

# Paternalism

*Kalle Grill*

## Glossary

**Hard Paternalism** Submitting to paternalism a person whose choice or action is known to be sufficiently voluntary to be genuinely hers.

**Interference** In this context: An action that may involve paternalism because it diminishes or fails to respect some liberal value, typically by limiting liberty or violating autonomy.

**Libertarian Paternalism** Designing or mildly affecting choice situations in order to promote the chooser's good.

**Moralism 1)** Interference with a person aimed at protecting or promoting morality. **2)**

Invoking the promotion or protection of morality as a reason for interference with a person.

**Paternalism 1)** Interference with a person aimed at protecting or promoting her good. **2)**

Invoking the protection or promotion of the good of a person as a reason for interference with her.

**Soft Paternalism** Submitting to paternalism a person whose choice or action is known not to be sufficiently voluntary to be genuinely hers. Sometimes also interference with a person aimed at establishing to what extent her choice or action is voluntary.

**PATERNALISM** is a theoretical concept. It means, roughly, interference with a person's liberty or autonomy for her good. The concept is much disputed, as will be reflected in this entry. Paternalism is opposed by the liberal tradition, with an early and significant criticism in John Stuart Mill's *On Liberty* (1859).

The Latin root "pater" suggests paternalism is being or behaving like a father. However, paternalism does not refer to all behavior that is typical of fathers, nor even behavior that is typical of fathers towards their children, but specifically behavior that involves some form of restraint and some form of benevolence. The word "paternalism" gained usage in the late nineteenth century. It was then mostly used to refer to more or less strategic benevolence in hierarchical relationships, such as that between king and subject, factory owner and worker, and owner and slave (revealing a rather pessimistic view of fatherhood). This use remains to some extent today. The word is also used pejoratively for government activity in general and in particular for expansions of government activity.

In contemporary moral philosophy, "paternalism" refers to benevolent restraint, regardless of power structures. In that context, no gender distinction is intended by the term,

which could, if we were more open to terminological change, be replaced with “parentalism”. The etymological connection with parenthood and hierarchical relationships may explain why paternalism is sometimes thought to involve an attitude of superiority, signaling that the person acted towards does not know her own good well enough, or is unable to protect or promote it satisfactorily. Actual psychological attitudes of this sort are sometimes taken to define paternalistic behavior. Paternalism can also be taken to manifest such attitudes regardless of the actual state of mind of the paternalist. However, the central aspect of paternalism is the combination of an action which interferes with some person and a reason for that action which concerns the same person’s good, regardless of attitudes.

## **I. Historical overview**

Paternalism has some historical connection with patriarchic theories of political power. In the seventeenth century, Robert Filmer defended the divine right of kings to rule by invoking the “right of fatherhood” which, he claimed, had passed down from Adam to the kings and princes of his own time. This theory was famously criticized by John Locke by invoking individuals’ natural right to life, liberty and property, independently of any government. Liberal philosophy from Locke to John Stuart Mill emphasized the limited right of the state to interfere with individuals. In Mill’s mid-nineteenth century work, individual liberty is defended not mainly against monarchs, but explicitly against democratic government, as well as against society more generally.

Mill’s position in *On Liberty* is that people should be free to do as they like so long as their behavior does not harm others. People may be harmed not only by action but also by inaction. However, in such matters that concern only a person herself and consenting others, society has no right to intervene. Mill is mainly concerned with social as opposed to individual paternalism, though not only with the law but also with public opinion, which he argues can be as invasive as political oppression.

Mill’s arguments are in part practical. He claims that every person knows best what is in her interest and is most concerned to promote that interest. He argues that making one’s own choices tends to develop one’s character. He calls for experiments in living, which enrich our collective view of the possible and so promote all our well-being. However, there is also a perfectionist streak in Mill. He embraces the values of development and individuality, which he sees as fundamental aspects of well-being. Underpinned by these value commitments, Mill’s liberty principle is not easily set aside by practical circumstances.

The contemporary debate on paternalism was ignited by the debate between Patrick Devlin and H.L.H. Hart following the 1957 report of the U.K. Wolfenden committee, recommending that homosexual behavior between consenting adults in private should no longer be a criminal offence. Devlin was critical of this recommendation, arguing that homosexual acts harm the moral fabric of society and can in the long run lead to its disintegration. Hart, supporting the recommendation, argues that Devlin's unacceptable legal moralism must be distinguished from acceptable legal paternalism. In defending paternalism, Hart argues that Mill had been overly optimistic in assuming that individuals are the best judges of their own interest and act accordingly. He points to inadequate attention to consequences, transitory desires, inner psychological compulsion and subtle manipulation by others. Such obstacles to fully voluntary choice have since been at the heart of discussions on paternalism.

Following the Hart-Devlin debate in the 1960s, the 1970s and even more the 1980s saw considerable philosophical debate both on the nature of paternalism and on its moral status. Among contemporary works, Joel Feinberg's *Harm to Self* (1986) is arguably the most influential. Feinberg defends a Millian anti-paternalism but develops the restriction to competent and informed agents that Mill only hints at. He calls the resulting principle "soft paternalism", as it allows benevolent restriction of insufficiently voluntary choice (see section Voluntariness).

