



# Behavioral Paternalism

Christophe Salvat

► **To cite this version:**

Christophe Salvat. Behavioral Paternalism. *Revue de Philosophie Economique / Review of Economic Philosophy*, Vrin ; De Boeck, 2014, 15 (2), pp.109-30. <<http://www.cairn.info/revue-de-philosophie-economique-2014-2-page-109.htm>>. <10.3917/rpec.152.0109>. <halshs-00947142>

**HAL Id: halshs-00947142**

**<https://halshs.archives-ouvertes.fr/halshs-00947142>**

Submitted on 14 Feb 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Behavioral Paternalism

## Or the Possibility of Paternalism

Christophe Salvat<sup>1</sup>

CNRS, Triangle, ENS de Lyon

### Abstract:

In this article I address the question of the moral legitimacy of paternalism. Paternalism is considered illegitimate a) when it acts under false pretenses to satisfy the interest of the paternalist agent, b) when it violates the individual autonomy of the people interfered with and c) when it does not respect the equality between people by singling out those who are deemed unable to decide for themselves. Over the last decade a new type of paternalism has emerged thanks to the groundbreaking works of some behavioral economists. This new type of paternalism, that I call here Behavioral Paternalism (BP), has become popular through Sunstein and Thaler's Nudges theory and challenges the view that paternalism is unacceptable today. The aim of this paper is to assess its moral legitimacy (not exclusively focusing on the autonomy proviso). The results of my investigation can be summarized as follows. Though BP is usually acknowledged for its 'libertarian' character, it does not satisfy the conditions of what is considered, since Feinberg, as 'soft paternalism'. Nevertheless, BP has a strong point that has been underestimated by its partisans: it withstands the equality argument. Unlike traditional forms of paternalism, BP is not demeaning and does not ostracize any category of people.

---

<sup>1</sup> Early versions of this paper have been presented at the Collège International de Philosophie in Paris (January 2010), at the University of Lille (Institut d'Economie Politique, June 2012), at the University of Strasbourg (BETA Seminar, November 2012) and at the University of Cambridge (Faculty of Philosophy, January 2013). I am grateful to all participants for their insightful remarks. I would also like to thank the referees of *Revue de Philosophie Economique* for their helpful comments. Whilst this paper was under review some interesting works have been published on paternalism. I wish to mention in particular *Paternalism: Theory and Practice*, edited by Christian Coons and Michael Weber (Cambridge University Press, 2013), as I would certainly have made extensive use of it.

Lastly, BP can be proved genuinely altruistic. This, however, demands that one abandons Sunstein and Thaler's main assumptions.

**Key words:** paternalism, behaviorism, equality, autonomy

## 1. Introduction

The idea of paternalism has suffered dramatically since the decline of the nineteenth-century model of entrepreneurial capitalism. As a social practice, it is widely considered as a patronizing and morally inappropriate conduct to adopt for any individuals benefiting from a superior economic, social or political position. The philosophical debate over the legitimacy of paternalism, however, twice endeavored to challenge this view. A first wave of arguments in favor of paternalism dates from the 1970s<sup>2</sup>, a second only started a few years ago under the influence of major behaviorist thinkers. It emerges from all these contributions that modern paternalism no longer bears resemblance to social or religious charity. In the 1970s paternalism was essentially considered from a legal point of view (Feinberg, 1983, Feinberg, 1986, Dworkin, 1971) or, if not, from a medical point of view (Buchanan, 1978, Gert and Culver, 1976, Dworkin, 1983). The questioning, however, remains the same: under what conditions can paternalism be morally acceptable?

In the nineteenth-century entrepreneurial or social paternalism was often perceived as serving the interests of the dominant class under false pretenses. Paternalism was thought immoral because it was not genuinely altruistic. Since WWII legal paternalism has been criticized for being intrusive and coercive. Paternalism was then thought immoral because it was perceived as a threat to individual freedom. A few years later paternalism suffered another philosophical blow when it emerged that it was also violating individuals' basic right to equality (Dworkin, 1978, Anderson, 1999, Arneson, 2005). As one can see, the moral demands made to paternalism have shifted over time and conformed to new ideologies. But none disappeared altogether. It is still considered immoral for paternalism to be

---

<sup>2</sup> Most of the original contributions as well as a number of unpublished ones (including Dworkin's Second Thoughts) have been gathered in the volume edited by Rolf Sartorius in 1983, SARTORIUS, R.E. (1983), *Paternalism*. University of Minnesota Press.

self-interested, and ensuring that individuals ‘paternalized’ are treated with equal respect does not encroach upon their basic right to freedom. Theories defending paternalism have to therefore adjust to continually meet more demanding moral standards.

At the beginning of the new millennium, spurred on by the success of behavioral studies, new theories on paternalism took up the challenge (Jolls et al., 1998, Sunstein and Thaler, 2003, Thaler and Sunstein, 2008, Camerer et al., 2003, Loewenstein and Haisley, 2008). Sometimes called ‘libertarian’, ‘asymmetric’, ‘light’ or simply ‘new’ paternalism, these theories are all related to behavioral sciences. For simplicity’s sake, I shall refer to them as Behavioral Paternalism (BP).

The aim of this paper is to assess BP’s ‘moral robustness’ and to check to what extent the three moral provisos have been taken into account. This paper is divided into five sections. In the first section I present what I consider to be the three main moral requirements that paternalism has to meet today, i.e. individual freedom, respect of equality and genuine moral intentions. The second part explains what I call Behavioral Paternalism. The following three sections respectively deal with the three moral requirements presented in the first section.

## **2. Can paternalism be moral ?**

The first and most intuitive argument opposed to paternalism is the respect of individual freedom. Opponents of paternalism defend a right to autonomy or self-direction (VanDeVeer, 1986). It is naturally not an absolute right. The nature of governments is to interfere with an individual’s decision-making process. Liberals believe that it is legitimate to do so only when individual actions are other-regarding. They hence justify legal restrictions such as the obligation for an individual to wear a helmet or a seatbelt whilst driving a car or a motorbike by the benefit reaped by others (the cost born by the collectivity) rather than by the utility the individual should expect from the measure. This means that to comply with this critic, self-regarding justifications for paternalism must always be proved inadequate to back interventions or they must be able to be otherwise justified. In practice, however, it is rather difficult to ascertain that no self-regarding justification for paternalism can ever

be found or that other-regarding justifications will be sufficient (Marneffe, 2006). The question is therefore not whether paternalism is a legitimate practice but to what extent autonomy provides a justification (or not) for paternalism.

