

# ATHENA

CRITICAL INQUIRIES IN LAW, PHILOSOPHY AND GLOBALIZATION


---

## Paternalistic Interventions: What do They Presuppose About Human Rationality, and When are They Justified?

MIGUEL FERNÁNDEZ NÚÑEZ

*Professor of Philosophy of Law, Autonomous University of Madrid (Spain)*

✉ [miguel.fernandezn@uam.es](mailto:miguel.fernandezn@uam.es)

 <https://orcid.org/0000-0001-7870-1817>

### ABSTRACT

In this article, I examine a pressing and recurring problem for liberal thought: paternalistic interference with freedom. I focus on its main premise: a deficit in the affected agent's deliberation that results in harm to the agent. I analyse how human beings are viewed by liberal legal norms — views about the effective rationality of normative subjects. I identify and characterise the four necessary and sufficient conditions to justify a paternalistic intervention. I outline several scenarios involving different defects in individual deliberation by identifying the agent's true self and basic interests. Finally, I contrast this repertoire with projects to build an “anthropology in law”, such as the one recently proposed by Moreso. My critical evaluation concludes that attributing “vulnerability” generically or hastily, as Moreso does, to the human collective, without precisely determining deliberative flaws, can lead to an excessive expansion of paternalistic interventions and an undue restriction of individual autonomy.

**Keywords:** paternalism, political liberalism, justifications of legal norms, legal presuppositions, rationality deficits, José Juan Moreso

This article is the result of a research stay at the Faculty of Law, University Roma Tre, where a previous version was discussed on 21st May 2025. This paper was written within the framework of the research project EcoGentium. Ecological justice and new vulnerabilities: global legal challenges (2024/00209/001), funded by the Agencia Estatal de Investigación (AEI). I would like to thank Juan Carlos Bayón, Ezequiel Monti, José Juan Moreso, Giorgio Pino, Michele Ubertaino and especially Giuseppe Rocchè and an anonymous reviewer from Athena for their comments on earlier drafts of this paper. Their suggestions were essential in enhancing the accuracy and depth of my article.

ATHENA

Volume 5.2/2025, pp. 142-185

Articles

ISSN 2724-6299 (Online)

<https://doi.org/10.60923/issn.2724-6299/22521>



*“Paternalistic intervention must be justified by the evident failure or absence of reason and will; and it must be guided by the principles of justice and what is known about the subject’s more permanent aims and preferences, or by the account of primary goods. [...] Paternalistic principles are a protection against our own irrationality, and must not be interpreted to license assaults on one’s convictions and character by any means so long as these offer the prospect of securing consent later on.”*  
J. Rawls

## 1. Introduction: The Problem of Paternalism for Classical Liberalism

### 1.1 Paternalism. What is it?

We can call “paternalistic” the intervention through which an agent (typically the state) interferes with another agent’s liberty, justifying the interference as being in the latter’s interest.<sup>1</sup>

Thus, they are the justifications (the intentions, the aims) behind the decisions that are paternalistic; such a decision may take the form of a moral or legal norm, a political or administrative measure.<sup>2</sup>

The standard normative effects of paternalistic norms and measures are obligations and prohibitions. However, there are also good reasons to include normative disabilities. For instance, the state’s refusal to confer legal validity on certain contracts, such as a contract of voluntary slavery.

A given justification is paternalistic if and only if it derives from the paternalistic principle.

---

<sup>1</sup> As long as an agent (hereafter, the state) interferes with an individual’s freedom of action, such interference is to be considered coercive. To be sure, this is “coercion” in a minimal sense. I will discuss these topics shortly.

<sup>2</sup> Therefore, it is clear that expressions such as “paternalistic interventions (decisions, norms)”, which are common in specialised literature, are metonyms (see Diciotti, 2021, 97; cf. Diciotti, 1988, 85). Nevertheless, for the sake of brevity, I often use them in this paper. Setting aside the metonym, a decision can be called “paternalistic” if and only if its sole or main justification is paternalistic.

Paternalistic principle<sup>3</sup> establishes that if an individual *X*'s action harms her interests, the agent *Y* (the state) should *pro tanto* intervene.<sup>4</sup>

### *1.2 Justifications of Decisions in the Context of Liberalism*

As stated, in this article I analyse the paternalistic principle within the normative framework of classical liberalism. By “classical liberalism” – or simply “liberalism” from now on – I refer to the ethical-political doctrine paradigmatically articulated by Mill, one of whose fundamental theses is precisely individual liberty and the principle of autonomy.<sup>5</sup>

As is well known, one of the main tenets of liberalism is that the state should guarantee individuals a domain of rights against invasion by others, and Mill's starting point is the defence of the individuals' “sovereignty” over their personal sphere, except for those actions that cause harm to third parties (Mill, 1864, 25). Nonetheless, Mill soon finds himself qualifying his initial theses towards paternalism, which itself highlights the problem we are concerned with (notably Mill, 1864, 172 ff.).

Very broadly, the ideal of autonomy — the idea that the individual should become “the author of her own life”, reflexively and free from significant external interference (primarily and above all from state coercion) — is the foundation of liberalism, as is individual liberty, insofar as it is instrumentally effective in achieving autonomy. This implies that liberalism requires specially qualified justification when it aims to restrict liberty (Moreso, 2021a, 650).

<sup>3</sup> Such a principle is, of course, ethico-political; therefore, moral norms are its paradigmatic embodiment. In this article, I do not specifically refer to “legal paternalism” (i.e., the paternalistic justification of legal norms) or to “moral paternalism”, assuming that the differences between them are secondary. Nevertheless, I will occasionally point out some differences when they are relevant.

<sup>4</sup> “Should intervene” means that such action is normatively required. Alternatively, we can phrase it as “the state has reasons for acting” or “the state is obliged to act”, and so on. The theoretical framework applied to this principle is not important to our concerns.

<sup>5</sup> I leave aside the fact that classic liberalism is a set of minimally heterogeneous proposals. This is an issue of the reconstruction of the history of ideas and can be ignored for the sake of expository economy. However, there is one point theoretically relevant to this paper: the contradiction between some theses put forward by Mill and Kant, forefathers of liberalism. I will address this issue specifically at the end of my essay (§3.4).

Indeed, liberalism adopts different views of a given intervention depending on its justification (see Feinberg, 1984, 26-27, for a similar presentation):

1. Harm principle establishes that, when the action of an individual (say Giuliana) harms another (say Gabriela), the state should *pro tanto* interfere with Giuliana's action. In other, more explicit words, if Giuliana causes or intends to cause harm to Gabriela and it is not permissible for Gabriela to suffer that harm, either if Giuliana's rights do not cover the harmful action or the normative considerations (whether substantive or procedural) in favour of preventive or reparatory interference outweigh those against it, then the state is all things considered justified in interfering with Giuliana's action.
2. Paternalistic and anti-paternalistic doctrines are concerned with the scenario in which Giuliana's action causes her harm. However, a distinction should be made based on Giuliana's rational capacities:
  - 2.1. Anti-paternalism establishes that, if she can discern, then *ceteris paribus* the state should not interfere with her action.
  - 2.2. Soft paternalism, instead, establishes that, if she is unable to discern, then *ceteris paribus* the state should interfere with her action.<sup>6</sup>

This holds in most cases (hence the *ceteris paribus* clause), but there can be exceptions.<sup>7</sup>

- 2.3. Genuine (hard or moderate) paternalism seeks to answer two questions: If she is formally and materially (at least normally) able to discern, but finds herself in a situation of (temporary or partial) material

---

<sup>6</sup>An exemplar of this is belonging to the category of *alieni iuris* subjects. The inclusion of a subject in a category of people who are formally incapable establishes the presumption that the subject is also materially incapable. This is a normative qualification which is often *iuris et de iure*, i.e., treated as indefeasible. Such a presumption is plausible as far as it is based on statistical generalisations, correlating the occurrence of certain empirical situations with a sufficient lack of relevant cognitive or practical abilities.

<sup>7</sup> Of course, it should be recalled that the question of formal (in)capacity is clear-cut – and, in addition, it must be so, as required by the material dimension of the legality principle – but this is not the case for material (in)capacity, as there is a continuum from radical incapacity to total capacity.

incapacity, should the state interfere with her action? Under what conditions?

As we will soon find out, [2.3] can be regarded as specifying the *ceteris paribus* clause in [2.1]: what changes when other things are not equal, namely, the conditions of rationality.

3. Anti-moralism establishes that, if Giuliana performs an immoral action, the state *pro tanto* should not interfere with her action. In other words, its immoral content is not in itself a basis for the state to intervene in a given action.

4. Anti-perfectionism establishes that, if Giuliana performs an action that contradicts a principle shaping the idea of “how Giuliana should become” – an idea not shared by Giuliana herself –, the state *pro tanto* should not interfere with her action.<sup>8</sup>

With the exception of [1], in which I specified the (usual) conditions under which the intervention measure should be taken all things considered, the remaining justificatory principles are merely *pro tanto*. Certain elements within the aforementioned definitions remain implicit and should therefore be spelled out. Furthermore, for these to become all things considered principles, the procedural and substantive considerations against state intervention must not outweigh or otherwise counterbalance the considerations in favour of state intervention. I will return to those that I call “background conditions”.

### *1.3 Where the Problem Lies: The Opportunity to Decide for Others, Based on Incapacity and Hypothetical Consent*

---

<sup>8</sup> As with paternalism, other forms of interaction between the state and citizens are certainly conceivable and feasible. The state may wish to persuade citizens – in a discursive, rhetorical, or dialogical sense, and therefore without resorting to coercion – to make certain choices so that they become the citizens they “ought to be” (whatever this may mean, which is not important here). Some persuasive strategies or some ways of exercising regulatory powers for the benefit of the competent agent herself (in a paternalistic or perfectionist sense) that are less forceful than coercion can be legitimate. However, this is a hypothesis I do not intend to explore in this work, except occasionally.

