ARTICLES

Universal Health Care and Enforced Beneficence
Christoph Hanisch

Compensation as Moral Repair and as Moral Justification for Risks
Madeleine Hayenhjelm

Effort as Responsibility
David Jenkins

SPECIAL SECTIONS:

LYING AND HYPOCRISY IN POLITICS AND MORALITY, WITH RUTH GRANT

RADICALISM AND COMPROMISE
ETHICS, POLITICS & SOCIETY
A JOURNAL IN MORAL AND POLITICAL PHILOSOPHY

Number 2-2019

Universal Health Care and Enforced Beneficence

Christoph Hanisch

Compensation as Moral Repair and as Moral Justification for Risks

Madeleine Hayenhjelm

Effort as Responsibility

David Jenkins

SPECIAL SECTIONS:

LYING AND HYPOCHRYSY
IN POLITICS AND MORALITY,
WITH RUTH GRANT

RADICALISM AND COMPROMISE
Title

Editor-in-chief
João Cardoso Rosas.

Editors
Alexandra Abranches, David Álvarez, João Ribeiro Mendes, Pedro Martins, Roberto Merrill.

Assistant Editor
António Baptista

Guest editors
Pedro Martins and Giuseppe Ballaci, for «Lying and Hypocrisy in Politics and Morality, with Ruth Grant»; Alexandra Abranches for «Radicalism and Compromise»

E-mail
epsjournal@ilch.uminho.pt

Editorial board
Acílio Estanqueiro Rocha, Alexandra Abranches, Ana Rita Ferreira, António Baião, António Luís Silva Baptista, Bernhard Josef Sylla, Bruno Serra, Catia Faria, Daniele Santoro, David Álvarez, Eze Paez, Giuseppe Ballacci, Hugo Rajão, João Ribeiro Mendes, José Augusto Colen, João Cardoso Rosas, Marco Marian, Pedro Miguel Martins, Roberto Merrill, Sara Bizarro.

Scientific Committee
Adam Swift, Alan Thomas, Ángel Rivero, Jeffrey McMahan, Jonathan Wolff, José de Sousa e Brito, Maria Xosé Agra, Peter Singer, Richard Bellamy, Samuel Scheffler, Seyla Benhabib, Steven Lukes, William Edmundson.

Edition
April 2019. Centre for Ethics, Politics and Society University of Minho, Braga, Portugal.

ISSN DOI
2184-2582 https://doi.org/10.21814/eps.2.1

Website
http://www.epsjournal.ilch.uminho.pt
TABLE OF CONTENTS

*Universal Health Care and Enforced Beneficence*
Christoph Hanisch .......................................................... 7

*Compensation as Moral Repair and as Moral Justification for Risks*
Madeleine Hayenhjelm ....................................................... 33

*Effort as Responsibility*
David Jenkins ................................................................. 65

LYING AND HYPOCRISY IN POLITICS AND MORALITY,
WITH RUTH GRANT

*Introduction*
Giuseppe Ballacci, Pedro Miguel Martins .............................. 95

*Lying and hypocrisy in morality and politics*
Ruth W. Grant .............................................................. 101

*Political pragmatism and ethics in Kautyla’s Arthashastra: contradiction or complementarity*
Shreya Bhattacharya, Kankana Saikia ................................. 111

*A contingência da política e a necessidade da hipocrisia em o príncipe de Maquiavel*
Albano Pina ................................................................. 123

*Can hypocrisy be a virtue? Hume on the morality of princes*
Alexandra Abranches ...................................................... 139

*Rescuing politics from lying and hypocrisy: utility and truth in Jeremy Bentham’s thought*
Benjamin Bourcier .......................................................... 155

*The Role of Lying in Politics*
Kathrin Bouvot 177

**Argumentative hypocrisy and constituent debates: the italian case**

Giovanni Damele 205

**Why the deliberative ideal justifies publicity – even if publicity may undermine deliberation**

John Pitseys 221

**RADICALISM AND COMPROMISE**

**Introduction**

Alexandra Abranches 255

**A Politics without Compromise: The Young Hegelians and Politics**

Vivien García 259

**Enacting Levinas’s infinite responsibility as an ethico-political compromise**

Julio. A. Andrade 273

**A reactionary fascination: Emil Cioran and Joseph de Maistre**

Paolo Vanini 287

**What Does Being An “Aristotelian” Really Mean?**

Sinan Kadir Çelik 303

**Compromise and/or Integrity? Some Normative Remarks**

Beatrice Magni 335

**The Virtue of Compromise**

Yuval Eylon 355
Abstract. I examine Allen Buchanan’s arguments for enforced beneficence and express a number of worries concerning his attempt to justify coercive distributive policies that guarantee (basic) health care services for all citizens. The central objection questions whether, given Buchanan’s own stipulation of universally-instantiated attitudes of moral beneficence amongst all society members, his arguments from, first, the coordination problem and, second, the assurance problem successfully establish a justification of enforced contribution. I defend alternative, non-coercive, responses to the aforementioned problems and show that a particular kind of institution (an “information service”) provides all citizens with the sufficient and reliable epistemic resources so that they can effectively help the sick and needy. I notice that Buchanan’s difficulties with justifying coercion can be regarded as providing indirect support for the view that developing a justice-based conception of moral health care rights remains, pace Buchanan, an important task to be completed.

Keywords: Coercion, Redistribution, Libertarianism, Health Care, Coordination Problem, Assurance Problem.

Sumário. Neste artigo examino os argumentos de Allen Buchanan a favor da beneficência obrigatória e formulo uma série de preocupações relativas à sua tentativa de justificar políticas distributivas coercivas que garantam serviços de saúde (básicos) para todos os cidadãos. Dada a estipulação de Buchanan sobre as atitudes universalmente instanciadas de beneficência moral entre todos os membros da sociedade, a nossa objeção central questiona os seus argumentos sobre o problema de coordenação assim como o problema de garantia estabelecem uma justificação da contribuição obrigatória. Defendo respostas alternativas e não coercitivas aos problemas acima mencionados e mostr o que um tipo particular de instituição (um “serviço de informação”) fornece a todos os cidadãos recursos epistêmicos suficientes e confiáveis para que possam efetivamente ajudar os doentes e os necessitados. Noto por fim que as dificuldades de Buchanan em justificar a coerção podem ser consideradas como um apoio indireto à visão segundo a qual o desenvolvimento de uma concepção baseada na justiça dos
Allen Buchanan has presented an influential approach to the two-part question of whether citizens are entitled to basic health care services and whether collective agents (e.g., government agencies) are justified in realizing this aim by employing coercive instruments and policies.