For Joel Feinberg (Feinberg, 1986), the violation of individual autonomy is the sole possible justification for anti-paternalism. Conversely autonomy can also be its sole possible justification. In a certain number of cases paternalistic interventions can thus be justified by the absence or insufficiency of individual autonomy. Four factors contribute, for Feinberg, to weaken the autonomous or voluntary quality of individual decisions. Those are:

1. Coercion or threat of coercion
2. Psychological dependency (drug, alcohol, hypnosis...)
3. Emotional ascendency (depression, anger, obsession...)
- 4a. Misinformation, deceit or wrong belief that misleads the agent relative to the consequences of his action (I.e. accidentally taking arsenic for salt)
- 4b. Misinformation or a wrong belief (produced by the individual inability to gather correct information) that eventually misleads him relative to the consequences of his action.

To remedy the totalitarian tendency of paternalism, Joel Feinberg proposes to vet its practice on the condition that it respects individual freedom. This ‘soft’ or ‘weak’ version of paternalism originates from Mill’s own ambiguity in his famous harm principle: ‘That principle is, 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. That the only purpose for which power can be rightfully exercised over any member of a civilized community, against his will, is to prevent harm to others. [...] He cannot rightfully be compelled to do or forbear because it will be better for him to do so, because it will make him happier, because in the opinion of others, to do so would be wise, or even right’ (Mill, 1989: 13). Due to the ambiguity of Mill’s phrasing (‘against his will’) – but also due to a

pervasive ambiguity throughout his work<sup>3</sup>— the discussion about the legitimacy of paternalism soon boiled down to a single argument: the degree of voluntariness in the action interfered with<sup>4</sup>. The notion of will can indeed either refer to a current (and temporary) desire or to a ‘true’ and rational will. From this information the nature (and hence the legitimacy) of paternalism could be inferred: a weak or soft paternalism meant to protect individuals against self-inflicting (but not fully voluntary) harm, while “in its extreme version”, a strong or hard paternalism was supposed “to guide them, whether they like it or not, toward their own good” (Feinberg 1983: 3). A criticism of Feinberg, however, is that he sets a sliding scale of Mill’s liberalism rather than properly offering a moral criterion of paternalism.

The second criticism addressed to paternalism is that it violates the fundamental principle of equality. Many philosophers understandably condemn paternalism for its condescending and offensive practices (Shiffrin, 2000, Anderson, 1999, Kleinig, 1983, VanDeVeer, 1986). A number of liberal thinkers consider the right to equality second to none (Dworkin, 1978, Anderson, 1999, Rawls, 1971, Arneson, 2000), but they often disagree on what constitutes an unfair inequality. Paternalism primarily originates from the desire to correct inequality of luck, that is to say inequalities that do not result from individual choices. There is little merit in becoming deputy director of a company when one is born in a rich and well-educated family in the nineteenth century, and when moreover the company belongs to one’s father. It is conversely extremely difficult to access such a position when one is born in a working class family or a poorhouse orphanage, for example. This is what is called (bad) brute luck. This is the reason why, historically, the wealthy and the well-educated often felt the need to share

---

<sup>3</sup> Commentators disagree on Mill’s positions towards paternalism. Richard Arneson, for instance, has argued that Mill’s political philosophy is not compatible with soft anti-paternalist positions ARNESON, R. (1980), ‘Mill versus Paternalism’, *Ethics*, 90, 470-89.

<sup>4</sup> Voluntariness is, for Feinberg, essentially based on responsibility. The five above circumstances obliterate the decision maker’s responsibility since the choices made do not express the individual character. Conversely, individuals take full responsibility for autonomous or voluntary choices “since they represent him faithfully, expressing his settled values and preferences” (Feinberg, 1989: 113). For Feinberg, misinformation always constitutes an autonomy-diminishing factor. Richard Arneson argues, however, that “a person’s actions may be authentic expressions of his personality without being deliberately chosen” (Arneson, 1980: 487). 4b is thus only acceptable if the behavior responsible for the misinformation is itself ‘out of character’, that is, which is submitted to one of the previous autonomy-diminishing factors. He concludes that voluntary choice is admittedly “important but [that it] does not plausibly have the make-or-break significance that soft paternalism attaches to it” and that it is therefore “a mistake to make a fetish of” it (Arneson, 2005).

what they believed to be undeserved luck with the working classes by providing them with free access to education. Although intuitively very appealing, luck egalitarianism has come under criticism for its demeaning attitude towards those who were deemed to be victims of bad luck. What Anderson reproaches about luck egalitarianism is that it categorizes the victims of bad luck according to what they are (stupid, untalented, disabled, ugly, socially awkward, etc...) <sup>5</sup>. No wonder that she finds its response, paternalism, offensive. Unsolicited, paternalist actions have disastrous effects on the ill-endowed people they are supposed to help. By granting people a (financial) compensation for their condition, they actually treat them as incomplete human beings. In its most traditional form (derived from the father-child analogy), paternalism is also often justified by the economic, cognitive or social superiority of the paternalist agent <sup>6</sup>.

The third criticism addressed to paternalism is that under the false pretension of being altruistic, it actually only serves the interest of the paternalistic agent. Paternalism is said to be immoral because it is falsely moral and that it uses immoral means to conceal its real nature. The captains of industry of the nineteenth-century who built social houses, school and hospitals for their workers have often been accused of being more interested in the productivity and the low turnover of their company than by the actual well-being of their low paid and hard working employees. Besides, it is not difficult to imagine cases where ‘good natured people’ resort to ‘wrong means’ in order to complete what they consider as being ‘a good action’ <sup>7</sup>. There is undoubtedly some truth in this. There is no point in denying it.

---

<sup>5</sup> Elizabeth Anderson argues that the point of equality was neither resources nor luck but status. And treating a category of people as children or adults as “too stupid to run their lives” (Anderson, 1999: 301) is incompatible with the respect all individuals should equally be granted. “Egalitarianism”, she writes, “ought to reflect a generous, humane, cosmopolitan vision of a society that recognizes individuals as equals in all their diversity. It should promote institutional arrangements that enable the diversity of people’s talents, aspirations, roles, and cultures to benefit everyone and to be recognized as mutually beneficial. Instead, the hybrid of capitalism and socialism envisioned by luck egalitarians reflects the mean-spirited, contemptuous, parochial vision of a society that represents human diversity hierarchically, moralistically contrasting the responsible and irresponsible, the innately superior and the innately inferior, the independent and the dependent. It offers no aid to those it labels irresponsible, and humiliating aid to those it labels innately inferior.” (Anderson, 1999: 308).

<sup>6</sup> For Gert and Culver, for instance, “paternalism can be practiced by anyone who has qualifications which he believes enable him to see better than S what is good for S’s good” (Gert and Culver, 1976: 50).