Let us return to the arguments listed above. Several observations should be made. Preliminarily, it should be emphasised that justifications can be made explicit, but they can also be implicit. Even when explicit, their formulation can be sufficiently indeterminate to allow for cumulative and alternative justifications, not to mention the possibility of manipulation by the interpreter. It is very easy for the same decision to be presented, either by the authority that promulgates it, that called upon to interpret it, or for the observer or critic, to identify, alongside paternalistic foundations, non-paternalistic foundations that are (or that are intended to be) complementary or alternative to the former.

In the second part of this article, through the critical analysis of the Lochner and Wackenheim cases, it will be sufficiently illustrated how one and the same judicial decision can be seen as conveying different justifications. In particular, we will see how the conceptual boundaries of paternalism with the other doctrines (namely moralism and perfectionism) are blurred; however, it is advisable to consider these definitional boundaries from the outset, as well as the autonomous conceptual profile that paternalism appears to have.

It is certainly difficult to distinguish perfectionism from unjustified paternalism, insofar as both apply standards to the subject that are foreign to her (cf. Alemany, 2006, 114-115, 162-163). On the other hand, perfectionism differs from justified paternalism because, as I will argue, only for the latter position, is “the view Giuliana has of herself” or “the view Giuliana would like to have of herself” important (cf. Hart, 1963, 30-31; Nino, 1989, 414, 427, 430-431; Alemany, 2006, 384-386). In contrast, [3] and [4] are positions that, *ex hypothesi*, exclude taking into consideration the agent’s motivational set.

With regard to [2], i.e. the paternalistic scenarios, in this work I focus on [2.3], which is the position that is most strongly or, more precisely, genuinely problematic for liberalism. [2.3] embodies the conceptual space occupied by what are generally called “hard paternalism”, “moderate paternalism”, or

“moderate anti-paternalism” (cf. Maniaci, 2025, 14-15). No doubt, we must determine to what extent the state should intervene, and thus which position regarding intervention is correct.

As for [2.1], as we shall see in a moment, it is a clear case of anti-paternalism for the liberal, i.e. a case where paternalism is not justified for the liberal. Regarding [2.2], the literature usually refers to such cases as “weak” or “soft paternalism” (Feinberg, 1986, 12 ff.; Alemany, 2006, 392 ff.). In a sense, it is a case where paternalistic intervention is clearly justified for the liberal. Indeed, in some of these scenarios, if we focus only on the limiting case where the individual is not autonomous at all, there are arguments to suggest that it is wrong to speak of “paternalism” if there is no autonomous exercise of liberty to interfere with. Given that, for the potential beneficiaries of soft paternalism, the assertion of a deficit of rationality and therefore the need to interfere with an embryonic, weak autonomy is more indisputable, the considerations applying to strong paternalism, i.e. the most problematic type of paternalism, will apply even more so (*“a fortiori, a maiori ad minus”*) to weak paternalism.

To better understand why [2.3] is the truly interesting position, it is worth examining the assessments in [2] in greater depth and considering their merits. Indeed, there is a connection between the incapacity of the subject and the consent that the subject can give or, ultimately, that can be considered valid. The underlying idea is that if we want to count a certain exercise of freedom as autonomous, it must satisfy the appropriate conditions (cf. Elster, 1983, 20 ff.). In [2.1], it is implied that Giuliana consents to the action she carries out; in [2.2], it is implied that Giuliana, given her incapacity, would not consent to her action if she were to abandon her state of incapacity.

The question in [2.3] is more difficult for two reasons. Firstly, unlike in [2.2], it is arduous to argue that the individual finds herself in a state of incapacity. Although she could be so, the incapacity must be qualified: what type it is, and how pervasive it is. For brevity, we can refer to this as the “problem of diagnosis”.

Nevertheless, the result of [2.3] aligns with that of [2.2]: *justified* cases of paternalistic intervention always presuppose some form of *rationality* deficit.

Indeed, the two questions that form the subtitle of this article are examined thoroughly, albeit far from exhaustively, and combined: the justification for paternalistic interference therefore always rests on some flaw in the agent's deliberation, along with other related, secondary conditions. Of course, it is important to note that not every deliberative flaw is sufficiently damaging; in a second step, it should be examined to what extent the flaw or the combination of flaws is relevant to justify paternalistic intervention. I will offer some indications regarding this question, especially concerning the interests affected by these flaws.

Incidentally, two presuppositions of paternalistic intervention must be explicated here: that the state knows better than the agent what is best for her well-being,<sup>9</sup> and that the state is precisely the one called to intervene.<sup>10</sup>

Concerning the potential alignment between [2.1] and [2.3], a further consideration is necessary. As argued in [2.1], it is presumed that Giuliana consents to the action she undertakes because, as stated, "Giuliana is able to discern". Nonetheless, it should be noted that the fulfilment of specific criteria is requisite in order to proceed. As Maniaci observes, moderate anti-paternalists allow individuals to freely pursue their well-being as they deem appropriate, provided they are

sufficiently rational [mainly, but not only in the sense of "coherent"], aware of the relevant facts, and sufficiently free from coercive pressures [and compulsion], and it must be the case that the desires

---

<sup>9</sup> As opposed to the classic "argument from paternalistic distance" (cf. Diciotti, 1986, 566). This claim will often prove false, especially since the state (e.g., the legislator) relies on empirical generalisations – exposed to exceptions and sometimes even refutable – and applies them to general and abstract categories of normative subjects. This caveat is of little importance: it is an easily granted and useful assumption from an explanatory-theoretical perspective for the purposes of normative arguments such as those examined and put forward in this paper.

<sup>10</sup> Indeed, while we can and should expect paternalistic decisions from a parent, the issue is whether the state should act as a "*pater*" towards its citizens.



they form are stable over time (see Maniaci, 2025, 5; cf. Maniaci, 2025, 13-15).<sup>11</sup>

Regarding the adjective, such an anti-paternalist is rightly labelled “moderate” because she recognises that certain conditions must be satisfied to respect the individual’s decision. Regarding the substantive, the anti-paternalist (even if moderate) remains anti-paternalist insofar as she presumes the rationality of agents and reviews it only when there is evidence of a potential deficit of rationality. I believe it is correct to assume the general rationality of subjects that belong to the class of [2.1]; nonetheless, in some problematic cases, we set aside this presumption and must carefully examine their competence. In those exceptional cases, we are already entering [2.3].<sup>12</sup> This does not deny the utmost importance, for liberals, of assuming individuals’ *pro tanto* material capacity in performing any action affecting only themselves. For the purpose of [2.1], *pro tanto* respect toward individuals’ self-referring decisions is required. Moreover, as I will show in this article, other conditions must be met to justify all things considered paternalistic interventions.

Finally, while in [2.2] it is undisputed that such consent is invalid – what is problematic, if anything, is whether the subject reaches a formal threshold –, it is debatable whether, in the absence of such incapacity (which is at issue in [2.3]) the subject would unfailingly consent to the intervention. The core task is to determine under which conditions the agent’s explicit consent to a supposedly free course of action *X* can be considered invalid or even non-existent, and under which conditions the subject would consent to the

<sup>11</sup> For a detailed disclosure of such exigences, see Maniaci (2025, 89-106).

<sup>12</sup> A potential distinction between Maniaci’s perspective – and other analogous perspectives, such as Alemany’s (2006, 422-423) – and my own pertains to the fluidity of the scenario delineated between [2.1] and [2.3]. This is due to the fact that, as Maniaci would acknowledge, it is possible to make deliberative errors of a significant nature without being formally incompetent, i.e. [2.2] (cf. Maniaci, 2025, 94-95). While my position may appear to adopt a more paternalistic stance than that of Maniaci when considered within the confines of the initial conditions for intervention, this is not the case when the procedure is continued to its culminating conditions and the various caveats attached to each condition are taken into account.

paternalistically interfered course of action *Y*. There is a phenomenology of deliberative elements involved, as well as a series of symptoms, that lead us to identify this consent (primarily, the basic interests). We can refer to this as the “problem of hypothetical consent”.

#### *1.4 Overview of the Paper: The Four Necessary and Sufficient Conditions for Paternalistic Intervention*

In the first part of this paper, I intend to address the problem of decision-making on behalf of others (hereafter referred to as “hetero-decision”) and, particularly, the problem of diagnosis and the problem of hypothetical consent. After examining the definitional questions above, I will focus on the problem that hetero-decision in the agent’s interest poses for liberalism. Next, I will show that the first step in developing and analysing a procedure for paternalistic intervention is to balance well-being and autonomy correctly. For paternalistic interference with freedom to be justified, well-being must outweigh or otherwise counterbalance autonomy. I will then discuss the connection between the subject’s well-being and the need for paternalistic intervention, developing a brief phenomenology of flaws in deliberation and the most important types of interests that can be affected by them.

In other words, I will present a brief case study of situations in which an individual is incapacitated, considering the more and less problematic aspects of paternalistic intervention in these situations. Only when the agent incurs some rationality deficit can intervention with her freedom of action be justified. This is the second step, or necessary condition, of paternalistic intervention.

In particular, I will explore a way to articulate the conceptual link between the subject’s well-being and the consent they would give to paternalistic intervention. Determining the affected agent’s “true self” and identifying it with the agent’s basic interests constitutes the third step or necessary condition for paternalistic intervention.

The fourth step consists in satisfying the “background conditions”, that is the substantive and procedural factors in favour of intervening itself should outweigh the opposing considerations. I will set out several considerations that give shape to these background conditions.

In summary, I aim to show that a paternalistic intervention is justified if and only if: [1] it results from correctly weighing well-being against autonomy (that is, if well-being prevails, while autonomy is rightly outweighed or otherwise counterbalanced, or not ultimately sacrificed, even if it appeared so *prima facie*); [2] it arises from one or more scenarios of rational deficits in a sufficiently damaging way; [3] it is consistent with the agent’s motivational set, particularly, in her “true self” and basic interests; [4] the procedural and substantive reasons for intervening outweigh those against intervening. Thus, these four steps, or sets of criteria, are what a procedure of paternalistic intervention must satisfy to be justified. I will address [1], [2] and [3] at some length, though not exhaustively or with the ideal depth, and I will provide several indications to determine [4].