Having caught its breath after the outburst of articles and books on the topic in the 1980s, the discussion on paternalism has again intensified over the last decade. Philosophers such as Richard Arneson and Peter de Marneffe have questioned the prevailing and, in their view, extreme resistance to all forms of benevolent restraint of voluntary choice. Meanwhile, drawing on research in behavioral economics, the economist Richard Thaler and the legal scholar Cass Sunstein have defended what they call "libertarian paternalism" – promoting healthy choice by conscious design of choice situations, allegedly without limiting liberty. More typical libertarians of course continue to oppose all forms of government benevolence. Liberals of all types tend to be anti-paternalists in the sense that they hold paternalism to be at least *prima facie* wrong – that is normally wrong, though perhaps acceptable in some circumstances.

## **II. Normative and non-normative definitions**

Philology aside, we are interested in paternalism because of its normative properties, most typically the claim that paternalism is always *prima facie* wrong, or, very similarly, that the

fact that an action involves paternalism is always a *pro tanto* reason against performing that action or judging it to be justified. However, we may or may not define paternalism in normative terms. Non-normative definitions must specify such terms as “interference” and “good” without using normative terms. Normative definitions can allow that these terms are specified by what exactly is, allegedly, *prima facie* wrong with paternalism. Philosophers disagree on what is best.

A non-normative definition has the advantage that we may agree on what paternalism is without necessarily agreeing on its moral status. This is an advantage if we want to use the term “paternalism” to describe the world without getting into normative questions. Once defined, we can then argue about the moral status of paternalism. Since a non-normative definition does not track moral sentiments, it is very likely that it will exclude some actions which are opposed on typically anti-paternalist grounds, and include some actions which are not opposed on such grounds. As a result, argument about the moral status of paternalism will likely be futile.

We get a normative definition if we take the anti-paternalist opposition as our starting point and define paternalism as that which this opposition is aimed at. The advantage of this approach is that we may avoid complex non-normative conceptual issues and engage directly the more pressing matter of what exactly is claimed to be *prima facie* wrong and whether or not this claim is convincing.

### **III. Paternalistic actions and paternalistic reasons**

Normally, the label paternalistic is attached to actions, including government actions, or policies. We think of laws, medical practices or private actions as being potentially paternalistic. This is also how philosophers normally define the concept. Definitions typically include:

- 1) an *interference condition* which delimits the kind of action that may be paternalistic, most often excluding non-intrusive actions such as greeting someone in the street (when this is not a sign to fellow paternalists to capture the person and force her to be more prudent).
- 2) a *consent condition* which limits paternalistic actions to such actions as have not been consented to – excluding actions that are performed in response to explicit consent, and possibly also tacit and inferred consent.

- 3) a *benevolence condition* which limits paternalistic actions to such actions as are motivated, and perhaps also justified, by the good of the person(s) interfered with.

Very often, there is also:

- 4) a *superiority condition* which restricts paternalism to such actions as are performed by an agent who considers herself in some way superior to the person(s) interfered with.

On some accounts, interference with substantially non-voluntary behavior cannot be paternalistic. Most often, however, such interference (when benevolent and not consented to) is called “soft paternalism”. For this and other reasons, voluntariness should not be considered a necessary condition of paternalistic actions. Voluntariness will be further discussed below (see section Voluntariness).

The four conditions are specified in different ways by different authors, giving rise to a plethora of definitions. The three major conditions 1-3 will all be discussed in the following three sections, and superiority will be discussed briefly in the section Anti-paternalism. In the short definitions of paternalism in the Glossary, condition 1 and 2 have been integrated into the single concept of interference and condition 4 has been omitted. There are two definitions of paternalism (and moralism) in the Glossary because there is an alternative to attaching the label paternalistic to actions.

The standard approach is to take for granted that what should be defined is paternalistic action. This approach is convenient because we are often interested in judging the moral status of actions. However, some philosophers claim that paternalism has more to do with reasons than with actions. They point to the fact that there may be multiple rationales for any given action, making it difficult to say when an action is motivated or justified by the protection or promotion of a person’s good. For example, one reason for prohibiting drunk driving is that it will reduce risks for drunk drivers, a typically paternalistic rationale. However, another important reason for that policy is that it will reduce risks to other road users, a typically non-paternalistic rationale. It may therefore be hard to tell if the policy as such is paternalistic.

One possible reaction to the prevalence of multiple rationales is to define paternalism not in terms of actions but instead in terms of the invocation of reasons for actions. This leads to the second, alternative definition in the glossary. On such more reason-focused definitions, condition 3 defines a type of reason and conditions 1 & 2 (& 4) define a

type of action. Paternalism is the invocation of such reasons for such actions. In other words, with this move the components remain the same but the way they are related to one another changes. Opposition to paternalism on reason-focused accounts is not opposition to any particular action or policy, but opposition to counting certain reasons for an action or policy as valid or good reasons. I will say more on this matter below, **see** section Benevolence.

#### **IV. Interference**

Paternalism can be described on a more general level as a value conflict between some liberty value and some other value, where both values concern the same person or group. In the typical, individual case, limiting one person's liberty promotes some other value for her, such as her physical, economical or moral well-being. As Mill recognized, we may certainly try to affect people, for their good, by non-interfering means such as reasoned argument (in most circumstances). Indeed, Mill had no problem with sin taxes either, as long as the revenue was needed for government essentials. Why then may we not promote people's good by slightly more interfering means? Defining interference, and so the limits of benevolence, is an unavoidable and difficult task for anyone who takes a general stand on paternalism.