<sup>7</sup> Allen Buchanan, for instance, has condemned a number of disgraceful medical practices involving withholding information or even bluntly lying to patients about their state of health to ‘avoid them any unnecessary harm’ (Buchanan, 1978). Bernard Gert and Charles Culver, studying the moral practices of medicine (a field known today as bioethics), also condemn the shams of medical paternalism (Gert and Culver, 1979, Gert et al., 2006). They contend in particular that “an essential feature of paternalistic behavior toward a person is the violation of moral rules (or doing that which will require such violations), for example, the moral rules prohibiting deception, deprivation of freedom or opportunity, or disabling” (Gert and Culver, 1976: 48). Most of the criticisms

Paternalism has indeed suffered from a long history of abusive practices. But what has been done in its name should not stop us looking for its real nature. I feel frustrated when I see a definition that automatically endorses the popular view that paternalism is a deceitful practice<sup>8</sup>. In that respect, I believe that paternalism does not significantly differ from altruism: a large number of so-called altruistic actions are in fact self-interested but it does not imply that altruism, as a general practice, is wrong. So let us admit that moral actions do exist, and that paternalism can be genuinely altruistic. By altruistic I mean disinterested and benevolent. I do not contest the possibility that the agent might have some indirect and/or long-term interest in his benevolent action. After all, we all ultimately gain from altruism. I merely consider that his action is altruistic if it is not motivated by self-interest. To borrow Thomas Nagel's phrase, it is "a willingness to act in consideration of the interests of other persons, without the need of ulterior motives" (Nagel, 1970: 79). Moral paternalism relies on the possibility of altruism.

### **3. Behavioral Paternalism**

In this section I present a new approach to paternalism that is referred to sometimes as 'New Paternalism' (Rizzo, 2009, Mead, 1997, Economist, 2006, Holt, 2006), 'Asymmetric Paternalism' (Camerer et al., 2003, Loewenstein et al., 2007), 'Light' (Loewenstein and Haisley, 2008) or 'Libertarian Paternalism' (Sunstein and Thaler, 2003, Thaler and Sunstein, 2003, Thaler and Sunstein, 2008). Despite this variety of designations I nevertheless prefer to call it 'Behavioral Paternalism' (hereafter BP). My justification for this is that, apart from 'New Paternalism' (but new is not a definitive characteristic), all these expressions prejudge the legitimacy of this new type of paternalism. Since my aim is precisely to discuss the terms of its legitimacy, I therefore believe that a neutral but

---

addressed to medical paternalism can nevertheless be handled through the principle of equality put forward by Elizabeth Anderson. What is, I believe, truly shocking in these practices is not the fact that doctors lie to their patients but the fact that they consider them incapable to cope with their own death.

<sup>8</sup> For Allen Buchanan, for instance, "[P]aternalism is usually characterized as interference with a person's liberty of action, where the alleged justification of the interference is that it is for the good of the person whose liberty of action is thus restricted." [I underline] BUCHANAN, A.E. (1978), 'Medical Paternalism', *Philosophy and Public Affairs*, 7, 371-90.

informative qualifying adjective is more appropriate. The term ‘behavioral’ imposes itself since the revival of paternalism observed for the last decade originates from behavioral sciences.

The American lawyer Cass Sunstein and economist Richard Thaler are the first and most important champions of BP (Jolls et al., 1998, Sunstein and Thaler, 2003, Thaler and Sunstein, 2003, Thaler and Sunstein, 2008, Jolls and Sunstein, 2006). Their groundbreaking works on libertarian paternalism have produced a vast academic literature and found a favorable echo with politicians. Their work, directly inspired by Richard Thaler’s own contributions to behavioral economics, has the merit of raising anew the question of the possibility and the legitimacy of paternalism in the 21<sup>st</sup> century. The paternalism they defend radically differs from the traditional philosophical approaches inherited from Mill’s ‘soft anti-paternalism’ (Feinberg, 1986: 15). They contest, in particular, the idea that paternalism is necessarily coercive and that it is predominantly institutionally-based.

Behavioral studies gave partisans of paternalism two major arguments. Firstly empirical studies have repeatedly shown that individual actions do not meet the rationality standard that the economic theory would expect. The idea of bounded rationality, introduced by Herbert Simon (Simon, 1957), has been decisive in explaining what was otherwise considered as economically inconsistent (Elster, 1985, Simon, 1982, Rubinstein, 1998, Kahneman, 2003). Ted O’Donoghue and Matthew Rabin hence show that their model of sin taxes acknowledging individuals’ bounded rationality has huge consequences in terms of economic policy: “Economic policy prescriptions might change once we recognize that humans are humanly rational rather than superhumanly rational, and in particular it may be fruitful for economists to study the possible advantages of paternalistic policies that help people make better choices.[...] In other words, economists ought to treat the analysis of optimal paternalism as a mechanism-design problem when some agents might be boundedly rational.” (O’Donoghue and Rabin, 2003). For some years now, economic models of bounded rationality have been supplemented by new evidence provided by neuroscientists, giving birth to a new discipline called neuroeconomics (Camerer et al., 2004, Camerer et al., 2005, Camerer, 2007, Kahneman, 2011). According to Colin Camerer neurosciences definitely undermine the ‘Revealed-Preference Approach’ by showing the role played by the difference between experienced utility (‘liking’) and decision-utility (‘wanting’) in

individual suboptimal choices. Benevolent agents (or governments), who are aware of the gap between the utility individuals think they will obtain (what they want) and the utility (or disutility) they will actually get (what they like), are then given the opportunity to act paternalistically without risks (Camerer, 2006, Camerer, 2008). In his terms “paternalism could be justified, in terms of a person’s own welfare, if the wanting system does not produce what the liking system likes and if the intervention creates more liking than the person would achieve on his own or with market-supplied help” (Camerer, 2006: 101-02).

A second argument for paternalism is individuals’ endogenous preferences. Preferences are claimed by behaviorists to be endogenous to social contexts and consequently inconsistent from one situation to another, the lack of any past or similar reference introducing an element of arbitrariness in the evaluation<sup>9</sup>. Over the last twenty years, behavioral economics has offered a great deal of evidence to support this thesis (Lichtenstein and Slovic, 2006). Framing, anchor or endowment effects have been identified as some of the key mechanisms at stake (Thaler, 1980, Loewenstein and Adler, 1995, Tversky and Kahneman, 1986, Tversky and Kahneman, 1981, Kahneman et al., 1991). But if endogenous preferences are responsible for individuals’ poor rational behavior, they can also be the medium for their own remedy; the remedy is in the disease. For Sunstein and Thaler, for instance, it is possible to significantly reduce the lack of individual rationality by switching the default rule settings from an ‘opt in’ to an ‘opt out’ formula. They illustrate the benefits of setting change in default rules with two case studies: the case of 401 (k) employee savings plans (Madrian and Shea, 2001, Choi et al., 2002) and their own experience of the University of Chicago’s car park payment system.