In the second part of the essay, I will examine the attempt to build an anthropology in law. By “anthropology in law”, I refer to the assumptions of rationality in legal norms, and the explicit or implicit representations in norms of human beings as more or less rational agents.<sup>13</sup> I will also examine how certain proposals seek to defend paternalism on the basis of a supposedly realistic (true or plausible) anthropology that the law must consider. I will focus on the recent proposal by José Juan Moreso, by pointing out how Moreso’s proposition is affected by several problems.

In general, and preliminarily, the very project of constructing a single anthropology in law can be considered questionable and, indeed, misleading. Instead, there is a theoretical need to clearly define and list cases of rational vulnerability; there is also a theoretical and practical need to examine

---

<sup>13</sup> These may be the (primary or secondary) addressees of the legal norms. The status of these statements which make up anthropology in law appears to be hybrid: in part, they are apophantic statements, concerning how normative subjects actually are; in part, they are evaluative statements, concerning what treatment should receive them.

remedies for such cases of vulnerability in a singularised way. As they are descriptively inadequate, and as they are theoretically and practically counterproductive, it will be necessary to reject those overly general anthropological accounts based exclusively on the image of a fully capable subject – such as defending the extrapolation to any area of law of the assumption of a *sui iuris* subject with full capacity to act, as a contractual party under the Civil Code –, and those overly general anthropological accounts based on the image of a (generically) incapable, fragile subject. Moreover, as we shall see, Moreso’s conclusion is based on a generic reconstruction of the justification of two judicial decisions and on a problematic recourse to the notion of “human dignity” that are too advantageous for paternalism.

## 2. Autonomy and Well-being, Selves and Typologies of Interests

### 2.1 First Step: The Correct Balance of Autonomy and Well-being

If we return to the initial definition of “paternalism”, i.e. if a paternalistic intervention consists in interfering with the freedom of action of a given agent (the agent’s right to decide as she pleases) for her benefit,<sup>14</sup> then we can reconstruct the first analytical step of scenarios of paternalistic intervention as the competition and, often, conflict between two values: autonomy and well-being.

We must therefore distinguish between two main possibilities: 1. that there is, all things considered, a concordance between well-being and autonomy, and 2. that there is, all things considered, a conflict between well-being and autonomy; 3. a third, marginal possibility is that, all things considered, there is a conflict between autonomy and autonomy.<sup>15</sup>

<sup>14</sup> These notes embody the primitive notions of many of the most relevant definitions of “paternalism”, such as those by Dworkin (1972, 175), Buchanan (1978, 371), Diciotti (2005, 99), Alemany (2006, 345, 352), and Maniaci (2025, 5).

<sup>15</sup> There is also a fourth possibility that, all things considered, there is a conflict between two components of well-being. I overlook this option because it partially overlaps with the second and third possibilities and, more importantly, can obscure the conflictual nature of the situation. The crucial point to keep in mind is that the two values differ in content, not that

1. If there is all things considered concordance between well-being and autonomy, there is no problem of justification; it will only be necessary to articulate how well-being and autonomy are reconciled, especially when *prima facie* seemed otherwise.
2. On the other hand, if there is all things considered conflict between well-being and autonomy, we must determine, from a consequentialist perspective, which solution should be adopted.
  - 2.1. If considerations relating to the well-being of the individual outweigh or otherwise counterbalance the value of autonomy, the state should *pro tanto* interfere with the individual's freedom of action.<sup>16</sup> To put it another way, in such cases the harm the individual would suffer from her free action is too serious (her action is not proportional), and therefore it is better to interfere with that action.

This appears to be a good reconstruction, for example, of the obligation to wear a helmet when riding a motorcycle in countries where such a mandatory legal rule applies. It makes sense to argue that this obligation serves (i.e., has the function of and is suitable for) protecting the motorcyclist and that such protection is more important than the freedom the motorcyclist might otherwise enjoy, such as the value of feeling the wind in her face, not being overheated by the helmet, or, more generally, the value of deciding as she sees fit. This case, as formulated, is an example of justified paternalistic intervention, and is certainly compatible with liberalism. This is especially true as the interference with the agent's freedom can be deemed minimal and is instrumentally necessary (there appear to be no equally suitable measures less invasive of individual freedom) for her well-being.

In such cases, the motorcyclist, as a conscious, extreme anti-paternalist agent, may ultimately consider that a proper balance is achieved when the

---

they are similar in kind, as various specifications of the same general value (as, for instance, a utilitarian or another defender of value monism would claim).

<sup>16</sup> I assume the activity of resolving a conflict between values by means of an axiological principle of weighing, in the sense of Chiassoni (2019, 168). As the procedural criteria for this, I use the traditional four proportionality test categories, albeit somewhat loosely.

pleasure of feeling the wind in her face (or similar considerations) outweighs the risk to life and physical integrity. I think there are various reasons to consider such balancing irrational. It is irrational if we focus on the likelihood of the occurring harmful event.<sup>17</sup> It is also irrational if we focus on the balancing of values or the weight assigned to the values themselves.<sup>18</sup>

In fact, Western legal systems enshrine various measures with a paternalistic justification, and it can be argued that this justification often does not contradict classical liberalism, as the necessary balance must be achieved and implies only a proportional and sometimes minimal sacrifice of freedom of action, required to improve or, more often, preserve the well-being of the agent affected.

The earlier statement that “considerations relating to well-being outweigh” or “carry more importance” is a somewhat hasty formulation and may appear overly favourable to consequentialism. It can also be argued, in a more deontological vein, that it is a matter of guaranteeing a set of goods considered

---

<sup>17</sup> For example, suppose the motorcyclist is not extremely prudent and does not drive at 10 km/h on a completely empty road, but is only moderately prudent, driving at 60 km/h on a road where other vehicles sometimes circulate. In such conditions, if she considers the statistical probability of a road accident to be extremely low, we can consider her irrational as she clearly underestimates the likelihood of such an event. Depending on the formulation, this is likely to be a case of statistical fallacy, or an inductive fallacy, concerning either the sample on which the generalisation is based or the generalisation itself. If she disregards the statistical frequency, out of excessive optimism or nonchalance, she could be accused of wishful thinking. Finally, if she does not even consider (at least intuitively, preliminarily) the statistical frequency of the event occurring, she can be deemed epistemically irresponsible. These are some of the possible charges against the motorcyclist concerning statistical frequency. For a different solution to a similar scenario, cf. Maniaci (2025, 52-54).

<sup>18</sup> Indeed, the reproach may emerge from the assignment of a different weight to the two values to be balanced. Consider that the motorcyclist prioritises the experience of freedom (the wind in the face, self-determination) over her well-being (her physical integrity, her life). This can be assessed in an objective or subjective manner. Objectively, by ascertaining that a rational plan of life cannot include a choice involving a significant likelihood of being severely injured, traumatised, or (absolutely irreversibly) killed. Subjectively, by determining that such a possibility is not included in the life plans of most individuals (the percentage of people who, after a period of deliberation and exposure to a minimum amount of relevant information, would choose this is likely to be very low). This argumentation can be based on the agent’s motivational set.

basic, which, regardless of their consequences or competing values, must be incorporated into *any* life plan an individual wishes to pursue.<sup>19</sup>

Some authors consider that autonomy (even in the form of freedom of action, which enables autonomy) is not merely another value to be weighed against well-being, but a side-constraint. This could be especially the case if in attributing a kind of lexical priority we are not thinking about any freedom, but the most important ones, such as those pertaining to the Rawlsian first principle of justice (Rawls, 1999, 52 ff.). Despite the apparent soundness of this solution, and the fact that a comprehensive examination of this issue exceeds the possibilities of this paper,<sup>20</sup> some perplexities emerged when a strictly deontological perspective on freedom is adopted. In cases such as that of the motorcyclist, we are at least implicitly, and preliminarily, comparing the relative importance of well-being. So, even if we do not put autonomy (or freedom) into a balance with well-being, because we consider that the very balance is precluded, in some sense we are presupposing a sort of comparison between the two values (a “pre-balance”, so to say). Additionally, as I have already pointed out, it is difficult to completely isolate autonomy from well-being, as two discrete values (even if most of my reconstruction of scenarios and thought experiments, for the sake of expository economy, does so). It can be argued that, in relation to the first point, and to a certain extent the second, that only nozickian libertarianism posits the liberty of the will as an absolute value that is not subject to evaluation and, consequently, will prevail over all other values. Even over the well-being of the very owner of that unproblematised freedom. This approach is not sound, as has been demonstrated in part.

## 2.2. There may also be cases where autonomy outweighs well-being.

For example, in the previous scenario, the motorcyclist does not wear a

<sup>19</sup> Also, if such a framing is plausible, this shows that, alongside the subjective element of satisfying preferences, we can identify an objective element (the so-called “objective list theories”).

<sup>20</sup> Specifically, I cannot address here the more fundamental question of whether a conclusive distinction or assimilation between the structures of moral deliberation of a consequentialist sort and those of a deontological sort can be traced. See Bayón (1989).

helmet because she wishes to commit suicide and considers this an ideal way to do so. Nonetheless, the *ceteris paribus* clause must be satisfied: the motorcyclist has made her decision as a result of a reflective process and stable wishes, with all pertinent information being made available to her and without the presence of external coercive pressures.<sup>21</sup> Again, the *ceteris paribus* clause also translates the assumption that the agent does not incur deliberation flaws, as we will see soon. Should autonomy prevail over well-being, and should the *ceteris paribus* clause be satisfied, it can be concluded that paternalistic intervention would not be justified.