#### Liberty and Autonomy

Liberty is most typically understood as the availability of an adequate range of options. Limitation of liberty can then be understood as the restriction of this range by the exclusion of some alternatives. This view fits well with typical cases of paternalism such as the prohibition of certain activities, for example recreational drug use or driving without a seatbelt. However, the view would seem to exclude other typical cases, such as mercy killing and treating an unconscious person in conflict with her preferences or values, for example by transfusion of blood to an unconscious Jehovah's witness (whose religion forbids such procedures). If we understand limiting liberty as diminishing choice, therefore, we must either accept that such actions are not paternalistic, or accommodate them by somehow extending the meaning of "choice".

Autonomy is most typically understood as self-determination, and so violation of autonomy may, in the context of paternalism, be understood as substitution of a person's judgment or agency, where the paternalist judges or acts in place of the subject. Substitution of judgment is the more common in the literature, in the sense of deciding for the subject. Substitution of agency can take place when the subject prefers some option but fails to act on

this judgment for some reason, such as lack of determination or weakness of will. The substitution strategy can accommodate cases where the subject is affected in other ways than having her options restricted. On the other hand, it may seem to exclude cases where some, perhaps marginal, options are restricted, but the person is left to decide and act among the remainder. Substitution must therefore be understood broadly to include partial restrictions which do not amount to full substitution of judgment in the normal sense of the term. Alternatively, interference can be defined disjunctively as limitation of liberty or violation of autonomy.

The restriction of choice and the substitution of judgment (or agency) strategies together account for many proposed specifications of interference in the literature. However, they share a common weakness. People may well wish to choose and make judgments concerning things that principally concern others and so are not under their legitimate control. Both strategies must therefore presuppose that the restriction or substitution takes place within an area of legitimate control where the person's choice, judgment or agency should be decisive. Only within this area can restriction or substitution be interference. The so-called private sphere is sometimes thought to be such an area. The private sphere is supposedly that area of life or conduct which affects only the subject herself and consenting others. However, almost any part of a person's life has some effect on non-consenting others. Moreover, the mere fact that some conduct does not affect non-consenting others does not imply that it should not be restricted in order to protect the subject herself, who is indeed affected. The delimitation of an area of legitimate control is therefore largely an unsolved puzzle. In the face of this puzzle, one option is to refer the matter to some more general moral theory and so strip paternalism of independent moral importance. For example, paternalism could be understood as wronging someone for her good, whatever wronging is according to the correct moral theory. Another option is to retreat to the general liberal notions of interference with liberty or violation of autonomy, vague as they are, or look to other specifications of these notions.

### Libertarian Paternalism

In recent discussions of so-called libertarian paternalism, the interference condition has allegedly been lifted. It is not interference, the thinking goes, to adjust a choice situation so that it promotes the chooser's interests, as long as she is free to make her own choice in that situation. In response, it may be thought that choice situations should support chooser preference. However, people very often have no preference independently of the choice

situations they face. The way information is presented, the context in which the decision is made and what the default option is – these things tend to have a very strong influence on choice, even when all alternatives are readily available. These effects have been noted in preferences over everything from medical treatment to retirement savings plans.

People may possibly have preferences over choice situations, but the same indeterminacy may just as well reappear on this meta-level. Very often, therefore, neither preferences over alternatives nor preferences over the way alternatives are presented can determine what choice situations should look like. Even when people do have well defined preferences over choice situations, it is not clear that it would be an interference if these preferences are not treated as decisive, especially so since choice situations must often have the same structure for everyone in a certain group (patients at the hospital, employees at the company, customers at the cafeteria). I will return to this aspect of group cases below, see section Legal Paternalism. If libertarian paternalism is truly non-interfering, it is not a form of paternalism in the normal sense of the word. Paternalism should be distinguished from the wider notion of benevolence.

The rationale for libertarian paternalism is that it is better to promote good choice than to leave matters to chance. There may of course be other reasons to design choice situations in certain ways, such as the designer's self-interest. When there are no conflicting interests, however, truly libertarian paternalism can be opposed only by citing side-effects, such as the standard argument that power corrupts, in this case the power to design choice situations. Sometimes, libertarian paternalism is supposed to allow not only that necessary features of choice situations, such as setting the default option, are manipulated or designed, but also that small or even substantial costs are imposed on choosing certain alternatives. This, however, clearly approaches standard, interfering paternalism, since there is no clear distinction between imposition of costs and outright prevention.

## **V. Consent**

If a person gives her genuine consent to some action that interferes with her for her good, she is not a target of paternalism. For example, I can consent to medical procedures that would clearly have been intrusive where it not for my consent. Forcing (old or confused) people to take their medicine against their will is paternalistic, but even cutting a person open and removing some of his organs may not be, if he has given his consent. This condition may be considered an instance of the old Roman formula *Volenti non fit injuria* – to a willing person, no harm is done. Explicit consent should not be confused with different forms of hypothetical

consent, which are not consent in a strict sense. However, consent need not always be given by uttering the words “I consent”, or even be verbal. In some circumstances, silent and tacit consent can be inferred from the circumstances. The consent, however, must be genuine, that is voluntary. The same standard can be applied to the voluntariness of consent as to the voluntariness of actions (see section Voluntariness below).