Buchanan addresses this issue from a moderately libertarian perspective, i.e., an individualistic view of society that assigns a fundamental role to negative rights and liberties, such as the rights to non-interference and private property. Since moderate libertarianism regards these negative rights as non-absolute, Buchanan is able to answer the aforementioned question affirmatively. He provides the resources for allowing a democratic society to implement an entitlement (a legal right) to basic health care by means of collective and publicly-coordinated endeavors. Buchanan is well aware that the crucial issue with regard to universal basic health care is the attempt to confront the categorical libertarian objection to any coerced financial contributions to such a project. Having conceptualized this aspect of the debate in such a clear manner makes Buchanan’s contribution timely and worth having another look.

Buchanan defends a “pluralistic account” of justifying this entitlement and its public implementation. His account assigns a central role to the idea of “enforced beneficence” which is, according to Buchanan, necessary to discharge the moral obligation to help those who face health-related needs in the most effective manner, without thereby appealing to the notion of a positive moral (as distinct from legal) right to health care. Buchanan’s contribution does not primarily lie in his rejection of the notion of a universal moral right to health care, though. Rather, Buchanan’s important thesis is that even if a moral right to health care cannot be defended, this failure does not render unfeasible the

---

aforementioned defense of coercion that gets employed in levying and coordinating individual contributions, based on beneficence.

I examine Buchanan’s arguments for enforced beneficence and express a number of worries concerning his attempt to justify coercive policies. The central objection questions whether, given his own stipulation of universally-instantiated attitudes of moral beneficence amongst all society members, Buchanan’s arguments from, first, the coordination problem and from, second, the assurance problem successfully establish a justification of coercive contribution. I defend alternative, non-coercive, responses to the aforementioned problems and show that a collective agent/institution might well provide all citizens with sufficient information concerning the question of how to effectively help the sick and needy. My criticism does not deny that Buchanan has identified a set of important collective action problems regarding the universal provision of health care. However, given Buchanan’s own premises, his argument falls short of vindicating his ambitious conclusion concerning enforcement as the remedy for these problems. Coercive and non-voluntary transfers of resources are simply not necessary in a society in which moral attitudes of beneficence are shared by all and when this fact is acknowledged by all members of the stipulated society. I notice that Buchanan’s failure to justify coercion can be regarded as providing some indirect support in favor of the claim that developing a workable conception of moral health care rights remains, pace Buchanan, an important task to be completed (but not taken up in this paper).

1. Universal Basic Health Care Without Moral Rights

The diversity of ethical issues involved in debates about health care policies has grown in both academic and non-academic discourse. Technological innovations in medical science, and the related rising costs of medical practice, have led to an intensified scrutiny of the notion of health care rights and of the question of what “equity” with regard to access to health care services amounts to in scientifically-advanced capitalist societies. The problem that is central to this paper concerns the efficacy of collective (and to-be-coordinated) acts of providing health care for all citizens, and, most fundamentally, how enforced contributing
to such policies can be justified on grounds of this desired efficacy even if positive moral rights to health care are called into question.

The idea that society (and its citizens) have an enforceable obligation to contribute to public health care endeavors, even if the sick and needy do not have a corresponding moral right to health care, lies at the heart of Buchanan’s work; work that has been influential partly because it rightly reminds us not to take moral rights as an unquestionable given. Buchanan does not only call into question liberal and social democratic conceptions of universal welfare rights; Buchanan also criticizes conservative and libertarian strategies of categorically rejecting any legally-guaranteed welfare provisions in the name of supposedly absolute property rights and liberties. The latter dimension of his writings is the central topic of the following reflections.

In his “The Right to a Decent Minimum of Health Care” Buchanan (1984, pp. 59-66) criticizes three prominent philosophical (liberal) proposals and shows that they cannot establish a positive right to health care (utilitarianism, Rawls’ “Justice as Fairness,” and Norman Daniels’ (1985) attempt to derive a right to health care from a Rawlsian principle of fair equality of opportunity). Its unsuccessfulness notwithstanding, Buchanan claims that we should nevertheless be careful when we consider the implications of this three-fold failure. Just because all these attempts to justify a positive right to health care fail, this failure does not lead, by default, to the libertarian triumph consisting in the successful refutation of any coercive arrangements designed to secure health care. In his discussion of rights’ enforceability Buchanan (1984, pp. 56-7) notes:

Indeed, the surprising absence of attempts to justify a coercively backed decent minimum policy by arguments that do not aim at establishing a universal right suggests the following hypothesis: advocates of a coercively backed decent minimum have operated on the assumption that such a policy must be based on a universal right to a decent minimum. The chief aim of this article is to show that this assumption is false.²

² One issue that I can only address in passing is the important question of why Buchanan’s challenges of coordination and mutual assurance (extensively discussed below) do not support a more generous universal healthcare system, but merely the mentioned “decent minimum of health care.” (I am indebted to a referee for this journal for raising this question.) Recently, Buchanan has directly addressed this issue by stating that “obligations of beneficence are traditionally understood to be limited by the proviso that rendering
Even if all rights-based theories fail, Buchanan continues, there is an alternative (pluralistic) justification of coerced contribution to health care provision available. Buchanan begins his positive case by noticing that politicians and philosophers who are attracted to rights language in health care debates attribute significant importance to the object of that right, *viz.* a secured minimum of medical care for all. Why should they insist on *this* project being realized if and only if we are able to tag the label of “rights” on the respective policy? Buchanan (1984, p. 66; my emphasis) suggests that they should not: “My suggestion is that the combined weight of arguments [*none of which is based on antecedent moral rights*] is sufficient to *do the work* of an alleged universal right to a decent minimum of health care.”

Buchanan is quick in assuming that his beneficence-based account will be able “to do all the work” that a justice- and rights-based account can do, without running into the aforementioned problems that come with moral rights. Even if we restrict ourselves to a consequentialist perspective, that focuses on the outcomes of the two competing justificatory approaches, Buchanan’s claim seems overly optimistic for a variety of reasons (on top of those that I spell out in the next sections). Especially with regard to the issue of the subjective experience of social and economic *security*, a universal and publicly guaranteed positive entitlement in terms of fundamental moral rights appears to contribute in ways that cannot be fully accounted for when society leaves everything to charitable impulses. As we will clarify in a moment, Buchanan (1984, p. 57) defends his charity-based approach by claiming that “[t]o the morally virtuous person the imperatives of charity may be as urgent as those of justice.”