In the first case, employers decide that instead of letting their employees join a savings plan, the American 401 (k) plan, they would automatically enroll them unless employees specifically stated that they did not wish to be included. When eligible, employees usually receive a 401 (k) plan form that they must complete in order to join. In this case, however, employees only receive a statement in

---

<sup>9</sup> The idea of ‘coherent arbitrariness’ of preferences developed by Dan Ariely, George Loewenstein and Drazen Prelec shows, however, that valuations are not entirely arbitrary: absolute valuation of goods or experience is for a large part arbitrary, but relative valuations can prove to be coherent ARIELY, D., LOEWENSTEIN, G. and PRELEC, D. (2003), "Coherent Arbitrariness": Stable Demand Curves without Stable Preferences', *The Quarterly Journal of Economics*, 118, 73-105.

which they are told that, unless they opt-out, they are automatically enrolled in the plan. In both studies, a significant increase in the enrollment rate was observed. In the second case, a change in the tax law made it possible for employees to pay for employer-provided parking on a pre-tax basis. The University of Chicago presumed that, although being in the employees' interest, many of them would not take the trouble to fill in and send back the form and decided to enroll all employees automatically. Those who preferred to pay with after-tax dollars were, however, free to opt out<sup>10</sup>. Had the university chosen a different default rule (like an 'opt in' strategy), the authors reckon that "many employees, especially faculty members (and probably including the present authors), would still have that form buried somewhere in their offices and would be paying substantially more for parking on an after-tax basis" (Sunstein and Thaler, 2003: 1171). A minor effort, such as filling in a form, is often sufficient not to claim a benefit people are entitled to.

Consider now their example of the cafeteria. The director is aware that the way she arranges the line of dishes will 'frame' the customers' choice. According to the authors, she faces four alternatives<sup>11</sup> (Sunstein and Thaler, 2003: 1164):

1. She could make choices that she thinks would be best for the customers, all things considered.
2. She could make choices at random.
3. She could choose those items that she thinks would make the customers as obese as possible.

---

<sup>10</sup> In practice, however, default rules are sticky and guaranteeing the conditions of free choice is difficult. Once a person has been enrolled in a default rule, he or she tends to overestimate its benefits (endowment effect). If opting in is a deterrent for many people, opting out from a default rule is even more dissuasive. To respect freedom of choice, a libertarian planner must set up a default rule that will be easy and costless for individuals to opt out of. Ideally, it should not be more demanding than a 'one-click' procedure (Sunstein & Thaler 2008: 249).

<sup>11</sup> Robert Sugden pointed out that one option is surprisingly missing from the list: the cafeteria director might simply choose to display the dishes in the most economically efficient way SUGDEN, R. (2008), 'Why incoherent preferences do not justify paternalism', *Constitutional Political Economy*, 19, 226-48. The authors' reasons for not including it are twofold. Firstly they argue that planners are not necessarily looking for profit maximization. Even directors of private companies, like school cafeterias, are not constantly subject to market pressures. Sunstein and Thaler do not distinguish profit driven from non-profit driven planners. Secondly, they claim that "market success will come not from tracking people's ex ante preferences, but from providing goods and services that turn out, in practice, to promote their welfare, all things considered. Consumers might be surprised by what they end up liking; indeed, their preferences might change as a result of consumption." SUNSTEIN, C.R. and THALER, R.H. (2003), 'Libertarian Paternalism Is Not an Oxymoron', *The University of Chicago Law Review*, 70, 1159-202. Although, technically speaking, one could argue that denying the existence of ex ante preferences does not logically imply that individual preferences are not being taken into account for profit making, it seems reasonable from the authors' point of view to exclude the profit option from the manager's choice. In the next section I present an amended version of the cafeteria example in which profit-making cannot be a motivation.

4. She could give customers what she thinks they would choose on their own.

Suppose for the moment that she chooses the first option (I shall discuss further her motivation to do so in the next section). She is well-aware that customers (much like herself) know what is healthy for them and what is not. But she is also aware (as her customers are as well) that the colors, the shapes, the odors and the disposition of goods arranged in front of them can influence their choice towards the unhealthiest dishes. Being director of the cafeteria and being in charge of the disposition of the line of dishes, she therefore has the opportunity to rearrange the dishes in order to reduce this behavioral bias.

#### **4. Morality of BP (I): autonomy**

To what extent is BP morally legitimate? In the first section I have suggested that – in order to be moral – paternalism ought to fulfill three criteria: respecting personal autonomy, respecting individual equality and being morally motivated. Let us consider these three criteria in connection with behavioral paternalism. The first condition is by far the one that has been given the most attention. Defenders of BP trust that its main asset is not to be coercive. This is the reason why they described it as ‘libertarian’, ‘soft’, ‘light’ or ‘asymmetric’. Sunstein and Thaler describe as their approach as “a relatively weak and nonintrusive type of paternalism” (Sunstein and Thaler 2003:1162).

Soft or weak paternalism typically describes situations of involuntary acts or misinformed decisions that must be interfered with for the person’s own good. Typically, a soft paternalist can legitimately prevent a person from using the saltcellar if they have not been informed that it contains arsenic instead of salt. Similarly, a person under the influence of drugs or alcohol can be temporarily prevented from making decisions that he/she may bitterly regret the next day (Feinberg 1986). Alternatively, if a person is appropriately informed of the risks incurred from crossing a dangerous bridge, smoking, or duelling, then no external interference is ever justifiable to stop that person

carrying out such an action<sup>12</sup>. Soft paternalism re-establishes what some regard as an essential feature of liberalism: the right to act irrationally.

Sunstein and Thaler's libertarian paternalism is not a form of soft paternalism. It is, on the contrary, rooted in a very weak notion of voluntariness that makes it closer to hard paternalism. Their analysis, based on behavioral studies, shows that most individual choices are sub-optimal because their valuations depend on arbitrary anchors and frames. Their actions reflect the context in which they have been carried out rather than the proper preferences of the agents who welcome an external aid to help them make the right choice. To a certain degree, most of their actions are then involuntary: they want to be thin but eat fatty foods, they want to have a comfortable retirement but don't save enough money, they want to donate their organs but don't say so etc. Actions are not made involuntary because information is hidden but because individuals misuse it.

The libertarian feature of their theory relies exclusively on freedom of choice. Yet there is an incongruity at simultaneously defending absolute freedom of choice and a weak concept of voluntariness. By discarding self-ownership, the authors actually undermined the possibility of libertarian paternalism. The number of opportunities a person gets is not relevant to assess their degree of freedom. Let me explain this by borrowing an example originally presented by Locke<sup>13</sup>. A prisoner is brought into a cell and is convinced that all the doors are locked when actually one has been left unlocked. Is this person free to get out? A libertarian like Nozick would answer in the affirmative. Technically, nobody physically stops him from getting out. He is not *coerced* into staying inside. A behaviourist, or for that matter a philosopher of autonomy like Dworkin, would argue, however, that since the prisoner is not aware of this opportunity, he is not any freer than if all the doors *were* actually locked. In the same way as in the cafeteria, where the order of the dishes is more significant to the consumers than their presence, the information given about the doors is more significant to the prisoner than their real state. If individuals were as rational as the economists suppose them to be (if they were 'Econs' to borrow Sunstein and Thaler's terminology), they would not be subject to framing

---

<sup>12</sup> Other means are available such as debiasing cognitively or emotional wronged choices (Blumenthal, 2007, Mitchell, 2005, Jolls and Sunstein, 2006).