Normally, there is a delay between the giving of consent and the action which is consented to. This raises the possibility that the person may change her mind. In research contexts, for example, it is established practice that subjects have the right to withdraw from a study at any time. On the other hand, we often have reason to restrict our future options, for example to help ourselves resist temptation. In an often cited example of consent, Odysseus asked his sailors to tie him to the mast when they approached the Sirens’ island, in order that he would hear their singing without risking his ship or his life trying to reach the island. Once the singing started, Odysseus clearly wanted his sailors to untie him. However, Odysseus had foreseen his own reaction and specifically instructed the sailors not to listen to his future self. The sailors’ refusal to untie Odysseus is therefore not considered a case of paternalism, or if it is a special case – self-paternalism. It is an open and difficult question, however, under what circumstances and in what timeframe a person may restrain her future self, and when others may help him to do so.

A related question concerns the possible scope of consent – whether consent must be specific to certain situations or whether it can be more open-ended. It seems we can consent to someone else directing parts of our life for us, for example by hiring a professional to manage our finances. Such consent can normally be revoked at any time. However, consent can be both wide in scope and future-oriented. For example, we can appoint someone to be the guardian of our interests if we become incapacitated and later be unable to revoke that appointment. If such consent is valid, our guardian can interfere with us for our good without engaging in paternalism.

When a person becomes unconscious or otherwise incapacitated and has not made prior arrangements, we face another sort of difficulty. No consent can be given, but we can make educated guesses concerning what the person would have chosen if she were capable, based on the available facts, including the person’s earlier decisions and behavior. Such an estimated consent is sometimes called “individual hypothetical consent” and is, though hypothetical and not strictly a form of consent, likely to be in harmony with the person’s values. This should be distinguished from “reasonable man hypothetical consent”,

which estimates what a rational and informed person would choose in the person's circumstances, with no regard for the person's specific character and preference.

Both kinds of hypothetical consent are sometimes argued to preclude paternalism even for conscious, mature and healthy individuals, who merely have got the facts wrong by human error or reasonable mistake. This is for example Mill's solution to the apparently paternalistic and acceptable action of restraining someone who is about to severely harm himself. If someone risks harm by unknowingly crossing an unsafe bridge, Mill reasons, it is not even an interference to physically restrain her. This makes sense, since there are many cases where people would simply be grateful for being prevented from making dangerous mistakes. However, invoking hypothetical consent whenever someone does not anticipate the full consequence of her action would open the door to a wide range of interferences, certainly more so than Mill would have been prepared to accept.

So called "subsequent consent" is not a form of consent either, but rather refers to expressing a positive attitude to something that has already happened. Unless a person has undergone great changes, however, subsequent consent is an indication, after the fact, that the person would have consented had she been able. Proxy consent is not consent by the person in question, but by another person who, perhaps temporarily, acts as the person's guardian, without her consent. Parents are generally allowed to give proxy consent for their children, though only within boundaries.

Depending on how interference is specified, it may be that there is no such thing as consenting to interference. In other words, just as consent precludes harm in legal contexts, it may preclude interference. If so, the consent condition and the interference condition are integrated, with consent one factor determining whether or not an action is interfering. Indeed, a possible specification of interference is "to affect matters that are under some person's legitimate control, without her consent".

## **VI. Benevolence**

Paternalism is always benevolent in some sense. Normally, this benevolence is directed at the person who is interfered with or affected. In so called "impure" cases of paternalism, however, the interference is primarily directed at someone other than the benefactor. For example, when a prohibition forbids the sale of recreational drugs, the supposed benefactor is the buyer, but the primary target for interference is the seller. Even in impure cases, however, there is arguably some interference with the benefactor, who is prevented from doing what she wants to do or getting what she wants to get, though indirectly. Interfering with

consensual interactions must be distinguished from interfering with one person or group in order to prevent harm or provide benefit to third parties – preventing a robbery for the good of the victim is not a case of impure paternalism.

On most definitions of paternalism, the benevolence is supposed to be first and foremost a psychological motive for the interference. This means that actions can involve paternalism even if they do not in fact benefit the person interfered with in any way, as long as the agent believes that such benefit is forthcoming. The focus is on the agent; assessments of the situation by third parties cannot be paternalistic (unless they amount to interference in their own right). Alternatively, the benevolence may lie on the level of justification. This means that actions can be evaluated independently of motives. If so, policies and laws can involve paternalism in force of their (alleged and/or actual) rationale, regardless of why exactly they were first instigated and enacted. Arguably, both motives and justifications can be paternalistic and so definitions should ideally accommodate both kinds of reasons.

### Mixed cases

Benevolence may or may not be the only reason for an interfering action. When benevolence is only part of the rationale, we have a “mixed case”. As mentioned above (see section Paternalistic actions and paternalistic reasons), paternalism may be defined either in terms of actions or in terms of reasons for action. Mixed cases pose no theoretical difficulty for reason-focused accounts, which are partly proposed as a way of dealing with such cases. If paternalism is the invocation of benevolence as a reason for interference, this is quite independent of whether or not other reasons are also invoked for the same interference. The interference as such cannot be paternalistic on reason-focused accounts.