3. In certain cases, autonomy (or freedom) conflicts with another instance of freedom, necessitating its primacy over freedom itself. For instance, the invalidity of voluntary slavery contracts can be conceived as a case of [2.1], or as a situation in which the subject's present freedom is limited to guarantee a future freedom that she could not enjoy if the choice to become a slave were granted (Mill, 1864, 184-185; Maniaci, 2025, 220-232).<sup>22</sup>

Therefore, the first condition for a justified paternalistic intervention, or the first step in the theoretical reconstruction of such an intervention, is to proportionally sacrifice autonomy for the sake of well-being. Alternatively, in some cases (such as a voluntary slavery contract), it may involve overlooking a hollow, *prima facie* version of freedom of action or autonomy, in favour of a sound, conclusive version of autonomy.<sup>23</sup>

Regarding the consequences of the actions in general, it is sound to establish the principle that, *ceteris paribus*, the more severe is the (risk of) harm, or the more irreversible is harm, the more the rational capacity should

<sup>21</sup> While the legal obligation still applies, from a critical moral standpoint, it can be argued that the agent's autonomy or other preferences outweigh or otherwise counterbalance those that justify the mandatory use of helmets for the general class of motorcyclists under normal circumstances.

<sup>22</sup> This is a case of positive freedom, as I will argue later (§2.5).

<sup>23</sup> It is important to be cautious when extrapolating this reconstruction, which is valid for cases such as voluntary slavery contracts, to other contexts, as it can be used as an apparently liberal approach while serving to disregard autonomy. For a careful discussion of important caveats see Colomer (2004, 176-180).



be required on the agent (apparently or effectively) willing to undergo such a harm (see Feinberg, 1986, 118-121; cf. Alemany, 2006, 404-405; cf. Maniaci, 2025, 43). Concerning the results of the balance, we can conclude that, on a continuum between anti-paternalism and paternalism, the most anti-paternalistic position is the one that gives autonomy lexical priority (and indeed, though this is highly problematic, does so without even considering the countervalue of well-being), while the most paternalistic position is the one that systematically disregards autonomy in order to prevent any harm (even minor) to well-being.

This is undoubtedly only the first step in a complex procedure, and even this initial step must be further qualified. Regarding autonomy, we must determine whether and to what extent consent is valid or could be obtained (that is, what would constitute a correct exercise of autonomy). Regarding both autonomy and well-being, whether and to what extent the conception of well-being is connected to the agent's more stable and most defining preferences, and a more complex or conclusive autonomous "true self". Both issues will be tackled together in the following section.

## *2.2 Addressing the Problems of Diagnosis and Hypothetical Consent*

We should now examine some prominent and recurring ways in which an agent can harm her well-being due to a deficit of rationality. Without excluding analogous situations, and without claiming to be exhaustive, I list situations in which an agent may find herself deciding a course of action and which support the view that the state should *pro tanto* intervene, when the deliberative flaw causes genuine and significant harm (all things considered, if the other conditions are also satisfied).

As we will see shortly, without necessarily adopting a Humean model of individual action – some shortcomings of which will soon become apparent –, it is important to distinguish, in intentional action, between beliefs, desires, and capacity for action. This already provides us with three distinct elements

constituting practical deliberation, each of which paternalistic intervention can affect.

We can establish a taxonomy and a structure of the elements affected by some rationality deficit. Depending on the deliberation component and the severity of the disorder or disorders, and the basic character of the interests involved, a greater or lesser need for state intervention is required.

Concerning the notions of “belief” and “capacity for action”, if taken in isolation and in a rather uninteresting sense (i.e., as lack of information and as lack of factual or normative capacity), these are two assumptions that are not particularly problematic in terms of justification. The state can assist the individual if they lack these two resources: by informing or empowering the citizen, respectively. In this minimal sense, these may not even be considered cases of paternalism: they lack the coercive element and, more fundamentally, they lack the hetero-decision element. In the first case, the state is merely helping the agent to freely make the decision; in the second one, the state is only supporting the decision the agent has freely made.<sup>24</sup>

More interesting, because evidently problematic, is the case in which desires and beliefs have some defect.<sup>25</sup>

### *2.3 Second Step: The Problem of Diagnosis. List of Rationality Deficits*

Preliminarily, we must exclude cases of unintentional (not deliberate, conscious) actions. When the issue concerns what the individual desires and believes, we may encounter various scenarios. For example, we can identify the following “disorders”, described in familiar terms in philosophy, which

---

<sup>24</sup> To be sure, even in these cases some salient problems remain, such as how to determine which information is relevant for the agent to take into account (consider, for instance, the epistemic responsibility of the agent herself), or to determine which options should be available.

<sup>25</sup> In fact, the Humean model is, at best, limited and its compartmentalising aspirations can be deemed as misleading, if we take into account considerations now commonplace among cognitive psychologists, such as the observation that there do not appear to be any mental states devoid of emotional charge.

often have corresponding manifestations of a clinical (especially, psychopathological) disorder.

Cases of more distinctly emotional problems, or of mixed epistemic and emotional problems, could be the following:

[A] Cases of weakness of will, or *akrasia*. Here, the agent believes that  $\varphi$  is the correct action to take and, to some extent, wants to  $\varphi$ , but is subject to impeding factors that prevent her from  $\varphi$ .<sup>26</sup> To this extent, there is no imposition of interests, but rather stimuli towards an already chosen option.

Instances of *akrasia* include severe addiction to substances and behaviours, which can erode the brain's ability to control desires and emotions. For example, narcotics act on the brain's reward system by increasing the quantity and duration of dopamine, causing individuals to refrain from choosing the genuinely preferred alternative option. Another significant case of weakness of will occurs in major depression, due to neural hypoactivity in the prefrontal cortex and hippocampus, which causes individuals not to dare to implement the preferred option.

Of course, minor addictions and depression can be superficial or transitory enough to be negligible as sources of harm to the agent's well-being, while it is cases of major addictions and depression that can be truly paralysing and invalidate genuine choices.

[B] Cases of what has traditionally been called "motivated irrationality", such as self-deception and wishful thinking.

In both cases, the agent – who is normally competent to evaluate evidence and sensible to reasons – holds patently false beliefs, that contradict easily available evidence and does so because of the desirability of the content of those beliefs (see Pedrini, 2013, 4; Mele, 2016, 132).

Self-deception and wishful thinking, in their milder forms, are also widespread phenomena, but among their manifestations we can find truly

---

<sup>26</sup> I am of course simplifying. For a detailed reconstruction, among others, see Kalis, *et alii* (2008).

problematic cases, where the agent's very identity or the main features of her environment are greatly distorted.

Cases of more distinctly epistemic problems could include:

[C] Cases of empirically based inferential problems, such as an erroneous identification of the causal course, the risk involved by the action or one or more of the main effects of one's action.

The aforementioned deficit manifests most acutely in certain individuals diagnosed with psychotic disorders. These individuals often demonstrate an inability to integrate risk-related information, which hinders their capacity to make optimal decisions and which precedes the formation of risk preferences (cf. Purcell, *et alii*, 2022).

[D] Cases of strictly evaluative inferential problems, such as an erroneous balancing of interests.

For instance, not attributing to some interest at the time of action the importance that the agent assigns to it in a more considered deliberation (cf. Alemany, 2006, 137). For more details, I refer the reader above (§2.1), where I critically analyse an example of this kind of error, regarding the extreme anti-paternalist motorcyclist.

It is imperative that some concluding distinctions are made. The first of these relate to the type of condition that could harm significant interests of the holder. The second issue pertains to the nature of the intervention that such a condition might necessitate, namely, an alternative intervention that may be more suitable and less invasive.

It is evident that certain of these deficits (of which a non-exhaustive list is provided here) manifest with such incidence that they do not compromise well-being; in other instances, they occur with a frequency that can be detrimental to the well-being of the agent in question. The diagnosis itself and its impact on decision-making and clinical solutions must be carefully examined. Neither of these deficits needs to be present to the highest degree to be disabling and sustain paternalistic state intervention (for instance, a minor obsessive pattern can paralyse important choices). Conversely, a deficit

at the highest degree cannot necessarily invalidate decisions (for instance, a person with major depression can still adopt appropriate decisions).

Even if the deficit is effectively disabling, this alone does not provide sufficient justification for state intervention in the face of such a rationality deficit. Most notably, it can be argued that paternalistic state intervention is not necessarily the least harmful or most appropriate criterion for addressing and solving these deficiencies in rationality. This might occur *ex post*. In certain cases, alleviating purely circumstantial and transient conditions (e.g., a shock, alcohol intoxication) enables the agent to engage in a cognitive process unencumbered by bias. In other cases, the problem may be alleviated or resolved through psychotherapeutic or pharmacological treatment. Alternatively, patients may be referred for an interview with a doctor who can help them identify and modify their defective psychological processes and certify the soundness of the new unbiased deliberative process (see Maniaci, 2025, 207). Or this might occur *ex ante*. In certain cases, a safety procedure, such as a waiting period before making a decision, can be implemented to ensure the decision is based on an accurate understanding of the relevant facts, the stability of desires and the lack of coercion or compulsion (see Maniaci, 2025, 109, 114-115).

Nevertheless, it is acknowledged that the effectiveness of some of these solutions may be questionable, as certain individuals in certain conditions may be partially or wholly unresponsive to psychological or systemic therapies, or impervious to the control, or safety strategies. Indeed, biases can be pervasive and permanent,<sup>27</sup> and self-deception can be so sophisticated that it deceives both the agent and the expert (cf. Maniaci, 2025, 207-208).

---

<sup>27</sup> In fact, as far as I can see, Maniaci's views on the extent of systematic cognitive errors are somehow limited. These errors can sometimes affect even basic cognitive competences, and even experts in those fields (*contra*, Maniaci, 2025, 189-190). Maniaci's rejection of the pervasiveness and severity of biases is based almost exclusively on the rejection of some arguments of Kahneman and Tversky's work (Maniaci, 2025, 187-191), rather than on other studies addressing this issue. To be sure, Kahneman and Tversky's studies in particular can be challenged, and even for reasons beyond those mentioned by Maniaci, such as serious replicability issues (see Schimmack, *et alii*, 2017). Nevertheless, many subsequent studies inspired by these researchers are conclusive, and the wide-ranging and incidence of many

Apart from these considerations, the problematic nature of these cases differs. In fact, we can focus on [B], as it is the most problematic case and, in a certain sense, the paradigmatic one for paternalistic intervention, all for the same reason: such cases can be reconstructed as an internal opposition (not necessarily a contradiction) within the individual. In these cases, the agent appears to be assailed by conflicting beliefs and desires. The idea, as will become evident in the following discussion, is to set the subject *prima facie* against herself and, all things considered, in favour of herself.