<sup>13</sup> I borrow this example myself from Dworkin who used it to support his distinction between liberty and autonomy (Dworkin 1988:14).

or anchor effects. They would correctly use the information they are given to maximise their welfare and all their actions would result from deliberate and well-informed decisions. Their actions would then be regarded as fully *voluntary*.

Despite these criticisms, it would – I believe – be unfair not to acknowledge the progress realized by the partisans of BP in their attempts to minimize all forms of coercive interferences. Despite its imperfections BP certainly has better chances (compared with other types of paternalism) to satisfy the autonomy proviso. But its main moral asset is not there. What BP does really change for paternalism is its relation to equality.

## **5. Morality of BP (II): equality**

In the first section I showed how the equality proviso was challenging for the defense of paternalism. I believe, however, that behavioral paternalism can partly elude the criticisms mentioned earlier. Paternalism, as previously argued, is usually motivated by a willingness to correct an (unfair) inequality between people. It thus implicitly establishes a distinction between the ‘lucky’ and ‘unlucky’ ones. Amongst the ‘unlucky’ ones, paternalism particularly favors those who have no personal responsibility in their misfortune (‘brute bad luck’): the ‘disabled’, the ‘ugly’, the ‘stupid’, and the ‘socially awkward’ ones (Anderson, 1999).

But paternalism is only demeaning if it finds its justification in such categories. I believe, however, that this ought not to be the case, and that a morally legitimate paternalism is possible if it is grounded on positional asymmetries. This means, in particular, that a superior economic rationality (a higher ability to compute or to order preferences) does not qualify as a good reason to interfere with someone else’s life. In addition to being demeaning for those deemed to be intellectually ill-endowed, it has disastrous totalitarian consequences as it logically grants the most intelligent individuals the ability to overturn the decisions of the normally endowed ones ‘for their own sake’ (Wikler, 1979). By contrast, it is – I contend – morally acceptable for a person to act when her relative position gives her a cognitive advantage. Take Mill’s famous example of the threatening bridge (Mill, 1989). Consider that A only knows about the poor state of the bridge because he is a local, whereas B – who is about to

cross it – is foreign to the area. If B were local himself, he would not need A’s warning and would also probably warn foreigners against the danger of this bridge. A is in a relatively better position to make a decision on this matter than B. This situation does not, however, say anything about A being inherently superior to B.

The same argument can apply to BP. It is possible to show that in the case of BP, paternalist agents cannot claim any superiority of body or mind. Take the cafeteria example: the director, or whoever arranges the dishes, is no different from her customers. In fact, she is a customer herself on her days off. The customers of the cafeteria are not victims of brute luck, waiting for her help. And she has no reason whatsoever to be kind to them (Sunstein and Thaler wrongly claim that she is necessarily benevolent, see next section). Nothing distinguishes her from the ‘paternalized customers’ except the position she occupies at the decision-making moment. If she does arrange the dishes so that her customers eat healthier, then her action undoubtedly qualifies as paternalistic. There is nothing demeaning about it: customers are not disgraced or marked out and the director is not motivated by a sense of justice. She arranges the dishes relatively to what she considers to be a ‘good’ order: meaning good for her and good for the others. BP does not require a cognitive or moral superiority from the paternalist agents. BP is what could be referred to as an ‘opportunistic paternalism’. If anything paternalism reflects individuals’ equality in front of behavioral bias.

Sunstein and Thaler do not appreciate this point in their works. They prefer seeing paternalist agents as ‘planners’ or ‘choice architects’. The Paternalism Behaviorism I am presenting here is not a defense of Sunstein and Thaler’s ‘Nudges Theory’. It is, I insist, an acknowledgement of its theoretical potentialities, including those they have chosen not to pursue. The difference between libertarian paternalism and BP should become even clearer in the next section, in which I shall argue that to be morally acceptable BP should rely on what I can only call a moral inconsistency, a point that is irreconcilable with Sunstein and Thaler’s own theory.

## **6. Morality of BP (III): altruism**

In the previous paragraphs I have demonstrated 1) that although it does not perfectly conform to the autonomy proviso, BP does considerably better than any other form of paternalism (enough to be socially tolerable) and 2) that it is the only kind of paternalism that actually conforms to the equality proviso. The last point to be considered is the genuineness and the (un)avoidability of its altruism.

In the first section I showed that paternalism is often dismissed as being falsely altruistic. Partisans of BP could argue, however, that this is not the case and that paternalist agents are genuinely morally motivated. What reason can agents possibly have to act altruistically?

Most attempts to explain altruism in decision theory rest upon an empathetic conception of altruism (Fehr and Schmidt, 1999, Gintis et al., 2008, Kolm and Ythier, 2006, Sen, 1991, Kirman A and Teschl M, 2010, Elster, 2009). An ‘altruistic’ agent is an agent whose well-being is functionally linked to other people’s well-being. She is therefore as interested as anyone else, the only difference being that her preferences are essentially other-regarding or ‘external’. Similarly, the philosophical notions of compassion and pity also involve the benevolent agent’s self-interest. Martha Nussbaum, who defends compassion as the foundation of ethics, defines it as “a painful emotion occasioned by the awareness of another person’s underserved misfortune” (Nussbaum, 2003: 301). Less instinctive and more positively orientated, sympathy plays a comparable role with compassion by connecting all individuals’ interest altogether (Fontaine, 1997, Sugden, 2002, Singer and Fehr, 2005). This is, however, not what I call altruism. Altruism, as I understand, it refers to purely disinterested actions. Besides, paternalistic actions motivated by compassion would undoubtedly violate the equality proviso described above. So far paternalism has then failed to prove its good intentions. Recent works in behavioral economics offer, however, new perspectives on the topic. Behavioral studies suggest indeed that people can behave out of interest and that their likeliness to do so largely depends on circumstances. In other words behavioral biases not only justify but also explain paternalistic actions. Consider the cafeteria example quoted above. Suppose now that the director of a cafeteria has to order fifteen dishes equally distributed between three classes of profitability level. She has to order all dishes, including the equally profitable ones. Since her interest as manager of the cafeteria is already

satisfied, she is otherwise indifferent to the way in which dishes are arranged within each profitability category. She could, for instance, adopt a random criterion of distribution. But moral options are also available to her: she could, as suggested by Sunstein and Thaler (Sunstein and Thaler, 2003, Thaler and Sunstein, 2008), promote healthy over fattening dishes, in which case she would act altruistically. Altruism – and consequently paternalism – is therefore theoretically possible.