Action-focused accounts, on the other hand, must somehow set a standard for which mixed cases, if any, qualify as paternalism. There are two main strategies. The first is to draw a line for how important the benevolent reason must be in the total set of reasons for an action to be paternalistic. Benevolence may most typically be required to be the only reason, the main reason or a necessary reason in some sense. The second strategy is to make the benevolence condition apply gradually so that interferences are paternalistic to the degree that their rationale is benevolence. This may be understood either relative to the strength of other reasons for the interference, or independently in terms of the absolute strength of the benevolence reason. The specification of paternalistic actions in mixed cases is further complicated by the fact that reasons may be either motives or justifications.

In other words, the prevalence of mixed cases make fully developed action-focused accounts very complex. It is difficult, perhaps even impossible, to take proper account of the moral importance of all other kinds of reasons for action, while opposing certain actions because they are partly motivated and/or justified by benevolence. Reason-focused accounts, on the other hand, though well suited to deal with mixed cases, are complex in their own right, with their refusal to label actions as either paternalistic or non-paternalistic. Perhaps this is the reason why most authors prefer to stick with action-focused accounts and deal only with unmixed cases.

### Theories of the good

Strictly speaking, paternalism is independent of conflicting theories of the good. Paternalism is interfering with a person for her good, regardless of what that good consists in exactly. However, which theory of the good one adopts affects how much room there is for paternalism. On subjectivist theories of the good, something can only be good for a person if she finds it good. If such a theory is correct, benevolent interference with a person can only be efficient in cases where she fails to bring about what she finds good or even allow that others bring it about. Such cases are not that uncommon, however. For example, a person may agree that coercive treatment would do her more good than harm and yet decline to consent to such treatment, perhaps out of fear, or because she is more concerned to spare her family the expense. Furthermore, a person may be mistaken about what means would contribute to her ends, and so resist taking them. On less subjectivist theories of the good, something can be good for a person even though she does not find it so, leaving even more room for paternalism.

In the literature, paternalism is often associated with objectivist theories of the good. This may be because justifying an interference with a person aimed at promoting some end which she does not endorse requires both an objectivist theory of the good and acceptance of paternalism. In that sense, the two are bedfellows. As we have seen, however, paternalism is consistent with subjectivist theories of the good. It does not require disregard for a person's view of her good.

A presumptive paternalist may of course fail to do good because she is confused or mistaken about the benefits of her interference. The mistake may be either practical – her interference will not bring about what she aims to bring about, or normative – what her interference brings about is not good. In either case, her interference is pointless and trivially undesirable.

## Moralism and types of good

The good that benevolent reasons concern or refer to can be of any kind. Most typically, the good is physical health or survival. Paternalism may also aim to promote people's finances, for example by mandatory retirement savings, or to protect a person's autonomy or promote her future liberty, for example by coercive drug rehabilitation or intrusive dissemination of information. A category of special interest is moral and religious paternalism, which aims to protect or promote some person's soul or virtue. For example, missionaries have often forced people to take part in religious services in order to save their souls from eternal damnation, and prohibitions on relatively harmless recreational drugs may be aimed at protecting people's good character from deteriorating into passivity or excessive creativity.

Moral paternalism should be distinguished from moralism, which is interference aimed at protecting or promoting morality in general, independently of any individual's virtue. Cursing may be thought to blemish a person's soul or character, but it may also be thought to make the world an uglier place. Less benevolent missionaries may try to save the heathens not for their own sakes, but for the greater glory of God and his kingdom on earth. The government may prohibit gay soldiers from announcing their sexual preference not for their own sakes, but in service to the nation's pride and purity. As shown by the latter example, arguments for moralism may overlap with arguments for the preservation or protection of society. If (implausibly) the essence of some society depends for its preservation on some set of general conduct or opinion, a shift away from that conduct or opinion would spell the end of that society. More practically, social trust is a common asset and if certain behavior damages this trust, perhaps without directly harming any one individual, interference with such behavior may be motivated by the protection of all in some sense, rather than by the preservation of morality as such. Such potential overlap can make it hard to distinguish moralism from paternalism in practice, though they are quite distinct in theory.

## **VII. Voluntariness**

Some behavior is so much an effect of forces external to the person that it can be questioned whether it is in any real sense her behavior – it is substantially involuntary. Reflexes and sleep-walking are obvious cases, but also actions performed under heavy intoxication, complete ignorance or extreme time pressure may qualify. Some human beings cannot act voluntarily at all, such as infants and some severely developmentally disabled. Interference

with substantially involuntary behavior is often called “soft paternalism” and is generally considered acceptable. (To complicate matters, some authors, especially economists, use “soft paternalism” quite differently to refer to paternalism where the interference is mild.) Naturally, interference with involuntary behavior can be morally problematic if it is harmful or unkind, but, the thinking goes, not simply because it is benevolent, not consented to, and perhaps done with a (very reasonable) sense of superiority.