First of all, this claim may be unconvincing to those who do not already believe that Buchanan’s envisioned moral community can be realized and consequently think that meeting the basic health care needs of all must be backed by an irreducible appeal to rights and social *justice*. (Kantians, like Ripstein (2009) for example, have good arguments for such claims.) In addition, even if aid to the needy is not to be unduly burdensome to the benefactor. Consequently, the enforced beneficence approach avoids objections to which more demanding egalitarian concepts of the right to health care are vulnerable” (Buchanan 2009, p. 74). I do not find this quick clarification fully satisfactory but a detailed analysis of this aspect of Buchanan’s view has to be postponed for another occasion.
all citizens were morally virtuous persons (but, in principle, retain the freedom to change this attitude), and gave enough to meet the health care needs of all, it may nevertheless be a constitutive feature of a guaranteed basic minimum that the beneficiaries know that their health care needs are taken care of as a matter of justice-based entitlements and rights. Of course, this first set of critical remarks is not a conclusive argument in support of such a moral right’s existence; Buchanan will probably (and rightly) highlight this point. However, it calls into question Buchanan’s optimistic initial assumption that beneficence can do all the work that rights-based approaches do. I put aside these worries. Having said that, especially the second worry from the good of rights-based public guarantees leads us to the core of Buchanan’s ambitious defense of enforced beneficence. To this central innovation we now turn.

2. Buchanan’s Pluralistic Justification of Coerced Contributing

Buchanan’s strategy for justifying an enforceable principle, guaranteeing a decent minimum of health care for everyone, is pluralistic. It establishes citizens’ access to basic medical services in the framework of four independent considerations. The combined weight of these considerations is supposed to provide an argument justifying a centralized agent using coercion in order to meet the moral obligation of beneficence (as distinct from an obligation of justice) to help those in need of medical assistance and services. A crucial element of Buchanan’s strategy is to present the obligation of beneficence in question as a collective one. Buchanan (1984, p. 70) justifies this shift from individual to collective beneficence by means of pointing to the significant financial and organizational efforts that are necessary to realize the goal of providing minimal medical services for all members of society.¹ A decent basic minimum of health care is such a collective good.² These goods come with a particular set of problems (other examples of such collective goods are national defense, environmental

---

² I am indebted to a referee for highlighting that the good in question is not properly referred to as a traditional “public good.” Hence, I replaced the language of public goods in many contexts. The good in question is health care that private and distinct individuals (the patients) enjoy. However, the crucial point is that we are considering the option of realizing and guaranteeing these goods in a public manner. “Publicly-realized-personal-goods” is probably the best label for what is at stake.
Universal Health Care and Enforced Beneficence

... protection, etc.) that must be confronted by societies in order to successfully generate these goods.

Before dedicating the remainder of this paper to these problems let me briefly present the other three considerations that Buchanan’s (1984, pp. 66-68) pluralistic strategy employs. The first argument focuses on “special rights” and is concerned with the rectification of past and present injustices (e.g. health problems related to discriminatory policies), the requirements of compensation (e.g. health problems related to a third party’s negligence), and entitlements to health care based on extraordinary sacrifices that citizens provide for their society (e.g. impairments due to compulsory military service).

The second consideration argues that a lack of certain kinds of collectively provided (and enforced) basic health care can be regarded as the violation of a specific negative right. This argument provides a justification for health care measures such as public sanitation and immunization programs and can be summarized under the heading of “harm prevention”. It is, for example, in everyone’s interest that certain infectious diseases are controlled by reliable public agents and services; collectively implementing (and financing) immunization programs (also, and especially5, for those who could not otherwise afford the vaccine) is in everyone’s interest.6 Thirdly, there is a number of prudential arguments in support of a guaranteed decent minimum to health care. These arguments emphasize a healthy population’s collective benefits such as a more productive labor force and fitter soldiers.

Buchanan is confident that these three arguments present strong support for the claim that every citizen should have access to a decent minimum of health care. He also reminds us that the three arguments do without any appeal to universal (as opposed to special) positive health care rights. In addition, and this

5 I say “especially” because historically it was the worst off members of societies that had been exposed to diseases that are now controlled by comprehensive preventive measures. Especially in today’s circumstances (globalization and urbanization) the better off can hardly argue at this point that they are not required to contribute to vaccination programs because they are free to avoid close contact with the poor.

6 It is worth noting that this second argument by Buchanan challenges radical libertarianism in two ways: not only does it ask citizens to contribute, at least some amount of resources, to the collective sanitation and/or immunization program. In addition, Buchanan’s argument seems to imply that a collective agent is justified in forcing citizens to undergo this immunization, if this is necessary to prevent harm not merely for the person in question but for those around her. I cannot pursue these issues here. Cf. Francis (2005).
is the major difference to Daniels’ and Rawlsian approaches, there is no need for a comprehensive theory of justice in order to support this entitlement to basic medical care.

The major component of Buchanan’s pluralistic strategy is still underdeveloped at this stage of the argument though. After all, Buchanan claims that a health care regime based on beneficence and charity (as opposed to justice and positive moral rights) can do all the work that a coercive (welfare) state could do, viz. to guarantee a decent minimum of health care for each individual member of society. So far Buchanan has established the justification for such a legal entitlement either for some subgroups of the citizenry only (e.g. former military staff), or for the populace at large, but merely with regard to an extremely minimal subset of essential medical services (e.g. vaccinations against some contractible diseases). I now turn to Buchanan’s argument for enforced beneficence that is supposed to solve this problem of providing the resources necessary for such a guarantee without thereby being committed to a universal moral right to health care. This is the fourth and final argument of Buchanan’s pluralistic strategy. Buchanan spends by far the most amount of time and effort on defending the argument from enforced beneficence. It is crucial for the success of his pluralistic strategy.

3. Buchanan’s Argument for Enforced Beneficence

Buchanan claims that reasonable secular and religious moral outlooks accept the existence of a moral duty of beneficence to help fellow humans in dire need. According to Buchanan, even (most) libertarians are committed to this duty. The latter can do so because this duty does not seem to imply any positive moral rights on part of the beneficiaries. With regard to the issue of providing a decent minimum of health care, however, discharging this obligation to help those in dire need consists, primarily, in contributing to the collective endeavor of providing basic medical services for those who cannot otherwise afford them. As mentioned above, this assumption is crucial. The most effective way to discharge the obligation in question, according to Buchanan, is assumed to be a collective regime, as opposed to individual, small-scale initiatives.
With regard to this collective effort, Buchanan asks us to envision two scenarios. First, a situation in which all agents are actually morally motivated to discharge their duty of beneficence to help those in need effectively. Secondly, we imagine a situation in which an individual benefactor cannot be sure that all others are equally motivated by this duty of beneficence as she is. Buchanan’s conclusion is that in both scenarios, beneficent and rational individuals have a decisive incentive not to contribute to the collective effort of providing a decent minimum to health care. However, acting on this rational incentive results in the most efficient policy not getting realized. This is so despite the fact that the agent is perfectly beneficent and consequently wants to help as effectively as possible.