#### *2.4 Third Step: The Problem of Hypothetical Consent. The Structure of Well-being and Basic Interests*

It is clear from the outset of this work that for liberalism, given the importance it attaches to autonomy, it is crucial to obtain the subject's consent to paternalistic intervention. A fundamental difference between paternalism (justified and unjustified) and perfectionism or moralism is that the latter two positions *ex hypothesi* are not interested in obtaining the consent of the normative recipient. In turn, within paternalism itself a distinction must be drawn between justified and unjustified paternalism, depending on whether consent is (finally) obtained or could be hypothetically obtained or not regarding the paternalistic decision. Preliminarily, if the intervention is based already on the agent's consent, then it lacks the coercive element and, most fundamentally, the hetero-decision element; therefore, for definitional reasons, the decision cannot be considered paternalistic. First, in some cases, e.g. when a transitory affliction recedes, the affected agent mediately gives consent to the paternalistic intervention. Second, sometimes factual consent cannot be obtained, but it is clear that it would be effectively given<sup>28</sup> or, more

---

biases are now widely accepted. Conversely, I am not suggesting that we should place undue emphasis on such errors in studies of paternalism, thereby risking overestimating their significance.

<sup>28</sup> Consider, for example, an individual who is aware of all the possible consequences of a debilitating disease is like to be diagnosed with and who leaves a detailed advance treatment statement for the eventuality he contracts the disease. The possibility may arise that the patient changes her decision, but this is quite another problem.

difficultly, that it would be rationally given. In such cases, rational-hypothetical consent is required with the assistance of empirical proof.

If this is the case, we must have some means of identifying the basis and the evidence for the existence and content of that consent; otherwise, we risk undesired effects, such as imposing normative standards on the subject that she does not share, as perfectionism and moralism do, or producing outcomes where any change could be self-confirming, as in cases of brainwashing. Indeed, we should seek a liberal response that does not involve deference to overly demanding justificatory theories (as perfectionism would require: “if the agent were virtuous, she would consent to the intervention”) nor excessive deference to the (partially) irrational subject.

It seems to me that the most suitable strategy for remaining within the bounds of justified paternalism is to refer, in these cases, to the motivational set of the subject involved, i.e. the set of beliefs and preferences the subject actually holds. Two strategies are combined here: the search for the true self and the identification of basic interests.

The first strategy is similar to those of authors who, in cases of individuals affected by persistent and severe addiction to narcotics or severe personality disorders, distinguish two selves: a true self, aiming to develop life plans, and a more superficial self, afflicted by invalidating circumstances (cf. Garnett, 2017, 8 ff.; cf. Husak, 1992, 71 ff.). I cannot pause to examine the merits of these strategies; suffice it to say that they are somewhat metaphorical and can be misleadingly used.

The second strategy involves identifying the subject’s most basic interests.<sup>29</sup> The two strategies can be fruitfully combined, by identifying the agent’s true self with her most basic interests. In cases where there is a conflict between basic and non-basic or less basic interests, it is reasonable to give priority to the former. Interests that have special final or instrumental importance in the agent’s life project can be considered “basic”. As I

---

<sup>29</sup> I presuppose Frankfurt’s strategy of stratification of desires and the importance of second-order desires. Cf. Álvarez Medina (2018, 18-19, 44).

anticipated, some basic interests can be found in the Rawlsian primary goods, which are important to any life plan the agent might wish to undertake (Rawls, 1999, 219-220). Other basic interests (which partially overlap in content with the former) are those that shape the agent's self-conception, based on the wishes the agent has deliberately acquired, chosen or modified – or chosen to maintain, even if she could reject them – (cf. Raz, 1986, 290-292) or the idea the agent would like to have of herself, based on second-order wishes – the desires the agent wish to have – (cf. Elster, 1983, 21; cf. Russell, 1999, 71-78). The paternalistic intervention must opt for the true self rather than the superficial one (affected by some invalidating situations), and for the basic interests rather than the less basic ones. Therefore, it can be reasonably posited that, provided we remain within the confines of basic interests, the subject would be in a position to consent or would have compelling reasons (without relinquishing her motivational set) to hypothetically consent to the paternalistic intervention.

How can we obtain actual empirical evidence that the subject would consent, without our argumentative shift becoming an insidious means of imposing interests on the subject that she does not share or will not share? There should be some empirical evidence for hypothetical consent (or, mediately, for the identification of the true self or basic interests): for example, when the subject finds herself tormented by decisions that frustrate what could be considered her basic interests, or when she regrets the options not taken, those that align with the view she wishes to have of herself, and so on. We might, for instance, examine the brain scan or psychological report of a person whose decisions are based on adaptive preferences. To be sure, there may be cases where none of this empirical evidence is present, where the subject fails to recognise (almost) any of her basic interests, such as cases of brainwashing or severe persistent dependence on narcotic substances that radically invalidate executive functions in the subject's life organisation, but these are tragic cases to which the model I am proposing cannot even attempt to respond.



I believe that establishing a connection with the agent's motivational set is fundamental in determining the requirements that must be met for a paternalistic intervention to be justified within the framework of liberalism. This enables the rational deliberative process (embodied in hypothetical consent) to be anchored in some of the individual's preferences, without imposing standards that are entirely foreign to her.

Of course, the content and articulation of the entrenchment of such basic interests must await a new, more detailed formulation. Nonetheless, I wish to make explicit some points of my proposal by distinguishing it from those theories framed in terms of altogether "subjective interests", as opposed to "objective interests".<sup>30</sup> As far as I can see, an articulation in terms of "basic interests", some of which are attached to the agent's motivational set, is richer and more theoretically powerful than the alternative strategy, for various reasons.

In general, I do not think the issue is to decide between a completely endogenous set of the agent's preferences and standards (let us call this "subjective", as do theorists such as Maniaci) and a completely exogenous set of preferences and standards, alien to the agent (let us call this "objective", as theorists such as Maniaci do).<sup>31</sup> Rather, in evaluating a decision as correct, the liberal should seek a normatively relevant element, intersubjectively justified (let us call this "objective"), which must also be anchored in some way to the agent's motivational set (let us call this "subjective"). However, this point remains overly general; I think much attention should be devoted to the ambiguity of the terms "subjective" and "objective" when paired with "interests", to see why in any of these senses (or dimensions), it is reasonable to refer at least partially to the objective element.

---

<sup>30</sup> Maniaci (2025) is a paradigmatic example.

<sup>31</sup> Maniaci (2025, 4) refers only to "the objective interests" as being completely decoupled from the agent's motivational set and from the agent herself and portrays them as "an ideal interests conception [which], I submit, is prescientific, premodern, deeply deceptive, and absurd" (see also Maniaci, 2025, 5-6). The question is whether the notion of "objective interest" can be used independently of such a radically idealist conception as depicted by Maniaci; of course, my impression is that it can.

Firstly, insofar as we require the agent to follow an inferential path so that their desires and beliefs are rational, not only in a logical-deductive, structural sense (requirement of non-contradiction, and so on) – a “thin theory of rationality”, in Elster’s terminology –, but also in the content of the premises of reasoning themselves (reasoning free from adaptive preferences, what count as a coercive pressure, etc.), we are already entering the realm of the rationally correct, which is material, substantive (in this sense, “objective”), and thus not deferential to the subject (in this sense “subjective”). The latter is clearly what Elster meant when he formulated the “substantive rationality of desires” and identified it with the “right causal formation process”: autonomous desires are only those that have been deliberately chosen, acquired, or modified (see Elster, 1983, 20, 21). So, in the first sense, it is possible to refer, albeit loosely, to the idea of “objective interests” when referring to the normative – whether cognitive or substantive – standards that a community uses to evaluate its members decisions.

Secondly, as I develop at several points (when referring to Rawlsian primary goods and objective list theories), in some sense the very conception of what constitutes the “well-being” that should be assured is, at least partially, objective, in the sense of intersubjectively valid normative standards. Of course, it is not totally objective, if that means excluding any anchorage to the agent’s motivational set.

Thirdly, in some of these cases, we are referring directly to correct decisions and, in fact, often to norms of critical morality. In this dimension and to this extent, it would be inadequate, by definition, to consider that we are not dealing with interests that are in some sense “objective” (at least as part of universally valid norms).<sup>32</sup>

The second and third points can be rephrased. Respecting the subject’s decision, which reflects their motivations, is one thing; considering why it is justified (or autonomous, if you prefer), or why it is not justified and why an

---

<sup>32</sup> In addition to what I have pointed out in the note on “moralism”, for these last two points I refer to what I have developed in Fernández Núñez (2023, 414 ff.).

alternative decision and intervention are justified, is another. If we must defer to the decision, we can speak of “subjective elements”, but if the decision can be criticised or, in any case, when it is evaluated, we can refer to “objective elements”. There is a threshold, which we can conveniently call “objective”, as its standards are intersubjectively shared, from which we evaluate the decision to determine whether it is justified.

### *2.5 A Suitable Interlude. Why Prohibitions and Criminal Sanctions are not the Central Problem. Hetero-decision and Varieties of Coercion*

In the introduction, specifically in §1.3, I explained why I regard hetero-decision as the most basic and acute problem that paternalism poses for liberalism. Here I attempt to clarify where the problem does not primarily lie, though it does so at a second step. This concerns the problem of coercion, especially when understood in terms of prohibitions and criminal sanctions. I hope, however, to have indirectly demonstrated a part of this thesis. Indeed, we have not needed to pay much attention to coercion until now when solving the scenarios at hand, but as we approach the background conditions, it becomes evident and more relevant as a problem.