Depending on one’s preferred definition of paternalism, soft paternalism may not be a form of paternalism at all. “Hard paternalism”, in contrast, is interference with voluntary choice and so unquestionably paternalism. Voluntariness comes in degrees and depends on such factors as how informed the agent is, how well she reasons about her available alternatives, how free she is from manipulation and external pressure, how free she is from internal compulsion, and how calm and collected she is. Most behavior of course is neither completely involuntary nor fully voluntary. The soft – hard categorization is normally supposed to be exhaustive, however, and so a line must be drawn between hard and soft at some degree of voluntariness. Because of the widespread acceptance of soft paternalism and rejection of hard paternalism, it is important where this line is drawn.

Very often, we do not know to what degree some behavior we observe is voluntary. Typically, if the behavior is harmful to the agent, it is considered acceptable to interfere with her in order to find out to what extent she knows what she is doing and whether she intends to do it. Such interference, therefore, is often considered a form of soft paternalism as well. As with degrees of voluntariness, lines must be drawn for what type and degree of uncertainty warrants such inquiring interference.

If it should be value-neutral, the distinction between soft and hard paternalism must be determined independently of the seeming reasonableness of the behavior under consideration. Naturally, some behavior is such that we automatically assume involuntariness, such as stepping into the road before an oncoming truck. To be sustained, however, the assumption of involuntariness must stand up to closer scrutiny. If we persist in classifying some behavior as involuntary or possibly involuntary, independently of the mental state and practical circumstance of the agent, but only in light of its undesirable consequences, then we are imposing our own standards of reasonableness on the agent. This may perhaps be morally permissible in some situations, but it is certainly paternalistic.

An agent’s degree of voluntariness naturally varies over situations. We know more about some things and we reason better in some situations. However, the threshold of sufficiently voluntary choice may either be rigid over situations or may depend on how much

is at stake. As an example of the latter, we may require a higher degree of voluntariness before we allow that giving up one's life is voluntary than we require before we allow that giving up one's blood for donation is voluntary. Indeed, the requirements in the former case may so strong that we can in practice never be sure that they are fulfilled.

In legal and medical contexts, voluntariness is often discussed in terms of (mental) competence or (decision-making) capacity. These are generally considered threshold concepts, a term that is natural in light of their practical implications – competent or capable persons are considered qualified to decide over their own treatment and their lives more generally. The threshold may or may not coincide with the threshold between soft and hard paternalism. Definitions of competence and capacity are normally expected to capture whether or not a patient or legal subject should be allowed to decide over her own affairs. In that sense, the concepts are normative. An alternative approach is to try to keep these concepts non-normative and allow that there may be other factors than the subject's degree of voluntariness which weigh on whether or not she should be allowed to make her own decisions.

The great importance often attributed to the soft – hard distinction depends on the assumption that interference with substantially involuntary behavior is not morally problematic, or only marginally so. This assumption can be questioned. I may for example resist interference with my private life even when I act contrary to my intentions, due to mistake, confusion or emotion, simply because of a preference against outside involvement. When the behavior is harmless or only trivially harmful the agent, such preferences are very reasonable and widespread. Moreover, behavior that is locally involuntary, so to speak, may still be globally voluntary in the sense that the agent considers it authentic and in line with her (foolhardy, careless or spontaneous) character.

### **VIII. Anti-paternalism**

Within the broad liberal tradition, paternalism is generally considered *prima facie* or *pro tanto* morally wrong. This is the anti-paternalist position or doctrine (though the term can also refer to the much stronger position that paternalism is always wrong all things considered). Anti-paternalism most obviously draws support from the fact that paternalism involves some kind of interference with individual choice or action (see section Interference above). This no doubt gives us a *pro tanto* reason against any action involving paternalism. However, paternalism is aimed at protecting or promoting individual good. If the interference fulfils this aim and we care about individual good (health, survival, finances, long term autonomy), this

should give us a *pro tanto* reason for the action. Indeed, we may even think that good intentions are good reasons independently of effects. At least in cases of effective paternalism, we would seem to stand before a moral conflict rather than a moral wrong. The anti-paternalist can respond to this observation in a number of ways.

The most straightforward defense of anti-paternalism is to accept that good-promotion gives us reason for action but insist that interference with (sufficiently voluntary) choice or action always gives us stronger reason against. This may seem reasonable given the liberal assumptions that people generally know their own interests and are the ones most capable to further them. However, there are plenty of cases where minor interference leads to enormous good, such as when a person is forced to step back from the edge of the cliff and is thereby saved from certain death, or when a population is forced to wear seatbelts and thousands of deaths are thereby prevented. It seems incredible that our (liberty) reasons against interference are stronger than our (survival) reasons for action in such cases. It is true that in many cases in which the interference is mild and the benefit large, the targeted behavior is not very voluntary. However, in other cases the person is relatively informed and rational, but only happens to make a mistake, or is inattentive or lazy, or is acting on a passing impulse.

Another way of defending anti-paternalism is to accept that promoting good may sometimes be more important than avoiding interference, but to claim that interfering with a person for her own good entails a special sort of insult that goes beyond the interference as such. The benevolent motive quite literally adds insult to injury, or rather to interference. This position implies that it is better to interfere with a person on a whim, for no special reason, than to interfere with her for her sake. This may seem implausible, especially for those who want to make room for care in ethics, but could be accepted by devoted individualists. Even so, the insult must be very large in order to underpin general anti-paternalism, since our reasons for interference can be very strong.