3.1. The Coordination Problem

The first scenario envisions a society of individuals, all of whom are motivated to act in accordance with their duty of beneficence to help those who cannot purchase a decent level of health care. It is exactly because of this universally-present genuinely beneficent motive that the agents in question want to discharge this duty effectively, i.e., they want their individual contributions to improve the situation of the poor to the greatest possible extent. Buchanan assumes that this motivation expresses itself in the fact that each beneficent agent maintains a willingness to provide (a portion of) the means necessary for the successful implementation of a decent minimum of health care policy. This is so because the agent is aware of the significant financial resources that are needed in order to achieve this aim. Buchanan presumes, for the sake of argument, that the agent in question is, in principle, willing to direct her individual contribution to this collective project because it promises to be the most effective way to discharge her endorsed duty of beneficence to help those in dire need.

There is one inescapable problem though, according to Buchanan. Since the benefactor is obligated (and willing) to help effectively, and since her individual contribution is going to be marginal (in comparison to the large number of individual contributions needed to realize the collective aim) she will conclude (for reasons detailed in the next paragraph) that the most rational thing

---

7 If we assume that our duty of beneficence must be discharged in an impartial manner and in accordance with some minimally egalitarian intuitions then a successful decent minimum policy gets all beneficiaries above the threshold of access to basic medical care. Once this goal is realized then the specific duty of beneficence, that Buchanan is concerned with, “disappears” so to speak.
to do is not to contribute to the collective endeavor. To the individual contributor, this conclusion appears to be the most rational, exactly because her contributing will very probably result in a contribution that less than maximally helps those in need. The rational and beneficent agent will, therefore, rather direct her individual contribution to small scale (but, overall, less effective) projects that aim to alleviate the health problems of those who cannot do so themselves, for example, on the level of local health care initiatives. Consequently, and assuming that all other beneficent agents are deliberating in a similar fashion, the most effective way to discharge the obligation in question will not be realized. Why does this paradox result? After all, are we not assuming that all agents are motivated by proper moral motives and genuinely want to help the poor as effectively as possible (and all society members know this about one another)? According to Buchanan, this first problem amounts to a variety of the “coordination problem”.

The dilemma that each potential contributor faces is that either her contribution unnecessarily adds to the good of universal health care because enough others have already contributed or she gives her resources when not enough others contribute. Buchanan (1984, p. 70) concludes: “In either case, my contribution will be wasted. In other words, granted the small scale of the investment required and the virtually negligible size of my own contribution, I can disregard the minute possibility that my contribution might make the difference between success and failure.”

In both cases the beneficent agent’s contribution is wasted and would have been of more effective (and more beneficial) use if it had been spent on individual and small scale projects, despite the fact that these latter projects turn out less effective overall in comparison to the collectively-provided decent minimum policy. Again, since the most rational thing to do appears to be not to contribute to the collective policy it will not be realized despite its acknowledged superior efficacy (the very fact that would make beneficent agents contribute to it in the first place).

---

8 At this point it is important to keep in mind that the other individuals are not refusing to contribute because self-interested motives overpower their beneficence-based motivation. This other problem is more relevant, it seems at first sight, for the second scenario that Buchanan discusses and that illustrates the assurance problem – as distinct from the coordination problem that is currently at stake.
The next step is the crucial one: Buchanan concludes that there is only one way to resolve this problem and to ensure that all citizens contribute to the decent minimum regime in a well-coordinated manner. Buchanan (1984, p. 70; my emphasis) says: “But if everyone, or even many people, reason in this way, then what we each recognize as the most effective form of beneficence will not come about. Enforcement of a principle requiring contributions to ensuring a decent minimum is needed.”

It is at this point that one wonders if the justification of enforced beneficence is in fact successfully established by Buchanan’s previous argument. Keep in mind that we are working under the assumption that all individuals are motivated by the stringency and force of an accepted moral obligation. Conflicts between individual self-interest on the one hand and duties of beneficence on the other are not the problem afflicting the scenario discussed. Given this reliable and society-wide presence of beneficent motives it appears ad hoc to claim that coercive mechanisms are needed and justified in order to overcome the coordination problem. My main objection to Buchanan’s first argument (and as it will turn out, also to his second) for enforced beneficence is that it does not establish a clear link between the coordination problem and any sufficient justification of collectively-imposed coercion. Given the current discussion of the coordination problem, it is difficult to see why other (non-coercive) mechanisms are incapable of overcoming this particular problem. The problem in question is “information based”, as opposed to being a challenge that requires coercion for its solution. Let me clarify and illustrate this objection to Buchanan’s first attempt to vindicate the enforcement of contributions to collective health care endeavors. I will then reply to objections and potential defenses of Buchanan.

Consider this alternative “mechanism.” The state, the government, or some private institution may provide a service that solves the coordination problem without using coercion by determining the amount of each individual contribution to the collective project of guaranteeing the decent minimum of health care (again, a good that all beneficent individuals are presumed to want to realize together). In returning to Buchanan’s discussion of the rational incentive that each of these agents has for not contributing, we can make my proposal clearer. The beneficent individual that we are asked to imagine is in a state of
epistemic, rather than a motivational and normative “uncertainty.” The task of my suggested institution, let us call it the “information service,” would be to remove this epistemic uncertainty and to determine each individual’s contribution that is necessary to achieve the publicly-guaranteed good of universal access to health care. Notice that the ultimate step of actually transferring the contribution, as determined by the information service, is then nothing anybody needs to be coerced to in a scenario of universally-maintained attitudes of beneficence. After all, thanks to the imagined highly-reliable information service, all individuals know that the contribution they give won’t be wasted. They voluntarily give their contribution that is needed to realize the most effective policy for helping the sick, i.e., the collective decent minimum policy, as envisioned by Buchanan.

Two objections to my proposal emerge immediately. Firstly, it appears unrealistic that the envisioned institution can determine the very exact amount of individual contributions in a way reliable enough to overcome the coordination problem. After all, in order to determine the size of the individual contributions, one has to know the exact overall budget necessary to realize the decent minimum policy. Considering the unpredictability of advancements in medicine and the similarly unpredictable rise or decline of the number of worse off citizens who must be covered by the decent minimum policy (and the poor’s diverse medical conditions), it seems rather unlikely that the information service can confidently guarantee each individual agent that her contributions are precisely as large as this is needed in order to avoid that her contributions are wasted.

One first thing to notice in response is that this same practical problem seems to apply to Buchanan’s policy of enforced beneficence. A real-world coercive institution might either fail to enforce enough contributing (i.e., not enough for successfully implementing the decent minimum policy). Or it collects too much (i.e., some of the enforced contributions are getting wasted at the end of the day) – both options raising the specter of the collapse of the institutions’ efficaciousness and, hence, legitimacy in the eyes of the coerced benefactors. If the enforcing institution, on the other hand, really knows what the precise contribution size is and enforces the extraction of the correct amount then it becomes again redundant regarding the enforcement part of the story:
Benefactors who are acting from a sincere duty of beneficence and have access to the accurate (and trustworthy!) information that is used by the centralized agency to determine the amount of needed contributions, do not have to be coerced to contribute to begin with.