Several insightful contributions about paternalism tend to divert attention from the problem of hetero-decision and focus it elsewhere. In particular, they often emphasise the coercive element, not only in general terms but also in terms of punitive and, moreover, criminal law, suggesting that this is one of the real problems, if not the real one.

This is the case, for instance, with Nino (1989, 420 ff.) and Monti (2019), to cite a pair of contributions that are at least suspicious of paternalism, or Thaler and Sunstein (2009, 4-6), to cite a contribution that instead tries to “sugarcoat” paternalism. The case of the latter authors is particularly significant: paternalism is presented as liberal or libertarian to the extent that nudges, undoubtedly, do not involve coercion. However, the conclusion may change to some extent if we pay attention to what I believe to be the most pressing and basic problem of paternalistic interferences: hetero-decision.

The notion of “hetero-decision” is presupposed by the paternalistic principle I formulated at the beginning of this article: agent *Y* intervenes for the well-being of subject *X* and does so because she has already *decided* for that subject.

The arguments of such authors and my response to them can be formulated as follows: they claim that coercion is the element that distinguishes genuinely or problematic paternalistic interventions from those that are not (which is certainly true),<sup>33</sup> but also that coercion is the real or primitive problem of paternalism (which is false), and particularly – if not exclusively – in the form of sanctions and criminal penalties (which is false). From this, they conclude that the real issue is not so much when and why one agent should decide for another, which is overlooked as the basic problem or at least not considered sufficiently important (which is false).

Instead, the significant justificatory question lies not only, or even primarily, in coercion, and certainly not in punishment: it is quite clear that for liberalism is highly problematic to punish an individual (especially through criminal sanctions) for that individual’s own good.

Of course, coercion and its modalities are highly relevant, for determining either whether something is a paternalistic intervention or whether it is justified. Regarding the assumptions of coercion, a question that easily arises is what conception of “freedom” is adopted when we refer to “interference with freedom” in paternalistic intervention. The tendency or reluctance to identify scenarios as paternalistic or non-paternalistic, or to qualify them as justified or unjustified paternalistic scenarios, can depend, and often depend, on the very conception of “freedom” presupposed.<sup>34</sup>

Firstly, it is evident that certain authors dealing with paternalism frequently allude to “negative freedom” (in the berlinian sense). However,

---

<sup>33</sup> It should be recalled that, from the initial definition, coercion is a conceptual component of paternalistic intervention.

<sup>34</sup> This is clearly demonstrated by Richter (2021), who compares the definition of “freedom” as presented by liberalism and republicanism, and shows how the republican conception of “freedom” allows much more room for paternalistic interventions than the liberal one.

this cannot be regarded as the sole conception of “freedom” in order to identify paternalistic interventions. Indeed, such a solution would be encumbered by two reconstruction problems. Firstly, if an intervention interferes with freedom understood as “negative freedom”, then, given the broad denotation of negative freedom, there will be many potential interferences. In other words, it is highly probable that the liberal proponent of negative freedom will be more inclined to identify instances of paternalistic interferences much more readily than someone defending an alternative conception (some of which may be more appropriate). Nevertheless, the most problematic issue is not so much the excessive proliferation of situations that can be qualified as “paternalistic”.<sup>35</sup> Indeed, the second problem pertains to the defence of negative freedom within liberalism itself. The argument is that maintaining negative freedom as the only possible understanding of “freedom” for liberalism would be too restrictive and, moreover, very unrepresentative, condemning both the problematic approach and the solution to misunderstanding.<sup>36</sup> This is because the main problems concerning the *correct* exercise of autonomy within the liberal context affect a freedom understood in the richer sense of “positive freedom”. The ideal situation to consider is that of an agent pursuing her own life plan. Therefore, we are not interested in scenarios of adaptive preferences, where the agent is not genuinely able to choose and is fine to perform actions without interference: consider the traditional example of the slave who is satisfied with her condition and only aspires not to be interfered with in the actions she must perform. The problematic character of such a scenario cannot be grasped with a negative conception of “freedom”, but it can be with a positive one. Instead, we are interested in the scenario in which the agent is master of herself, i.e. she is self-determined, through reflection and her choices, and identifies with what she is (see Celano, 2013, 183-184). Moreover, only by adopting a

---

<sup>35</sup> From a similar perspective, it has been noted that republicanism would tend towards the opposite diagnosis. See Richter (2021); cf. Pettit (1997).

<sup>36</sup> This is indeed the main reason why the analysis by Richter (2021) can be considered partial: it is confined to negative freedom as if this were the only sort of liberal freedom.

conception of “freedom” as “positive freedom” can it be understood that certain paternalistic interventions are justified, insofar as they aim to or succeed in promoting autonomy, contrary to what might initially appear. Conversely, certain paternalistic interventions are unjustified insofar as they do not aim to or succeed in doing so.

What I present is particularly clear and crucial in the case of so-called “libertarian paternalism”. It is true that artefacts such as nudges do not exert coercion, but rather influence their addressees.<sup>37</sup> However, if they are to be considered paternalistic and problematic as a form of paternalism, it is because, in a sense, they presuppose hetero-decision by the agent who implements the nudges.<sup>38</sup> Overlooking this aspect allows the authors to bypass (but only apparently) the problems of the noun “paternalism” and to create the apparently oxymoronic, and in any case paradoxical, expression “libertarian paternalism”.

## 2.6 Fourth Step: Some Indications Regarding the Background Conditions

What I refer to as “background conditions” are the normative conditions favouring intervention, that must outweigh or otherwise counterbalance those against intervention. In this fourth step we are centred in the intervention itself. For instance, intervention is justified if: 1) it is not too invasive; 2) it is not ineffective; 3) it is not too inefficient; and, as a specification of [3], 3’) it is not too costly.

I focus on the invasiveness of the action as a crucial condition. In the previous section we saw that coercion is important. Thus, it is not only a question of whether a measure is coercive, but also of how coercive it is.

---

<sup>37</sup> It can be surmised that in cases such as nudges, an important variable to clarify is what is meant by “influence”, as well as the strength and modalities with which it can be exerted (not to speak of the opacity of the justification or of manipulation). This could be important as a basis for assimilating, at least partially, such “influence” to the phenomenon of coercion.

<sup>38</sup> On top of that, the definition of “paternalism” explicitly formulated by Thaler and Sunstein (2009, 5) is itself partial and unduly advantageous to their position, as it includes only the beneficial element, not the coercive element or the most conceptually basic elements of autonomy and hetero-decision.

Consider the state's abstention from interfering with adult's freedom to determine their own diet: it is generally accepted that there are reasons not to interfere because, although such interference may (if sound)<sup>39</sup> promote the individuals' well-being, self-determination is considered to outweigh the promotion of well-being.<sup>40</sup>

Some further indications must be added to outline the background conditions more precisely. Adding a nuance to the first condition for justified paternalistic interventions, not only do the ends (like the prevalence of well-being over autonomy, or the other way round) matter, but the means are also crucial in evaluating how severe and invasive state coercion (or eventually influence) can be. Consider the general case of a law containing an exhaustive list of healthy products and requiring that such products be mandatory or highly recommended. In a first, more specific scenario, if the state were to ban all products not included in that list, by prohibiting either their importation or their production, this measure would constitute a very significant interference with individual freedom. In a second scenario, if the state were to impose high direct taxes on those products excluded from the above list, this could be considered a moderately significant interference with individual freedom. Finally, in a third scenario, if the measure took the form of a simple recommendation and resulted in each food shop being encouraged to place the products at eye level we could not speak at all of an interference with individual freedom. Even if the coercive element is absent from the third scenario, unlike the first and second, it could still be deemed a case of paternalism in a loose sense insofar the state is conditioning the norm addressees, and there is a hetero-decision by the legislator (regarding what course of action is better), even if this (meta)decision does not exclude the individuals' final decisions at all.

---

<sup>39</sup> It must be kept in mind that a prerequisite is the assumption that the state knows better than the agent what is good for her.

<sup>40</sup> The balance may change if the measure is applied to minors: naturally, because the legal system does not take into account, or takes little account of, the self-determination of minors, but also because even if it were argued that it is the self-determination of adults (i.e., parents or guardians) that is considered, the well-being of minors carries greater weight.

In the different scenarios, state action is incisive to varying degrees (and also, among other things, varies in effectiveness), so some interventions can be justified, while others cannot.

### **3. Anthropology in Law: Competent Agents, Vulnerable Subjects. Critique of Moreso's Proposal**

In this second part of the article, I complement, contrast and deepen some initial distinctions (the definitions of the different justifications of a legal norm) and highlight into the importance of the four conditions for justified paternalistic intervention. However, I will do this mostly indirectly, by demonstrating the need to apply such distinctions through the examination of strategies that do not make them and, as far as I can see, would benefit from doing so.

The reconstruction of an anthropology in law is a strategy very similar to the one I have undertaken but much broader in scope. However, it focuses not on the individual recipient of the law, but on the generic subject of the law, understood either as a capable subject or as a vulnerable subject. An even more striking difference, as we shall see below, is the narrowness of my project compared to the breadth of anthropologies in law.

#### *3.1 Pelagius versus Augustine. An Overly Ambitious Project*

Depending on what kind of anthropology we defend (that is, depending on how we conceive human rational action), we would allocate more or less space to paternalism: the more capable a subject is, the less conceptual and normative space there will be for justified paternalistic interventions; conversely, the more vulnerable a subject is, the more space there will be for such interventions.

I will examine some recent proposals by José Juan Moreso, which are problematic but very illuminating and stimulating, to develop my argument critically.



Moreso (2021, 131; 2023, 158-159) contrasts two general conceptions of individuals, a contrast he traces back to the theological controversy between Pelagius and Augustine of Hippo.