Perhaps what is insulting is not the benevolent interference as such, but the attitude displayed by the paternalist. On some definitions, as we have seen (see section Paternalistic actions and paternalistic reasons), the paternalist always considers herself in some way superior to the person she interferes with. This may seem insulting. However, if an interference would benefit a person, this must be because something is stopping the person from securing this same benefit on her own. The paternalist is in fact superior in the specific sense that she is more able to produce this benefit. For example, I may have brought a pack of cigarettes for the hike in the wild, failing to take this chance to quit smoking because it is too

hard for me. My friend, on the other hand, may have no trouble tossing the package into the mountain stream. In general, of course, it is not morally problematic to consider oneself more able than others in certain respects. The anti-paternalist must explain why it is especially problematic to consider oneself more able to promote a person's good than she herself is.

The most promising defense of anti-paternalism is arguably to point to our often strong reasons against interference and boost them with some empirical assumptions about our tendency to selfishness, corruption and incompetence when it comes to benefitting others, especially in institutional settings. Even though we can sometimes, in theory, benefit others by interfering with them, a large enough dose of pessimism will entail that whenever we try we will either fail or cause greater harm – by setting dangerous precedents or by discouraging individual responsibility or by strengthening existing power concentrations or something along these lines.

A final option, of course, is to reject general anti-paternalism and accept that benevolent interference with voluntary choice is sometimes justified, all things considered. Anti-paternalism is the doctrine that paternalism is always *prima facie* or *pro tanto* wrong. If this doctrine is false, liberty and autonomy can still be very important values, just not values which automatically trump other values. Given a strong commitment to liberal values, avoiding paternalism will very often be a sensible rule of thumb, though it can be set aside when the circumstances are grave enough and the means available to improve them efficient enough.

## **IX. Legal paternalism**

The liberal criticism is mainly focused on paternalism by the state. When such paternalism is backed up by the criminal law it may be called legal paternalism. Legal paternalism most typically consists in prohibitions, such as against driving a motorcycle without a helmet or using heroine. Economical policy such as taxing and subsidizing may also be paternalistic if it is interfering. Examples might include sin taxes on alcohol and cigarettes and subsidies for exercise. Legal paternalism is often impure in that it targets consensual interactions, for example drug sale or assisted suicide (on impure paternalism, **see** section Interference above).

Legal paternalism is special in several respects. The state is normally the supreme authority in terms of physical force within its domain, determining the basic rules by which a society functions. On the one hand, this makes paternalism by the state especially intrusive, since government policy can have far-reaching and comprehensive effects on what opportunities are available in society. On the other hand, the basic rules of a society partly

determine people's legitimate expectations and so limiting liberty on this basic level may not upset those expectations and may in that sense be especially non-intrusive. Either way, since the effects of state policy are so far-reaching, it is important that it tends to protect and promote people's good.

## Large numbers

Legal paternalism normally affects large numbers of people with no regard for their individual preferences and special circumstances. The state may have to limit the liberty of all in order to protect or promote the good of some. For example, the state may want to restrict the sale of potentially harmful recreational drugs for which some people develop a dangerous addiction (such as alcohol). Such cases involve not only weighing some people's liberty against their good, but also weighing the liberty of some against the good of others. If the rationale for a policy is to benefit certain people through interference with them (for example by limiting their access to alcohol) this is undoubtedly paternalistic. That other people are also interfered with, without further benefit, is simply a negative side-effect, whether or not it is severe enough to make the policy unjust all things considered.

People often welcome liberty-limiting policies because they help them secure important benefits. Examples include seat belt laws and mandatory retirement savings. If all those affected welcome a liberty-limiting policy, that policy arguably does not involve paternalism. More often, however, only some of those affected welcome the policy, making it non-paternalistic for them but paternalistic for the remainder. Such cases pose a difficult challenge for action-focused accounts of paternalism, which must be fine-tuned to determine whether or not the policy is, on the whole, paternalistic. The standard approach is to go with the motive of the policy-makers, judging a policy non-paternalistic if it is intended to benefit the consenters, but paternalistic if it is intended to benefit the non-consenters. This approach gives no guidance to policy-makers who intend to benefit all, or who simply wonder what their intention ought to be in this situation. Also, there remains the general problem with defining paternalism in terms of motives, namely that on such definitions a policy is non-paternalistic if the policy-makers' motive is not to benefit, but, for example, to further their own careers, even if the policy is accepted and preserved because it is generally considered to be justified by the benefits to those it interferes with.

On reason-focused accounts, what is paternalistic is invoking some people's good as a reason for interfering with them. If such reasons can be avoided, a policy can be enacted without paternalism. In the case of liberty-limiting policies that are welcomed by some but not by all, paternalism can be avoided if the benefits to the consenters are large enough to justify the interference with the non-consenters. At least this is true if the consenters consent out of self-interest. If they consent because of the benefits to the non-consenters, they join purposes with the policy-makers, arguably becoming paternalists themselves and further complicating the analysis.