Secondly, it is an empirical question how my information based approach can avoid these problems of recommending slightly too little or slightly too much voluntary contributions. One expects this empirical issue to be settled by continuous political and practical processes.9

In addition, one might propose that the beneficent and non-coercive contributions, determined by the information service, are deliberately set slightly above the expected budget that is needed to ensure the decent minimum for all. This seems to be justified because of the specific features of the collective good in question, i.e., the unpredictable nature of health-care-related public policy goals (mentioned above). The resulting surplus will then not count as “wasted.” It can be used in the following year(s) and/or for other basic minimum projects that have an indirect and long term impact on the society’s overall health situation (such as dietary initiatives in public schools, etc.).

Moreover, one can argue that it is part of the idea of a guaranteed decent minimum, briefly mentioned in section one, that there are at least some additional reserves available in case the information service was too conservative in its projections – something that will always remain a possibility, regardless of how well the information institution is set up. According to this rejoinder, contributing to an already sufficiently-financed decent minimum program turns out not to be wasteful after all; on the contrary, it contributes to this minimum being guaranteed for all potential patients in the face of the inherently unpredictable variables characterizing the complex project in question.

---

9 Another issue is that the “right” amount of individual contributions is dependent on a particular society’s conception of the decent health care minimum. Determining what services count as basic enough and what medical resources are to be dedicated to the implementation of these services (e.g., decent technology vs. the best technology available regarding cancer treatment) is not a value-neutral and apolitical project. I assume for the purposes of the discussion in the text that societies have settled that issue, that is, the decent minimum policy is reasonably codified and limited and its administrators have a precise idea of what this policy is going to cost. Many have stressed the point that continuous democratic deliberation is called for in order to settle these complex questions in real-world circumstances in a legitimate way. Cf.: Gutmann (1982, pp. 556-8) and Bole (1991, pp. 10-7).
A different way to formulate an objection to the suggested non-coercive solution to the coordination problem highlights that Buchanan repeatedly emphasizes the, seemingly unavoidable, negligible size of the individual contributions.\textsuperscript{10} Why should the information service make a difference regarding this particular dimension, given that any morally-motivated agent, faced with the negligible impact that her contribution potentially has, will again judge that her contribution does more good if transferred to small scale and local initiatives and policies? We seem to run into the initial problem, the presence of the powerful information service notwithstanding.

As mentioned above, it must be acknowledged that this version of the objection also calls into question the applicability of the information-service-solution to real world circumstances, with all their political and empirical complications. The service would have to provide an enormously precise, detailed, and (morally problematic) intimate set of information, determining exactly what the individual contributions would have to be in order to realize the publicly-funded health care infrastructure. The information has to be that detailed, exactly in order to make sure that the individual contributions in question are never negligible. If (and yes, it remains a big “if”) such a service is delivered in this reliable and unambiguous manner, no contributor would ever be justified in judging her potential contribution a negligible one. The information service would make sure that the contribution is exactly as it ought to be, relative to the goal of efficaciously-guaranteeing the target endowment needed for the most effective strategy of helping those in need.\textsuperscript{11} Again, all this calls into question the practical feasibility of my proposal. At the same time, this vagueness is permissible at this point in the development of the argument, if one keeps in mind

\textsuperscript{10} I am indebted to a referee for this journal for framing the objection in question in terms of the negligible size of the individual benefits.

\textsuperscript{11} There’s an interesting further complication arising from this response (that I put aside in the paper). Does my proposal, counterintuitively, lead to the discrediting of any additional (voluntary) contributions into the public system, because these unforeseen contributions would then invalidate the information that the information service attempts to generate? I postpone this further complication until a later opportunity. However, the next paragraph in the text hints at the likely response, defending the moral praiseworthiness of contributors who, in a supererogatory spirit, opt to transfer even more towards the publicly-financed health care regime than would be required to precisely guarantee its realization.
that Buchanan’s own reflections are, first and foremost, conceptual and abstract ones. Real-world implementation is a different story.

Moreover, the second part of my above argument applies to the problem of negligibility of individual contributions. Recall, that the nature of the collectively-guaranteed good (universal health care) calls into question the idea that its absolute size can be fully determined. Who gets sick at what point in time and with what kind of disease? What about accidents? (Expensive) innovations in the health care industry? Regardless of how sophisticated and nuanced the information service turns out, it will surely not possess clairvoyant abilities, right?

At first, all this had looked like a severe blow to the information service proposal (and this problem has readily been acknowledged). However, the unpredictable and “unplannable” nature of individual human health works even more decisively in the other direction, rendering problematic a contributor’s personal judgment not to contribute to the public project due to some vague worry regarding the potential negligibility of her resource transfer. Realistically, the (real world) information service will provide a certain, reasonably-broad, spectrum of projected overall costs, taking the unpredictability of individual health into consideration (for, e.g., a certain number of fiscal years, based on assumptions regarding population development, life expectancy, etc.). Hitting one of the many reasonable targets within that range, will then be considered a satisfactory outcome. Given the reliable availability of the information (service) regarding the spectrum and range of efficacious outcomes, individual and private judgments in favor of non-contribution, based on the negligibility of one’s contribution, are even harder to justify.¹²

3.2. The Assurance Problem

Buchanan’s second scenario appears more promising with regard to justifying enforced beneficence. We are still deliberating whether to contribute or not from the standpoint of a beneficent agent, i.e., an agent who accepts her duty

¹² Furthermore, do not forget the feature of my proposal that, on top of the information service, we might introduce an institutional mechanism that returns individual contributions to the contributors, in case the contributions in question turned out to have remained underused. I readily admit that also this part of the account will run into very complex issues of implementation.
of charity to help the poor with their health care needs and, a fortiori, wants to do so as effectively as possible. In contrast to the first scenario, however, the beneficent agent does now seem to have a new incentive not to contribute, namely that others might be prone to free-ride. She does not know whether or not enough others will actually contribute. This time the reason why these others might fail to contribute is not the inherent paradox that comes with the universal presence of the beneficent motive to help the poor as effectively as possible but, supposedly, it is the possibility of this very motive getting overpowered in others’ practical deliberation by self-interest.

Unfortunately, Buchanan does not consistently motivate this shift towards morally-deficient agents in the set up of his thought experiments. I therefore proceed in two steps in order to execute my discussion of the assurance variety of the enforcement argument: First, I grant for a brief moment that we are now considering an imagined society, in which at least some agents might deflect from contributing to the collective health care policy because they end up being overpowered by self-interested, non-beneficent, motives. Second, however, I highlight that this very scenario is not consistent with most of the passages in Buchanan’s writings, i.e., assumptions that remain committed to the idea that we are considering a society of universally shared attitudes and motives of beneficence.