But how likely is it for the director of the cafeteria to choose the altruistic option? I believe that Sunstein and Thaler are wrong to consider that paternalism is her only real choice (Sunstein and Thaler, 2003: 1164). Malevolence is fortunately uncommon but it nonetheless exists (Kolm, 1995). Random choice set apart<sup>14</sup> there are moreover many other possible criteria to arrange dishes: one can order them by size, by color, by shape, by composition, by season, by name, by origin etc. Altruism is then just one of many options available to the director of the cafeteria. Why would she prefer this particular option to all the others?

Behavioral economists have repeatedly shown that self-interested individuals often prove unable to act rationally. In a similar way morally motivated individuals can also be incapable of acting correspondingly. Behavioral biases divert decision-makers away from their ‘ideal’ choice in morals as well as in prudence. Numerous experimental studies support this statement, a number of which have been presented by John Doris in a compelling book on individuals’ (absence of) character (Doris, 2002)<sup>15</sup>. In other words, individuals do not act altruistically because they are morally good persons,

---

<sup>14</sup> Can a choice be absolutely random? Statisticians and neuroscientists have observed that the human brain is ill-disposed toward randomness and that individuals frequently prefer adopting any given criterion of choice rather than (unsuccessfully) attempting to simulate a random statistical distribution. MLODINOW, L. (2008), *The Drunkard's Walk: How Randomness Rules Our Lives*. Knopf Doubleday Publishing Group.

<sup>15</sup> Amongst the vast array of experiments quoted by Doris, let me mention two cases to illustrate the importance of framing effects in moral choice. The first one, undertaken in 1972 by the American psychologists Alice Isen and Paula Levin, took place in a shopping mall near a phone box (Isen and Levin, 1972). Passersby and customers using the phone box were the unwitting participants of the experiment. At the end of each call, a young woman ‘accidentally’ dropped a folder, thus spreading all her papers across the floor. The aim of the experiment was to study whether individuals are more likely to behave altruistically when they are in a good mood. Phone callers are divided in two groups, a control group for which nothing is done and a test group composed of good-humored subjects. In order to put the subjects in a good mood the experimenters place a ‘forgotten’ dime in the phone receiver for the callers to find. The results of the study were striking: 14 out of 16 callers who found the dime came to the aid of the ‘clumsy’ young woman, whilst only one out of the 25 who did not find the dime did! The second study belongs to a series of experiments led by the social psychologist Robert A. Baron (Baron and Thomley, 1994, Baron and Bronfen, 1994, Baron, 1997) highlighting the effect of aromas on human behavior. In the last experiment Baron shows that individuals located near to a pleasant source of aromas (such as hot croissants or coffee) were more disposed to help passersby (in need of change in his

but because favorable circumstances prompt them to act in such a way at this particular moment. The same person in the same situation but under different circumstances might act differently. A simple example could clear things out. When I go to town, I always park my car in the same parking. There is invariably a beggar (usually the same) waiting for some change at the till. Sometimes I give him money, sometimes I don't. Situations and persons are exactly the same. The only thing that varies—and that ultimately motivate my action – is the state of my mood. When I am in a good mood I give away, when in a bad mood I don't. Being a sunny day can make the difference. We all have similar experiences in our lives. I believe the same thing applies to the director of the cafeteria. Having the opportunity to act altruistically or paternalistically is not a reason *per se* to do so. And this, I shall argue, is fortunate as it not only demonstrates that paternalism can be genuinely morally motivated (3<sup>rd</sup> criticism) but it also reinforces BP's position in terms of respect of personal autonomy (1<sup>st</sup> criticism) and individual equality (2<sup>nd</sup> criticism).

Paternalist agents are not more rational or more insightful than the people they propose to help. But what if they were morally 'superior'? What if they owed their status to their moral rather than to their cognitive excellence? This would, I contend, have dramatic consequences. Paternalism has been widely (and partly correctly) criticized for letting the self-proclaimed 'uppermost intelligent' or 'supremely rational' individuals interfere with those they considered 'too stupid to run their lives'. If paternalism were to be the prerogative of the 'morally superior', then they would undoubtedly interfere with the lives of those they deem as 'too corrupt to live a morally worthy life'. What then differentiates moral paternalism, as Gerald Dworkin calls it, from normal or legal paternalism "is the idea that a person's welfare can directly involve, and therefore be increased by, changes in their moral character and actions – changes which do not necessarily result in their being happier or healthier, etc." (Dworkin, 2005: 308). These changes are coercive (and thus illegitimate) if they are not later endorsed by the people concerned (Feinberg, 1986). From an equality point of view they are, however,

---

experiment) than those who were situated in a neutral aromatic environment (Baron, 1997). All these observations, as well as many others, corroborate recent neurobiological studies that show the impact of (positive) emotions on altruism (Damasio, 2004).

always illegitimate as they specifically target (and thereby ostracize) a specific category of the population.

The fact that individuals who have the opportunity to act paternalistically do not systematically decide to do so has a second advantage: it limits the practice of paternalism and consequently contributes to its social acceptability. Most of the philosophical debate on freedom focused on the legitimacy of paternalistic actions taken separately. Very little attention has been given, however, to the disagreement and the sense of violation felt by individuals who continuously have their self-regarding choices interfered with, however legitimate each of these interferences might be. Despite its limits, voluntariness might be a good criterion to assess the legitimacy of an interference (Olsaretti, 2004) but it does not take into account the moral cost born by the people who are interfered with. Imagine Maria decides to dine out with some friends to celebrate her 40<sup>th</sup> birthday. Maria is seriously overweight. She tries to lose weight but has little willpower. At the restaurant she orders a pizza and some Italian wine. They all have an enjoyable evening and then the waiter comes to take the dessert orders. Maria is about to choose a dessert, when a man at the neighboring table suddenly stands up and urges her ‘for her own sake’ not to have dessert. Pizza, he says, was already highly inappropriate for her, considering her weight; dessert would be ‘suicidal’. Technically speaking his intervention is legitimate since her eating fat is not ‘fully voluntary’. His action is, however, socially inconsiderate and morally harmful. Even if Maria has willpower issues she should have the right not to have all her alimentary choices being questioned and she should certainly not be subjected to such public humiliation. All paternalistic opportunities should not be exercised, however legitimate or well-intentioned they seem to be. If everyone were to act altruistically our life would soon become absolutely unbearable.

