectly voluntary choice cause or risk severe harm to self and others, not because the agent rejects our reasons to interfere with her for her sake, but only because she rejects our reasons to avoid harms to others.

Taken together, I find that the four arguments amount to a strong case for abandoning the form of principled anti-paternalism defended by Joel Feinberg and implicitly accepted by many others. Very often we should let people make their own mistakes and suffer the

consequences. However, there is no moral principle that forbids limiting or interfering with a person's liberty for her good. In particular, there is no moral principle which makes reasons which concern central human values such as health and happiness invalid when they conflict with liberty reasons.

Sufficient voluntariness is a factor in many other contexts than paternalism. A claim that a person has a right to something often means that she may have or do this thing on the condition that her choice is sufficiently voluntary. The right to marry freely, to vote, to enter contracts – these rights are arguably conditional on voluntariness. Annuling or failing to accept a marriage, a vote or a contract does not violate a right if the parties concerned did not act voluntarily. We may want to annul or refuse to accept these things for personal good reasons, but also for the good of other people, or for whatever other reason. The arguments against anti-paternalism can therefore rather straightforwardly be used as prototypes for arguments against most any doctrine of invalidation. Though not as straightforwardly, the arguments may also inspire arguments against other forms of influence-regulation on moral grounds, such as side constraints and lexical ordering. I believe all morally based influence-regulation should be rejected, but have not argued for this more ambitious claim.

To oppose influence-regulating doctrines is not to deny that there are intricate relationships between reasons. Things of value may be empirically related in the sense that actions tend to affect many such things simultaneously. Things of value may also be conceptually related – value may for example be disjunctive in the sense that it is of value that one of several things happen. Rather than formulate influence-regulating doctrines, we should investigate the empirical and conceptual relationships between values and so between reasons.

When Isaiah Berlin tells us that “[t]he extent of a man's negative freedom is, as it were, a function of what doors, and how many, are open to him; upon what prospects they open; and how open they are” (2002 [1969], 41); when Amartya Sen develops his concept of freedom as capability (1992, chapter three) or when Joseph Raz develops his ideal of autonomy (1986); when Mill develops the notion of individuality (1991 [1859], chapter III), and even when he briefly states that “[t]he only freedom which deserves the name, is that of pursuing our own good in our own way” (17) – in all these cases we see significant contributions to our system of values, to our views on what is important in life. When, on the other hand, Mill tells us that “the sole end for which mankind are warranted, individually or collectively, in interfering with the liberty of action of any of their number, is self-protection” (14), we are left confused as to what may then be the value of liberty and what other values there may be, that they should be related in this way.¹⁴

kalle.grill@yahoo.com

[14] This article has been long in the writing. I am sure I have some unrecognized debts. For very helpful comments, I wish to thank especially Sara Belfrage, Alon Harel and Lars Lindblom. Two anonymous reviewers for other journals exposed several weaknesses as well as the many ways in which my argument could be misunderstood. Hopefully it is now more transparent.

REFERENCES

- Arneson, Richard. 1980. Mill versus Paternalism. *Ethics* 90: 470-89.
- . 2005. Joel Feinberg and the Justification of Hard Paternalism. *Legal Theory* 11: 259-84.
- Beauchamp, Tom L. and Childress, James F. 2001. *Principles of Biomedical Ethics* (5th edition). Oxford: Oxford University Press.
- Berlin, Isaiah. 2002 [1969]. Five Essays on Liberty: "Introduction." In *Liberty*. Oxford: Oxford University Press.
- Broome, John. 2004. Reasons. In *Reason and Value: Themes from the Moral Philosophy of Joseph Raz*, eds. Jay Wallace et. al., 28-55. Oxford: Oxford University Press.
- Day, J.P. 1970. On Liberty and the Real Will. *Philosophy* 45: 177-92.
- De Marneffe, Peter. 2006. Avoiding Paternalism. *Philosophy and Public Affairs* 34: 68-94.
- Dworkin, Gerald. 1972. Paternalism. *The Monist* 56: 64-84.
- Feinberg, Joel. 1984. *Harm to Others*. Oxford: Oxford University Press.
- . 1986. *Harm to Self*. Oxford: Oxford University Press.
- Gert, Joshua. 2007. Normative Strength and the Balance of Reasons. *Philosophical Review* 116: 533-62.
- Griffin, James. 1986. *Wellbeing - its Meaning, Measurement and Moral Importance*. Oxford: Oxford University Press.
- Grill, Kalle. 2007. The Normative Core of Paternalism. *Res Publica* 13: 441-58.
- Groarke, Louis. 2002. Paternalism and Egregious Harm. *Public Affairs Quarterly* 16: 203-30.
- Kamm, F.M. 2006. *Intricate Ethics*. Oxford: Oxford University Press.
- Kleinig, John. 1983. *Paternalism*. Manchester: Manchester University Press.
- Mill, John Stuart. 1991[1859]. *On Liberty*. In *On Liberty and Other Essays*. Oxford: Oxford University Press.
- Nozick, Robert. 1974. *Anarchy, State, and Utopia*. Malden MA: Basic Books.
- Parfit, Derek. Manuscript. *Climbing the Mountain*.
- Rawls, John. 1971. *A Theory of Justice*. Cambridge MA: Harvard University Press.
- Raz, Joseph. 1986. *The Morality of Freedom*. Oxford: Oxford University Press.
- . 1990[1975]. *Practical Reason and Norms* (2d edition). Princeton: Princeton University Press.
- Scanlon, Thomas. 1998. *What We Owe to Each Other*. Cambridge MA: Harvard University Press.
- Scoccia, Danny. 2008. In Defense of Hard Paternalism. *Law and Philosophy* 27: 351-81.
- Sen, Amartya. 1992. *Inequality Re-examined*. Oxford: Oxford University Press.
- Sunstein, Cass R. and Thaler, Richard H. 2003. Libertarian Paternalism Is Not an Oxymoron. *The University of Chicago Law Review* 70: 1159-202.
- Van de Veer, Donald. 1986. *Paternalistic Intervention*. Princeton: Princeton University Press.