It turns out that regardless of which of the two readings we endorse, the remainder of the argument from the assurance problem is relevantly similar to Buchanan’s argument from lack of coordination discussed above and, therefore, open to an epistemic (and non-coercive) resolution. Buchanan begins the relevant argument with the reasonable claim that without the assurance that enough others actually contribute, the most rational thing to do, again, appears to be to direct one’s individual contributions to (suboptimal) projects of dispersed and small-scale health care endeavors. According to Buchanan, the assurance worry suggests that in order to achieve the optimal and most effective outcome, an agency must be established that enforces contributions in order to disarm the beneficent agents’ (rationally-warranted) incentive not to contribute. Once the

---

13 Cf. Agich (1991, pp. 191-196) for a helpful reconstruction of this reading of Buchanan’s argument.
beneficent agents rest assured that their morally-deficient, i.e., narrowly self-interested, co-citizens are forced to contribute, the former will regard their contribution as not being wasted and will contribute their share.

Granting for a moment this way of setting-up the assurance problem in the context of Buchanan’s wider argument, one first observation concerning his reflections is that they seems to assume that each individual’s contributions are strictly fixed. In particular, Buchanan appears to presume that beneficent individuals are not willing to contribute even slightly more than their “fair and equal share” of the overall sum that would have been sufficient to realize the decent minimum regime if all others had done the same right thing. If we imagine a small scale society of ten, equally well off, members and stipulate that a universal decent minimum policy costs one hundred dollars, then each individual’s fair and equal share amounts to ten dollars. If one out of the ten is overpowered by self-interested motives and prefers to keep these respective ten dollars, then the other nine would waste nine times ten dollars – as long as they remain unwilling to contribute more than the fixed amount of ten dollars. Does this fact by itself establish Buchanan’s conclusion, according to which the society in question is justified in forcing the one to be “beneficent” and to contribute her fair and equal share?

Not necessarily, it seems to me. One should stress at this point the potentially problematic aspect of the assumption mentioned above, viz., that the potential benefactors are depicted as inflexible (and unwilling!) when it comes to giving even slightly more than their fair share. In the case of our model society, the fact that the one imperfectly-moral agent fails to give results in asking the beneficent nine for individual contributions of a bit more than eleven dollars (assuming that the medical services comprised by the decent minimum remain available to all, including the one deflecting member). If, and Buchanan is committed to this assumption, the nine others are ready to act from genuinely beneficial motives (as opposed to, for example, justice-based or contractarian motives of reciprocity) then they will not, at least as long as free riding remains rare (and it has to, according to Buchanan’s own presentation of the assurance problem – even according to the first of my two readings), pull out of the collective endeavor, because a small minority of morally deficient individuals
refuses to contribute its fair shares. And the above-introduced information service will again stand ready to alleviate that deeper epistemic worry, pertaining to determining the precise level of contributions and the number of active contributors that are now necessary, in the face of non-ideal levels of beneficent compliance. Crucially, coercion remains unwarranted, even if we currently go along with Buchanan and accept that the envisioned hypothetical society consists of both, benevolent and self-interested agents.

A qualification such as “as long as a certain point of widespread refusal is not reached” is, of course, a critical feature of my proposal. I admit that. If a large number of individuals is constantly overpowered by self-interested motives and stops being motivated to meet their obligations of beneficence, then the decent minimum policy breaks down. But notice that this would be a rather unsurprising outcome and not original news at all (and it is partly for this reason that I believe that this first interpretation of Buchanan’s assurance problem cannot be what he had in mind. In line with Buchanan’s other assumptions, mentioned above, such a society would be deeply morally deficient, a point granted by all major philosophical and religious doctrines and playing a major role in Buchanan’s argument in the first place (further discussed below). Under stable moral conditions of widespread beneficent attitudes, agents are not going to be obsessed with ruling out every possible occurrence of the free rider problem as a necessary precondition for them to contribute to the project of realizing the collectively-pursued good of decent health care for all.

Be that as it may, my main response to Buchanan’s treatment of the assurance problem is more fundamental, namely that his overall presentation of the collective action problems appears to rule out the assurance problem in its traditional formulation from the get go. When Buchanan describes the society that we are supposed to imagine for the sake of his arguments, its central features contradict those versions of the assurance problem that would be required to vindicate coercion and enforcement (as opposed to establishing the information service).

Just consider that Buchanan (1985, p. 74) explicitly states that in regard to both (!) collective action problems he “proceeds on the assumption that the individuals in question are motivated by a desire to be charitable, not simply by a
desire that the needy be provided for (by someone or other).” But given this assumption that applies to his argument for enforced beneficence as a whole, Buchanan’s version of the assurance problem ultimately collapses into the epistemic and knowledge challenge, initially introduced in response to the coordination problem above. Hence, given Buchanan’s own assumptions about the motivational states of the individuals populating his envisioned libertarian society, even the “assurance problem” falls short of constituting the kind of challenge that we need to start vindicating coercively-enforced beneficence.

3.3. The Three Brothers’ Problem and Buchanan’s Assumptions

We can see this more clearly when we discuss another, closely-related, objection to my discussion of the assurance problem, namely the so-called “three brothers’ problem” in evolutionary theory; also discussed in the economics of altruism.14 This scenario poses a challenge to my critique of Buchanan because it describes another situation in which altruistic individuals seem to be rationally compelled not to engage in an action that they all acknowledge as necessary for generating a universally-desired collective outcome. We are supposed to imagine three brothers, one of whom (the “recipient”) is in some dire emergency. Let us say he fell into a pond, can’t swim, and would drown if no one helped him. Each of his two brothers (the two potential “donors”) is equally far removed from him, can swim, but is able to rescue him only by incurring some non-negligible risk to his own life.

Evolutionary theorists highlight that even if each of the two donor brothers is basically altruistic and acknowledges that his genetic endowment (shared by the recipient brother in need) will be maximally promoted only when both he and the recipient survive, this attitude alone seems to fail to get the rescuing action going in the case at hand. The result might well be, the presentation of the puzzle concludes, that the brother in need drowns in the pond, leaving only two instead of the maximum three relatives alive and in a position to pass on their genes.

14 I am indebted to an anonymous referee for this journal for drawing my attention to the three brothers’ problem and to the literature discussing it. Two varieties of the problem are discussed in two important papers by Eshel and Motro (1988a; 1988b). My discussion focuses on the presentation of the problem in Cohen and Motro (1990, p. 56).
Obviously, this constitutes an outcome that all three had rationally acknowledged to be less advantageous than the one that would have been feasible.