## **7. Conclusion**

In this article I have endeavored to show that the new type of paternalism defended by Sunstein and Thaler (amongst others) is better prepared to respond to moral attacks than any other type of

paternalism so far. Although its libertarian feature is not fully asserted, behavioral paternalism – correctly stated – does not fall under Anderson’s equalitarian critic. Unlike its more traditional expressions it can also prove to be truly altruistic. Its disinterestedness and absence of condescendence make BP the only acceptable form of paternalism in modern society. BP is not morally irreproachable yet. To some extent it still does infringe on individual freedom and it certainly cannot be described as ‘libertarian’. I have not, however, been interested here by its ideological potentialities but by its moral implications. Paternalism, as I said, is a form of altruism. Yet, one is intuitively regarded as ‘good’ when the other is always regarded as ‘wrong’. What exactly makes paternalism wrong? At what conditions could it be ‘good’? Those are the questions I tried to answer to here. But to complete this investigation one would also need to question the nature of altruism: ‘What exactly makes altruism ‘good’?’ and ‘At what conditions could it be ‘wrong’?’ My very last point suggests indeed that what people believe to be wrong in paternalism could well also be something that is wrong with altruism.

ANDERSON, E. (1999), 'What Is the Point of Equality?', *Ethics*, 109, 287-337.

ARIELY, D., LOEWENSTEIN, G. and PRELEC, D. (2003), "'Coherent Arbitrariness": Stable Demand Curves without Stable Preferences', *The Quarterly Journal of Economics*, 118, 73-105.

ARNESON, R. (1980), 'Mill versus Paternalism', *Ethics*, 90, 470-89.

ARNESON, R. (2000), 'Luck Egalitarianism and Prioritarianism', *Ethics*, 110, 339-49.

ARNESON, R. (2005), 'Joel Feinberg and the Justification of Hard Paternalism', *Legal Theory*, 11, 259-84.

BARON, R.A. (1997), 'The Sweet Smell of... Helping: Effects of Pleasant Ambient Fragrance on Prosocial Behavior in Shopping Malls ', *Personality and Social Psychology Bulletin*, 23, 498-503.

- BARON, R.A. and BRONFEN, M.I. (1994), 'A Whiff of Reality: Empirical Evidence Concerning the Effects of Pleasant Fragrances on Work-Related Behavior<sup>1</sup>', *Journal of Applied Social Psychology*, 24, 1179-203.
- BARON, R.A. and THOMLEY, J. (1994), 'A Whiff of Reality: Positive Affect as a Potential Mediator of the Effects of Pleasant Fragrances on Task Performance and Helping', *Environment and Behavior*, 766-84.
- BLUMENTHAL, J.A. (2007), 'Emotional Paternalism', *Florida State University Law Review*, 35, 1-72.
- BUCHANAN, A.E. (1978), 'Medical Paternalism', *Philosophy and Public Affairs*, 7, 371-90.
- CAMERER, C. (2008), 'The Case of Mindful Economics', in CAPLIN, A. and SCHOTTER, A. (eds.) *The Foundations of Positive and Normative Economics: A Handbook: A Handbook*. New York: Oxford University Press, USA.
- CAMERER, C., ISSACHAROFF, S., LOEWENSTEIN, G., O'DONOGHUE, T. and RABIN, M. (2003), 'Regulation for Conservatives: Behavioral Economics and the Case for "Asymmetric Paternalism"', *University of Pennsylvania Law Review*, 151, 1211-54.
- CAMERER, C., LOEWENSTEIN, G. and PRELEC, D. (2005), 'Neuroeconomics: How Neuroscience Can Inform Economics', *Journal of Economic Literature*, 43, 9-64.
- CAMERER, C.F. (2006), 'Wanting, Liking, and Learning: Neuroscience and Paternalism', *The University of Chicago Law Review*, 73, 87-110.
- CAMERER, C.F. (2007), 'Neuroeconomics: Using Neuroscience to Make Economic Predictions\*', *The Economic Journal*, 117, C26-C42.
- CAMERER, C.F., LOEWENSTEIN, G. and PRELEC, D. (2004), 'Neuroeconomics: Why Economics Needs Brains', *Scandinavian Journal of Economics*, 106, 555-79.
- CHOI, J., DAVID LAIBSON, BRIGITTE MADRIAN and METRICK, A. (2002), 'Defined Contribution Pensions: Plan Rules, Participant Decisions, and the Path of Least Resistance', *Tax Policy and the Economy*, 16, 67-114.
- DAMASIO, A. (2004), *Looking for Spinoza. Joy, Sorrow and the Feeling Brain*. London: Vintage.

- DORIS, J.M. (2002), *Lack of Character. Personality and Moral Behavior*. New York: Cambridge University Press.
- DWORKIN, G. (1971), 'Paternalism', in FEINBERG, J. and GROSS, H. (eds.) *Philosophy of Law*. Wadsworth Publishing Company.
- DWORKIN, G. (1983), 'Paternalism: Some Second Thoughts', in SARTORIUS, R. (ed.) *Paternalism*. Minneapolis: University of Minnesota Press.
- DWORKIN, G. (2005), 'Moral paternalism', *Law and Philosophy*, 24, 305-19.
- DWORKIN, R.M. (1978), *Taking Rights Seriously*. Harvard University Press.
- ECONOMIST, T. (2006), 'The new paternalism  
The avuncular state  
A smarter, softer kind of paternalism is coming into style ', *The Economist*. London.
- ELSTER, J. (1985), *Sour Grapes: Studies in the Subversion of Rationality*. Cambridge University Press.
- ELSTER, J. (2009), *Traité critique de l'homme économique: Le désintéressement*. Éditions du Seuil.
- FEHR, E. and SCHMIDT, K.M. (1999), 'A Theory of Fairness, Competition, and Cooperation', *The Quarterly Journal of Economics*, 114, 817-68.
- FEINBERG, J. (1983), 'Legal Paternalism', in SARTORIUS, R. (ed.) *Paternalism*. Minneapolis: University of Minnesota Press.
- FEINBERG, J. (1986), *Harm to Self: The Moral Limits of the Criminal Law*. New York: Oxford University Press.
- FEINBERG, J. (1989), *The Moral Limits of the Criminal Law: Volume 3: Harm to Self*. OUP USA.
- FONTAINE, P. (1997), 'Identification and Economic Behavior Sympathy and Empathy in Historical Perspective', *Economics and Philosophy*, 13, 261-.
- GERT, B. and CULVER, C.M. (1976), 'Paternalistic Behavior', *Philosophy and Public Affairs*, 6, 45-57.
- GERT, B. and CULVER, C.M. (1979), 'The Justification of Paternalism', *Ethics*, 89, 199-210.
- GERT, B., CULVER, C.M. and CLOUSER, K.D. (2006), *Bioethics: A Systematic Approach*. Oxford University Press.