Pelagius holds an optimistic view of rationality and human action, arguing that agents are responsible for their own actions (in theological terms, “they can obtain salvation by themselves”). In terms suitable for our discussion: actions are fully attributable to agents, who do not need assistance to perform them.

Augustine, on the other hand, argues that Pelagianism is an overbearing conception, as human nature is “vulnerable, wounded, torn, ruined” (*De natura et gratia, caput LIII, 62*). Again, in terms more pertinent to us: subjects are fragile, dependent on others (on God, according to Augustine; on other people according to secular views) and therefore must be grateful to other beings.

John Rawls developed a theory that, according to Moreso, can easily be compared to Augustinian anthropology: there is a clear rejection of the concept of “merit” as the basis for allocating goods in a well-organised political community, and a rejection of the thesis that attributes to personal merit what is actually the result of natural luck and the efforts of others (parents, teachers, and so on). This also implies, as in Augustine, the importance of gratitude: individuals are interdependent, and personal improvement comes from those around them (see Moreso, 2023, 159-160; Moreso, 2025, VII).

As I have pointed out, Moreso’s project is too ambitious and, in a sense, misleading. It is too ambitious because it does not seem reasonable to argue, without further qualification, whether human agents are capable even without the help of others (or whether they are altogether rational or powerful) or whether they are incapable without the help of others (or irrational or vulnerable). It will be necessary, among other conditions, to see in which social context we are examining their actions and identifying the dimensions in which subjects need help, because they are fragile.

Sometimes, however, Moreso presents this anthropology not as a general description or a completed project, but as an idea to be considered within a broader, anticipated project, serving as a counterbalance to the allegedly traditional emphasis on the rationality of human agents (see Moreso, 2023, 169). The first interpretation makes the project overly ambitious and misleading; the second makes it more feasible and sound, to the limited extent we recognise that law and legal philosophers tend to overemphasise human rationality and competence, which is far from self-evident. Therefore, we can consider Moreso's proposal (2025, VIII), viz., the vulnerability element in the notion of "autonomy embodied in vulnerable beings" not so much as a way of "thinking by systems" but rather as an "idea to take into account".<sup>41</sup> However, such an interpretation reduces its explicit ambitions.

This is intended to draw on a realistic anthropology, which is also Moreso's theoretical purposes: an anthropology that is or at least aspires to be partly descriptive (by identifying a deficiency and the causal mechanism linking that deficiency to the possibility of suffering harm) and partly normatively meaningful (by assessing of harm and the desirability of avoiding it, as well as identifying other normative components in the very notions of "fragility", "vulnerability", etc.).<sup>42</sup>

Therefore, Moreso's objectives and method must also be adjusted to achieve the theoretical and, even more so, practical outcomes that he himself seeks, namely to palliate, to solve situations of fragility and vulnerability.

"Being vulnerable" means, with minimal lexical precision, "being especially susceptible to suffering harm".<sup>43</sup> Even without openly

<sup>41</sup> I am employing Vaz Ferreira's (1971, 78-94) distinction. Moreso wrote those contributions as a corrective to the overly Pelagian, to his taste, contributions formulated by Atienza and Maniaci.

<sup>42</sup> In the last decades, the term "vulnerability" has been used to counter to the assumption of (supposedly excessive) rationality or capacity attributed to the subject. To this extent, I am following to a very widespread usage, although, as we shall see later, it is not without its drawbacks.

<sup>43</sup> The literature on the jurisprudential uses of "vulnerability" is extensive. Particularly relevant to my criticism are observations that the concept of "vulnerability" is often used in an opaque manner – as noted by Diciotti (2018, 30 ff.) – or even in a circular way – as noted by Maniaci (2025, 164 ff.).

acknowledging it, as in Moreso's case, we are asserting two theses with significant theoretical and practical implications: [1] The vulnerable subject is unable to decide for herself; [2] the subject's lack of decision or, more often, her erroneous decision is likely to cause her harm.<sup>44</sup>

There is a need to trace both the causes and effects of vulnerability, as well as the problem of diagnosis: to identify the scenarios of vulnerability and, above all, the contexts in which this vulnerability is underpinned by some form of "cognitive fragility". Demonstrating why this is important and how it can occur were the purposes addressed in §2.

If my preceding arguments are sound, the theoretical reasons why these moves by Moreso seem questionable are now quite clear. Moreover, regarding vulnerability, it is important to carefully analyse: 1. the reasons why someone is especially susceptible to suffering harm; 2. the sources of such susceptibility; 3. the causal relation between the susceptibility and the harm; 4. the dimensions and depth in which harm is experienced. For instance, regarding [2] it is interesting to note that Moreso's claims, following some of Rawls' path, seem to be centred on "natural vulnerability", omitting particular or group vulnerability (for the categories see Álvarez Medina, 2021, 75). The practical reasons, on the other hand, include the fact that attributing fragility to an individual can lead to the long-term crystallisation of that fragility and, moreover, erroneously so if the individual was not originally fragile. This results from a series of effects: the individual's self-perception, the perception held by a particular social group, psychological mechanisms that make the situation a self-fulfilling prophecy, and social mechanisms that ascribe an identity to the individual and prescribe a conventional response to that identity (among others, cf. Lukianoff and Haidt, 2018, 19 ff.).

What I have just said pertains to the project of building an anthropology in law but Moreso's conclusion that we need to embrace more cases of

---

<sup>44</sup> We can imagine milder forms of these theses: [1'] the agent is unable by herself and without some help (support, supervision), to decide; [2'] through the missed or erroneous decision, the agent does not fully achieve her interests. However, even in such attenuated scenarios, the typical problems associated with the lack of competence persist.

paternalism than classical liberalism commonly accepts is not only a response to his anti-Pelagian anthropology, but also to an argument about the resolution of certain legal cases and the value of human dignity. I will conclude my work by critically examining these two features.

### *3.2 The False Positives of Paternalism*

Moreso's argument is based on a somewhat indeterminate resolution of the legal cases under consideration: both the *Lochner* case and the *Wackenheim* case can be reconstructed not as cases of paternalism, but also as cases of harm to third parties (*Lochner*), perfectionism, and moralism (*Wackenheim*). In the way they are linguistically formulated, I will try to show that the latter is a better reconstruction of these cases.

### *3.3 Lochner Case: Harm to Third Parties as a Concealed Validation for Paternalism*

In *Lochner v. New York*, the US Supreme Court ruled that New York State legislation (the "Bakeshop Act"), which established a maximum of 60 hours per week and 10 hours per day as working hours for bakers, was unconstitutional because, according to the Court, contractual freedom cannot be justifiably limited for the well-being of the contractual parties themselves.

The Court's central reasoning was indeed that such a measure constituted unjustified paternalistic measure. As Moreso quotes:

There is no reasonable ground for interfering with the liberty of person or the right of free contract by determining the hours of labour in the occupation of a baker. There is no contention that bakers as a class are not equal in intelligence and capacity to men in other trades or manual occupations, or that they are not able to assert their rights and care for themselves without the protecting arm of the

State, interfering with their independence of judgement and of action. They are in no sense wards of the State.<sup>45</sup>

Moreso comments:

It is conceivable that a baker might wish to work more than ten hours a day, for example, because he is young and ambitious and wishes to save money to set up on his own. Our political philosophy must show why this measure, which is undoubtedly paternalistic, is justified (Moreso, 2023, 162).

Moreso's treatment of the case is somewhat hasty and, in fact, there are several problems relating to the identification of a paternalistic justification. First, if we reconstruct the *Lochner* ruling with a paternalistic justification, a distinction should be traced between justified and unjustified intentions. As we have analysed, they will be justified if and only if the measure in question: [1] strikes a proper balance between well-being and autonomy; [2] ensures, for instance, a minimum amount of rest for bakers, that is, it compensates for the incorrect balancing or erroneous recognition of values or the way in which they are realised (that is, in general, it remedies some deficit of rationality on the part of the subjects involved); [3] such a resolution is consistent with the basic interests of bakers and with their motivational set. If these conditions are not met, given the coercive nature and benefit to the subject of the intervention, we are faced with an unjustified paternalistic measure.

In any case, aside from this, we may be dealing with a case of harm to third parties. In fact, in the *Lochner* case, it appears particularly clear that we are not so much dealing with a paternalistic justification (i.e., the measure protects each baker from their own decisions) as with a justification based on harm to third parties (i.e., the measure protects bakers who do not want to work at night from those who do).

---

<sup>45</sup> *Lochner v. New York*, U.S. 57 (1905).

Thus, there seems to be a problem of strategic interaction that is resolved either by unfair competition from some bakers or by a deterioration in the quality of life for more bakers. Certain bakers who would not work at night on their own initiative feel compelled to do so if other bakers are allowed to work at night. Therefore, it is not so much a question of protecting the latter from themselves,<sup>46</sup> but rather the former from the latter.

Furthermore, concerning specific mechanism, the idea is that guaranteeing the right to paid holidays, daily rest periods and not having exorbitant working hours as unwaivable would be the ideal mechanism for protecting bakers from employers, to whom they would otherwise be subjected to strong coercive pressures.<sup>47</sup>

Incidentally, the cognitive abilities (more competent, more fragile) that can be assumed of the subjects in each reconstruction of the case are very different.

These two justifications for the case (first, as justified paternalism, and second, as harm to third parties) are absent from both the *Lochner* case itself and the Moreso ruling. However, we must distinguish explicit justifications put forward from implicit justifications presupposed. Indeed, my impression is that the *Lochner* case, as critics of the Supreme Court's decision argue, seems to support paternalism, but especially because harm to third parties, the most notable and solid justification that can be offered for such a case, is only tacitly present and camouflaged as paternalism.

### 3.4 Wackenheim case: *Perfectionism and Moralism Disguised as Paternalism and the Reappearance of Human Dignity*

While in the *Lochner* case it is easy to find cumulative justifications, two of which show respect for the individual's motivational set, the *Wackenheim* case is more insidious, as the justifications for the court's decision appear

---

<sup>46</sup> This misconception is also evident in Moreso (2025, V). This point is made very aptly by Maniaci (2025, XII-XIII; cf. 2025, 20).