Why do the three brothers end up with this sub-optimal outcome? The answer is that from the narrow perspective of each of the two donors’ rationality, the individually best outcome is to stand by and let the other brother jump into the pond, let him incur the risk of drowning, and to have the three brothers survive (who then have their shared genetic endowment promulgated to the maximum extent). The evolutionary biologists’ take home lesson is that this scenario seems to always (!) support the less altruistic brother in terms of her “inclusive fitness” over other relatives who end up engaging in the risky, life threatening, rescuing action. In addition to having his brother getting rescued, the less altruistic brother enjoys the additional evolutionary benefit of having his personal genetic endowment not getting endangered by any risky rescue. In summary, Cohen and Motro (1990, p. 56) state, “this [the rescuing brother’s decision in the face of all other potential donors remaining passive] entails an even greater increase in the inclusive fitness of the relatives which decided not to offer their help. It seems, therefore, that if there is any altruistic relative in the vicinity natural selection will always favour the other selfish relatives.”

At first sight, the three brothers’ problem appears to support Buchanan (and undermine my information focused proposal) because it presents at least one case, in which some coercion and enforcement (not mere information services) seem unavoidably necessary to bring about the optimum outcome, in order to overcome the impact that evolutionary forces have on kin selection in the presence of more than one relative. Coercing one of the brothers to rescue the one in dire straits seems necessary in order to realize the outcome of one’s kin’s genetic endowment being maximally spread. Similar to old-fashioned prisoner’s dilemmas, without any enforcement mechanism, a merely suboptimal collective outcome gets produced when all agents act in accordance with what seems the most rational thing to do (from the individual perspective), i.e., to wait for others to take the risk involved in rescuing the brother. This aspect of the three brothers’ problem parallels Buchanan’s description of what is happening in the case of altruistic individuals failing to provide a shared good that they all deem worth
realizing but, due to one or the other collective action problems, are only capable of realizing if an external enforcement mechanism compels them to contribute.

In response to the three brothers’ problem, and in concluding my investigation, let me apply another time the crucial distinction between scenarios in which assurance is absent because of some collective knowledge deficit or, alternatively, because the motivational states of the agents involved are unpredictable, unreliable, and unstable. Recall my above reflections on Buchanan’s assurance problem as well as on his coordination challenge: With regard to both Buchanan (1985, p. 73) presumed that we are dealing with “a society of morally upright, altruistic libertarians,” i.e., a group of individuals, with respect to which “the barrier to successful collective action is [neither] egoism [n]or self-interest in any significant sense.” My central proposal has been that in these and in many other passages, Buchanan commits himself to a crucial and consequential presumption. If his argument for enforced beneficence includes this presumption from universal altruism, it undermines his argument for coercion and centralized enforcement mechanisms. The “morally upright libertarians” in question need institutions that overcome the distinctively epistemic deficits characterizing their predicament. Once a planning and knowledge agency provides the exact pieces of information regarding the empirical facts of what each person has to contribute in order to hit the target of effective health care provision for all, the universally-shared and acknowledged altruistic motives take care of the rest. No coercion and enforced contributing enter the picture at all. Hence, no argument is even necessary to justify such practices to begin with.

Alternatively, and this is the second horn of what we might call “Buchanan’s dilemma”, if we allow that some (many?) members of Buchanan’s envisioned libertarian society are prone to free riding, deception, etc. then this not only contradicts many other things that he says (and that I quoted above) but, more problematically, this alternative set of premises lets his argument run into the standard problem that the enforcement in question will be executed against the preferences (and “the will”) of non-consenting others, who will then simply reject the claim that they are members of “a society of morally upright, altruistic libertarians” (as it is defined by Buchanan). In that case the issue of coercion
indeed becomes a relevant one and an enforcing, not just information-providing, authority must be introduced to realize the collective goods in question by ensuring that enough others contribute. However, framing the collective-action-challenges this way would amount to Buchanan engaging the controversy concerning the enforcement of (controversial) virtues and actions; a debate in which the libertarian will readily insist that imposing beneficent actions and policies on dissenting agents is a morally impermissible thing to do on part of public institutions. Moreover, Buchanan’s writings that I currently examine do not challenge the libertarian on that front. This, in turn, lends further support to my claim that Buchanan’s overall argumentative strategy must be interpreted as resting on the alternative assumption of universally-shared beneficence amongst all parties.

Now a similar Buchanian dilemma emerges when we revisit the three brothers’ problem. While I cannot fully develop an analogously-structured response to that problem here, it should be clear at this point that, given the above reflections and claims, the three brothers case must be further specified in order to really present a challenge to the alternative solution of Buchanan’s two collective action challenges. We have to ask, if the two brothers’ (that is, the two potential donors’) problem is an epistemic predicament or a matter of internal motivational deficiencies? If the latter, then it can be readily agreed that the only way to overcome their hesitance to help their brother is an external enforcement mechanism, forcefully “coordinating” the rescuing effort of their brother and countering the evolutionary pull to free-ride by simply waiting for the other brother to take care of the risky rescue.

As I tried to highlight throughout this essay, this does not at all appear to be a formulation of the three brothers’ problem that fits Buchanan’s analogous scenario regarding health care provision and its collective action hurdles. A parallel version of the three brothers’ problem would presume universally shared altruistic attitudes and motives on part of all brothers, ruling out the desire to free ride from the beginning. Rather, the two potential donor brothers must be envisioned as facing a variety of the above described epistemic problem. That problem, however, can be resolved by the acquisition of information concerning the required act of rescue; again, an act that both are presumed to be willing to
undertake. I call this a “variety” of the epistemic problem because in the case of the three brothers we need not merely an information gathering and distributing agency – and it is also for this reason that discussing the three brothers’ problem is an enlightening exercise. Different from the good of collectively-provided health care in Buchanan’s argument, rescuing the third brother is an *indivisible* good (that is, it can only be realized by each of the two donors individually). Hence, in addition to the information concerning the exact contribution that is required to realize the desired collective good in an effective expression of beneficence, the two brothers need an unambiguous procedural mechanism that determines whose turn it is, so to speak. A lottery, for example, might be one way of settling the question of which of the two brothers actually ends up performing the rescue. Again, under the assumption of universally-shared attitudes of beneficence, this lottery is *not* insisted upon because the brothers distrust each other regarding their attitudes and motives. Rather, they need to *generate* a specific kind of belief content in order to overcome their currently vague situation in terms of their actions. After all, *it would* be an irrational waste of resources if both brothers were to jump into the pond, together overdetermining the act of rescue through their uncoordinated individual decisions.¹⁵

These additional issues regarding the three brothers’ problem are certainly important and more work needs to be done to spell out the details and their relevance to Buchanan’s dilemma. However, the main response continues to consist in the observation that given (!) the assumption that all relevant parties are predisposed altruistically, also the three brothers’ problem is susceptible to a non-coercive solution. *Enforced* contribution (enforced rescuing) is only necessary in case (some) parties’ beneficent motives and attitudes are unpredictable and unreliable.