Paternalism may also be avoided if a policy that interferes with a group in order to benefit that group produces the benefit for each member through interference with other members. This is typically true for policies that are not considered paternalistic, such as prohibitions on theft and assault – these core elements of the criminal law restrict the liberty of everyone, in everyone’s supposed best interest. More typically “paternalistic” policies may have the same structure – product safety regulation partly aims to protect people from being harmed by each other with dangerous products, such as cars or explosives.

Since legal paternalism involves weighing the interests of various groups, issues of fairness arise. Those who benefit most from a liberty-limiting policy are typically the bad choosers – those least able to make decisions beneficial for themselves in the absence of interference. People who are bad choosers in one area tend to be bad choosers in other areas. Therefore, avoiding benevolent legal interference will tend to favor the better off at the expense of the worse off.

## **X. Medical paternalism**

Paternalism is close at hand in medical settings, since caregivers are normally in a privileged position vis-à-vis their patients. Treatment choices are complex and patients’ decision-making capacity is often impaired. In contrast to an earlier acceptance of medical paternalism, advocacy of patient autonomy is now the norm. Sometimes this is taken to imply that patients should always be informed of available treatment options and encouraged to actively choose between them. However, patient autonomy may also be respected by allowing patients to choose not to get informed, but rather trust their caregiver or their families to decide for them.

There are good practical (consequentialist) reasons to respect patient autonomy. Like any person, a patient is most likely to know her own circumstances and preferences, including non health-related circumstances like her family responsibilities and financial situation, which the caregiver may not know anything about. Adding to the general instrumental and possibly intrinsic value of autonomous choice, patients have to some extent lost control over their own lives and bodies and so have an extra need to re-establish their sense of self-direction. Also, adherence and so effective treatment is promoted by some respect for patient autonomy. Furthermore, at least in modern, individualist societies, trust and respect in the relationship between caregiver and patient generally makes people more likely to seek treatment and more receptive to recommendations.

On the other hand, the medical context provides ample examples of how interference can be beneficial. Generally able people can refuse best treatment because they

are anxious or upset, because they mistrust their doctors, because they are misinformed or make mistakes, or simply because they fail to bring themselves to accept a risky or painful procedure.

Several ethical codes and declarations after the Second World War have emphasized the importance of informed consent, first to participation in research and later also to treatment. Early on, informed consent served mainly as protection against exploitation and misconduct. Over time, it became more and more integrated with respect for patient autonomy. Today, asking patients to give informed consent is standard procedure, but has been criticized for inefficiency, for failing to protect autonomy, and for inducing caregivers to shrink from their professional responsibility. The formal use of informed consent forms should be distinguished from actual consent by informed people, whether or not the former is intended to capture the latter.

Medical paternalism in institutional settings is much like legal paternalism in that it targets groups with little possibility for individual variation. However, medical paternalism is also a feature of personal relationships between individual caregivers and their patients. Relationships can be openly paternalistic in that the caregiver assumes the role of benevolent authority, issuing well-intended directives and expecting to be obeyed. Relationships can also be secretly paternalistic in that caregivers manipulate patients for their own good, for example by mixing medicine into food or drink. In both cases, an alternative is an equal and honest relationship in which caregiver and patient reason together about the available options and perhaps even make decisions together. What is the right approach in a certain situation depends on whether anti-paternalism is a correct doctrine, and on what is the value of health, of equality and of self-direction.

### **Further reading**

Archard, D. (1994). For our own good. *Australasian Journal of Philosophy* **72**, 283-93.

Arneson, R. (2005). Joel Feinberg and the justification of hard paternalism. *Legal Theory* **11**, 259-84.

Arneson, R. (1989). Paternalism, utility, and fairness. *Revue Internationale de Philosophie* **43**, 409-437.

De Marneffe, P. (2006). Avoiding paternalism. *Philosophy and Public Affairs* **34**, 68-94.

Dworkin, G. (1972). Paternalism. *Monist* **56**, 64-84.

Feinberg, J. (1986). *Harm to self*. Oxford: Oxford University Press.

- Gert, B. and Culver, C.M. (1976). Paternalistic behavior. *Philosophy and Public Affairs* **6**, 45-57.
- Grill, K. (2007). The normative core of paternalism. *Res Publica* **13**, 441-458.
- Hart, H.L.A. (1963). *Law, liberty and morality*. Oxford: Oxford University Press.
- Husak, D.N. (1981). Paternalism and autonomy. *Philosophy and Public Affairs* **10**, 27-46.
- Kleinig, J. (1983). *Paternalism*. Manchester: Manchester University Press.
- Mill, J.S. (1991 [1859]). On liberty. In: *On Liberty and Other Essays*. Oxford: Oxford University Press.
- Shiffrin, S. (2000). Paternalism, unconscionability doctrine, and accommodation. *Philosophy and Public Affairs* **29**, 205-250.
- Thaler, R.H. and Sunstein, C.R. (2003). Libertarian paternalism is not an oxymoron. *The University of Chicago Law Review* **70**, 1159-1202.
- Van de Veer, D. (1986). *Paternalistic interference*. Princeton: Princeton University Press.

### **Websites**

<http://plato.stanford.edu/entries/paternalism/>

<http://plato.stanford.edu/entries/mill-moral-political/>

### **Cross-references**

Advance directives; Autonomy; Euthanasia; Informed consent; Liberalism; Public Health Ethics; Right to Know and Right Not to Know.