<sup>47</sup> I thank an anonymous reviewer of *Athena* for highlighting this point.

rather alternative and do not show respect for the agents' motivational set. Indeed, if the *Lochner* case was mostly a case of harm to third parties disguised as a paternalistic case to implicitly benefit from the greater legitimacy of harm to third parties for liberals, the *Wackenheim* case is largely a case of moralism and perfectionism disguised as a paternalistic case to implicitly benefit from the greater liberal legitimacy of paternalism.

In the *Wackenheim* case, a person with achondroplasia whose main job consists of being thrown against a wall with appropriate safety measures for the entertainment of the audience, is on trial. The French towns where the show takes place prohibit it, but the court of first instance overturns the administrative act of prohibition. On appeal, however, the French Council of State overturns the first judgement and ratifies the towns council's decision. The case is settled according to this *ratio decidendi*:

The Council of State has ruled that respect for human dignity is a component of public order. Consequently, the authority vested with municipal police powers may prohibit an attraction that infringes upon it, even in the absence of specific local circumstances, by exercising its general police powers (Moreso, 2023, 163).

In this occasion, the textual basis for identifying a substantive justification for the judicial decision is even more laconic than in the previous cases.

One possible justification, which neither the ruling nor Moreso's reconstruction overtly considers, is the harm caused to third parties: that Wackenheim is exposed to public ridicule could undermine not some vague abstract dignity of the group of people with achondroplasia, but the self-conception that many people with achondroplasia have, which would be substantially undermined by practices such as those of Wackenheim, leading to the perpetuation of public denigration of people with achondroplasia as a group.

This resolution based on harm to third parties, very similar to the proposal for *Lochner*, seems to have some merit. Only some merit, however, as it is

necessary to carefully weigh the causal attribution of how Wackenheim's conduct affects other people with achondroplasia and ensure that we do not excessively subordinate the freedom of action of individuals (in this case, Wackenheim) to the will of others.<sup>48</sup> This is especially important when it comes to an ascriptive characteristic that places the subject in a group, where belonging is not a freely chosen variable.

Incidentally, as in the *Lochner* case, this also presupposes a lesser capacity for self-determination on the part of the normative subject (Wackenheim and anyone who may find themselves in a similar dispute).

The main point is that the Council of State's decision and Moreso's reconstruction offer justification elsewhere, in a much more objectivist direction focused on the notion of "dignity".

Of course, "human dignity" is a very important concept in contemporary legal systems and serves significant functions (among others, Azzoni, 2012). However, there are often several problems with its use.<sup>49</sup> One of these concerns is the content of such dignity and whether it always prevails over other competing considerations. For example, Moreso, following Atienza, provides a Kantian version of dignity that presupposes a naive compatibilism between dignity and autonomy.

In fact, the issue is that if, as Kant maintains, individual dignity must be respected because it consists in respecting individuals as agents capable of acting according to what they believe to be right (cf. Kant, 2005), are we not assuming that agents do not make mistakes, i.e., that they always choose one of the many options of what can be considered right? The only limit recognised by Kant in this regard is the infringement of the freedom of third parties. As I have tried to show, at least indirectly, this anti-conflictual view of well-being and autonomy is neither realistic (since genuine conflicts arise

---

<sup>48</sup> Cf. Maniaci (2025, 171) for additional conditions that a third-party aggression based on "human dignity" must meet.

<sup>49</sup> For the insidious uses of "human dignity" in paternalistic interventions cf. Maniaci (2025, 167-173).



and, in some cases, the agent is right and in others wrong, without necessarily causing harm to third parties) nor, therefore, theoretically illuminating.

A final problem with human dignity, and the most pressing one for liberalism, is that its use makes it easy to slip into perfectionism and moralism. In fact, in considering a certain conduct “contrary to human dignity”, we often impose on the subject a standard of personal perfection or a standard of moral correctness to which the subject does not adhere. In the *Wackenheim* case, it is genuinely debatable whether Wackenheim himself, whose main job was to be thrown, given the limited employment options available to him, could be persuaded of the fact that his dignity, the ideal to which, if enlightened, he would supposedly subscribe, requires him not to perform that job. As I have already shown, this is a case in which we impose a course of action and certain ideals on the individual against their preferences and, unlike paternalism, we are not even interested in considering their desires and beliefs.

### *3.5 The Moral of the Story About Alternative Justifications*

To sum up, Moreso views *Lochner*’s judicial decision as a case of justified paternalism, whereas I see it as a case of justified application of the harm principle.

Moreso regards *Wackenheim*’s judicial decision as a case of justified paternalism, while I see it as a case of unjustified perfectionism and moralism.

We must think twice about the presuppositions of rationality endorsed by legal norms and judicial decisions. Moreso’s reconstruction of these cases presupposes a lower rational capacity – indeed a generic vulnerability – on the part of the normative parties judged; instead, we can assume higher rational capacities for them. We must await the accurate resolution of the “problem of diagnosis” before providing a solution, and we should be more confident in the rational abilities of the norm addressees. If this is so, we must not be overly deferential to paternalism.

If the detailed determination of the problem of diagnosis is so important for resolving these kinds of cases, *a fortiori*, we should consider attempts to construct an anthropology in law to be an overly ambitious and misdirected project.

## References

- Aleman M. (2006). *El paternalismo jurídico* (Iustel).
- Álvarez Medina S. (2018). *La autonomía de las personas* (CEPC).
- Álvarez Medina S. (2021). *La protección de la vida privada y familiar* (Marcial Pons).
- Azzoni G. (2012). Dignità umana e diritto privato, in *Ragion pratica*, n. 38.
- Bayón J. C. (1989). Causalidad, consecuencialismo y deontologismo, in *Doxa*, n. 6.
- Buchanan A. (1978). Medical Paternalism, in *Philosophy & Public Affairs*, vol. 7, n. 4.
- Celano B. (2013). *I diritti nello stato costituzionale* (il Mulino).
- Chiassoni P. (2019). La balanza inesistente, in *Analisi e diritto*, 2019.
- Colomer J. L. (2004). Autonomía y derechos, in J. Betegón, *et al.* (eds.), *Constitución y derechos humanos* (CEPC).
- Diciotti E. (2021). Alcune osservazioni sull'antipaternalismo moderato di Maniaci, in *Diritto e questioni pubbliche*, no. 2.
- Diciotti E. (2018). La vulnerabilità nelle sentenze della Corte europea dei diritti dell'uomo, in *Ars interpretandi*, no. 2.
- Diciotti E. (2005). Preferenze, autonomia e paternalismo, in *Ragion pratica*, 2005/1.
- Diciotti E. (1988). La giustificazione paternalistica di norme, in *Studi senesi*, 1988/1.
- Diciotti E. (1986). Paternalismo, in *Materiali per una storia della cultura giuridica*, n. 16.
- Dworkin G. (1972). Paternalism, in *The Monist*, vol. 56, n. 1.

- Elster J. (1983). *Sour Grapes: Studies in the Subversion of Rationality* (CUP).
- Feinberg J. (1986). *Harm to Self* (OUP).
- Feinberg J. (1984). *Harm to Others* (OUP).
- Fernández Núñez M. (2023). *Derechos e intereses* (CEPC).
- Garnett M. (2017). Agency and Inner Freedom, in *Noûs*, vol. 51, n. 1.
- Hart H. L. A. (1963). *Law, Liberty, and Morality* (OUP).
- Husak D. (1992). *Drugs and Rights* (CUP).
- Kalis A., et al (2008). Weakness of will, akrasia and the neuropsychiatry of decision making, in *Cognitive, Affective, & Behavioral Neuroscience*, vol. 8, n. 4.
- Kant I. (2005) [1785]. *Fundamentación de la metafísica de las costumbres* (Tecnos).
- Lukianoff G. and Haidt J. (2018). *The Coddling of the American Mind* (Penguin).
- Maniaci G. (2025). *The Unbearable Lightness of Legal Antipaternalism* (Springer).
- Mele A. (2016). *El autoengaño desenmascarado* (Cátedra).
- Mill J. S. (1864). *On liberty* (Longman).
- Monti E. (2009). Neutralidad, autonomía y paternalismo, in P. Capdevielle, et al (eds.), *Libres e iguales: estudios sobre autonomía, géneros y religión* (UNAM-III-UBA).
- Moreso J. J. (2025). Foreword: Pelagius or Augustine?, in G. Maniaci, *The Unbearable Lightness of Legal Antipaternalism* (Springer).
- Moreso J. J. (2023). Sobre el semipelagianismo de Atienza, in E. Díaz and F. Laporta (eds.), *Leer a Manuel Atienza* (CEPC).
- Moreso J. J. (2021a). *Lo normativo: variedades y variaciones* (CEPC).
- Moreso J. J. (2021b). Acerca del pelagianismo de Maniaci, in *Diritto e questioni pubbliche*, no. 2.
- Nino C. S. (1989). *Ética y derechos humanos* (Ariel).
- Pedrini P. (2013). *L'autoinganno* (Laterza).
- Pettit P. (1997). *Republicanism* (OUP).

- Purcell J. R., *et al* (2022). A review of risky decision-making in psychosis-spectrum disorders, in *Clinical Psychology Review*, n. 91.
- Rawls J. (1999). *A theory of justice* (HUP).
- Raz J. (1986). *The morality of freedom* (OUP).
- Richter A. (2021). ¿Cuán paternalista es el republicanismo?, in *Analisi e diritto*.
- Russell B. (1999). *Russell on Ethics* (Routledge).
- Schimmack U. *et al* (2017). Reconstruction of a Train Wreck, in *Replicability-Index*, <https://replicationindex.com/2017/02/02/reconstruction-of-a-train-wreck-how-priming-research-went-of-the-rails/>.
- Thaler R. H. and Sunstein C.R. (2009). *Nudge* (Penguin).
- Vaz Ferreira C. (1979). *Lógica viva: moral para intelectuales* (Ayacucho).