- GINTIS, H., HENRICH, J., BOWLES, S., BOYD, R. and FEHR, E. (2008), 'Strong Reciprocity and the Roots of Human Morality', *Social Justice Research*, 21, 241-53.
- HILL, C.A. (2007), 'Anti-Anti-Anti-Paternalism ', *NYU Journal of Law and Liberty*, 2, 444-54.
- HOLT, J. (2006), 'The New, Soft Paternalism ', *The New York Times*.
- ISEN, A.M. and LEVIN, P.F. (1972), 'Effect of Feeling Good on Helping: Cookies and Kindness', *Journal of Personality and Social Psychology*, 21, 384-88.
- JOLLS, C. and SUNSTEIN, C.R. (2006), 'Debiasing Through Law ', *Journal of Legal Studies*, 35, 199-241.
- JOLLS, C., SUNSTEIN, C.R. and THALER, R. (1998), 'A Behavioral Approach to Law and Economics', *Stanford Law Review*, 50, 1471-550.
- KAHNEMAN, D. (2003), 'Maps of Bounded Rationality: Psychology for Behavioral Economics', *The American Economic Review*, 93, 1449-75.
- KAHNEMAN, D. (2011), *Thinking, fast and slow*. London: Allen Lane.
- KAHNEMAN, D., KNETSCH, J.L. and THALER, R.H. (1991), 'Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias', *The Journal of Economic Perspectives*, 5, 193-206.
- KAHNEMAN, D. and TVERSKY, A. (2000), *Choices, Values, and Frames*. Cambridge University Press.
- KIRMAN A and TESCHL M (2010), 'Selfish or selfless? The role of empathy in economics', *Philos Trans R Soc Lond B Biol Sci.*, 365, 303-17.
- KLEINIG, J. (1983), *Paternalism*. Manchester University Press.
- KOLM, S.-C. (1995), 'THE ECONOMICS OF SOCIAL SENTIMENTS: THE CASE OF ENVY', *Japanese Economic Review*, 46, 63-87.
- KOLM, S.C. and YTHIER, J.M. (2006), *Handbook of the Economics of Giving, Altruism and Reciprocity: Foundations*. Elsevier Science.
- LICHTENSTEIN, S. and SLOVIC, P. (2006), *The Construction of Preference*. Cambridge University Press.
- LOEWENSTEIN, G. and ADLER, D. (1995), 'A Bias in the Prediction of Tastes', *The Economic Journal*, 105, 929-37.

- LOEWENSTEIN, G., BRENNAN, T. and VOLPP, K.G. (2007), 'Asymmetric paternalism to improve health behaviors', *JAMA: The Journal of the American Medical Association*, 298, 2415-17.
- LOEWENSTEIN, G.F. and HAISLEY, E. (2008), 'The Economist as Therapist: Methodological Ramifications of 'Light' Paternalism', in CAPLIN, A. and SCHOTTER, A. (eds.) *The Foundations of Positive and Normative Economics. A Handbook*. New York: Oxford University Press.
- MADRIAN, B.C. and SHEA, D.F. (2001), 'The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior', *The Quarterly Journal of Economics*, 116, 1149-87.
- MARNEFFE, P.D. (2006), 'Avoiding Paternalism', *Philosophy and Public Affairs*, 34, 68-94.
- MEAD, L.M. (1997), *The New Paternalism: Supervisory Approaches to Poverty*. Brookings Inst Press.
- MILL, J.S. (1989), *On Liberty and other writings*. Cambridge: Cambridge University Press.
- MILLER, D. (1983), 'Constraints on Freedom', *Ethics*, 94, 66-86.
- MITCHELL, G. (2005), 'Libertarian Paternalism is an Oxymoron. Review Essay', *Northwestern University, School of Law Northwestern University Review*, 99, 1248-77.
- MLODINOW, L. (2008), *The Drunkard's Walk: How Randomness Rules Our Lives*. Knopf Doubleday Publishing Group.
- NAGEL, T. (1970), *The Possibility of Altruism*. Princeton: Princeton University Press.
- NUSSBAUM, M.C. (2003), *Upheavals of Thought: The Intelligence of Emotions*. Cambridge University Press.
- O'DONOGHUE, T. and RABIN, M. (2003), 'Studying Optimal Paternalism, Illustrated by a Model of Sin Taxes', *American Economic Review*, 93, 186-91.
- OLSARETTI, S. (2004), *Liberty, Desert and the Market: A Philosophical Study*. Cambridge University Press.
- RAWLS, J. (1971), *A Theory of Justice*. Cambridge, Massachusetts: Harvard University Press.
- RIZZO, M.J. (2009), 'Little Brother Is Watching You: New Paternalism on the Slippery Slopes', *Arizona Law Review*, 51, 685-739.
- RUBINSTEIN, A. (1998), *Modeling Bounded Rationality*. Mit Press.

- SARTORIUS, R.E. (1983), *Paternalism*. University of Minnesota Press.
- SCOCCIA, D. (1990), 'Paternalism and Respect for Autonomy', *Ethics*, 100, 318-34.
- SEN, A. (1991), *On Ethics and Economics*. Wiley.
- SHIFFRIN, S.V. (2000), 'Paternalism, Unconscionability Doctrine, and Accommodation', *Philosophy & Public Affairs*, 29, 205-50.
- SIMON, H.A. (1957), *Models of man: social and rational; mathematical essays on rational human behavior in society setting*. Wiley.
- SIMON, H.A. (1982), *Models of Bounded Rationality, Vol. 3: Emperically Grounded Economic Reason*. MIT Press.
- SINGER, T. and FEHR, E. (2005), 'The Neuroeconomics of Mind Reading and Empathy', *American Economic Review*, 95, 340-45.
- SUGDEN, R. (2002), 'Beyond sympathy and empathy: Adam Smith's concept of fellow-feeling', *Economics and Philosophy*, 18, 63-87.
- SUGDEN, R. (2008), 'Why incoherent preferences do not justify paternalism', *Constitutional Political Economy*, 19, 226-48.
- SUNSTEIN, C.R. and THALER, R.H. (2003), 'Libertarian Paternalism Is Not an Oxymoron', *The University of Chicago Law Review*, 70, 1159-202.
- THALER, R.H. (1980), 'Toward a positive theory of consumer choice', *Journal of Economic Behavior & Organization*, 1, 39-60.
- THALER, R.H. and SUNSTEIN, C.R. (2003), 'Libertarian Paternalism', *The American Economic Review*, 93, 175-79.
- THALER, R.H. and SUNSTEIN, C.R. (2008), *Nudge. Improving Decisions About Health, Wealth, and Happiness*. New Haven & London: Yale University Press.
- TVERSKY, A. and KAHNEMAN, D. (1981), 'The Framing of Decisions and the Psychology of Choice', *Science*, 211, 453-58.
- TVERSKY, A. and KAHNEMAN, D. (1986), 'Rational Choice and the Framing of Decisions', *The Journal of Business*, 59, S251-S78.

VANDEVEER, D. (1986), *Paternalistic Intervention: The Moral Bounds of Benevolence*. Princeton University Press.

WIKLER, D. (1979), 'Paternalism and the Mildly Retarded', *Philosophy and Public Affairs*, 8.