Both, Buchanan’s society of “morally upright libertarians” and a three brothers scenario in which *all* are genuinely and reliably benevolent, indeed present collective action problems that ask for shared solutions – Buchanan has done a lot to correctly identify *that* point and its relevance for the health care debate. However, the problem, as described and contextualized by Buchanan, is

¹⁵ Other mechanisms and procedures to *generate* (not merely to distribute) the kind of information that is needed to overcome the assurance as well as the coordination problem might be better than the lottery. A question that I cannot pursue in this paper.
open to getting resolved by entirely non-coercive means. On the other hand, if Buchanan’s society and the three brothers were prone to deflection and free-riding, coercion would indeed be necessary to realize the shared goods in question. However, this latter scenario would then shift Buchanan’s argument into the familiar territory of standard political-philosophical debates on the justifiability of enforcing contested public moralities on non-consenting members of society.16

References


16 I am very much indebted to George J. Agich for insightful conversations on Buchanan’s argument and for his extensive written comments on an earlier version of my essay.


COMPENSATION AS MORAL REPAIR AND AS MORAL JUSTIFICATION FOR RISKS

A COMPENSAÇÃO COMO REPARAÇÃO MORAL E COMO JUSTIFICAÇÃO MORAL DOS RISCOS

Madeleine Hayenhjelm*
madeleine.hayenhjelm@umu.se

Abstract. Can compensation repair the moral harm of a previous wrongful act? On the one hand, some define the very function of compensation as one of restoring the moral balance. On the other hand, the dominant view on compensation is that it is insufficient to fully repair moral harm unless accompanied by an act of punishment or apology. In this paper, I seek to investigate the maximal potential of compensation. Central to my argument is a distinction between apologetic compensation and non-apologetic compensation. Apologetic compensation, I argue, is an act that expresses regret and apology by means of some offer of money, goods, or services. Non-apologetic compensation is an act that seeks to restore loss or harm without expressing regret or apology. In the paper, I defend the view that acts of compensation can be apologetic and argue that such apologetic compensation is sufficient for moral repair.

Keywords: compensation, apologies, reparations, restitution, moral repair.

Sumário. Será a compensação capaz de reparar o dano moral causado por um acto errado anterior? Por um lado, há quem defina a função da reparação como reposição do equilíbrio moral. Por outro, a perspectiva dominante acerca da compensação considera que esta é insuficiente para reparar totalmente o dano moral a não ser que seja acompanhada de um acto de castigo ou de um pedido de desculpas. Neste artigo procuro investigar o potencial máximo da compensação. Uma distinção entre compensação apologética e compensação não apologética é central para o meu argumento. A compensação apologética é um acto que exprime arrependimento e um pedido de desculpas através de uma oferta de dinheiro, bens ou serviços. A compensação não apologética é um acto que procura restaurar a perda ou dano sem exprimir arrependimento ou um pedido de desculpas. Neste artigo, defendo a perspectiva segundo a qual os actos de compensação podem ser apologéticos e argumento que uma compensação apologética deste tipo é suficiente para a reparação moral.

Palavras-chave: compensação, pedidos de desculpas, reparações, restituição, reparação moral.

* Senior lecturer (associate professor) at the Department of Historical, Philosophical and Religious Studies, Umeå University.
0. Introduction

In this paper, I am interested in whether something we might refer to as the strong compensation claim can be defended. This is the claim that full compensation is sufficient to right wrongs and to fully restore the moral situation for the victim such that it would be as if the original harm had never occurred. Feinberg, for instance, writes that compensation “functions not only to repair the damages but also to ‘restore the moral equilibrium,’ as would an apology or expression of remorse” (Feinberg, 1970a, p.74). The strong compensation claim is perhaps expressed most clearly by Goodin: “From a moral point of view, the function of compensation is straightforward. Compensation serves to right what would otherwise count as wrongful injuries to persons or their property” (Goodin, 1991, p.257). Contrast this with the weak compensation claim: that, given a previous wrong, compensation can repair the harm but not fully restore the moral balance. Gaus (1991), for example, argues that compensation can only restore moral imbalances in those cases where the right’s infringement was justified in the first place. The standard view of the repairing qualities of compensation is much closer to the weak claim than to the strong claim. It is often argued that compensation cannot address moral wrongs at all, but merely losses; or that compensation, though it contributes to righting wrongs, does so to a lesser degree than apologies or punishments do. Roberts (2002), for example, suggests that compensation must be accompanied by an apology in order to right a wrong. Griseri (1985, p.402) suggests that compensation is often demanded in lieu of complete annulment, that is, when things cannot be made fully right.

What would be required of an act of compensation in order for it to fully rectify a wrong? There are three distinct positions about what it would require to morally repair a wrong via compensation: 1) Moral reparation requires only reparation of harm (the wrong consists in nothing but the harm); 2) Moral reparation requires reparation of wrong (where the wrong is distinct from the harm); 3) Moral reparation requires reparation of both harm and wrong (where these are distinct). I shall argue position 2 is the most reasonable one. Position 1 rests on the assumption that the wrong of an action consists in nothing but the harm. This seems reasonable enough as a utilitarian assumption but seems
insufficient on any non-consequentialist account. Position 3 might seem preferable to 2 in that it covers both harm and wrong. However, in many cases the reparation of harm may not be possible or necessary in order to right the wrong. The position argued for in this paper is that full repair requires compensation such that the right kind of meaning of repair and regret is conveyed. I will argue that, if we assume an expressive interpretation of moral wrongdoings and formulate an equivalent expressive conception of compensation, then there is a case for the strong compensation claim. When compensation fails to be expressive in the right way, compensation cannot fully repair moral damage on its own.

In this paper, I focus on three clusters of problems divided over five sections. The first problem (sections 1-2) consists in finding a plausible account of compensation that could explain how wrongs could be righted. The second problem (section 3) is to work out the finer details of apologetic compensation: what makes a particular act of compensation successful in communicating regret and re-writing the message of the wrongful act? The third problem (sections 4-5) takes us back to the main question: is apologetic compensation sufficient to right a wrong, and can the strong compensation claim be defended? In the last section (5) I briefly apply this to the moral problem of risk impositions.

1.0. The Indifference Criterion.

Is there an account of compensation that could make the strong compensation claim convincing? According to Nozick, full compensation is provided when the victim is on the same indifference curve as they were before the harm occurred.

Something fully compensates a person for a loss if and only if it makes him no worse off than he otherwise would have been; it compensates person X for person Y’s action A if X is no worse off receiving it, Y having done A, than X would have been without receiving it if Y had not done A. (Nozick, 2006, p.57)

This means that I could be compensated in a number of ways, as long as I am indifferent between being harmed but compensated and not being harmed. Compensation is achieved by making the victim “as well off” as they would have
been had the original incident not occurred. The basic idea of the indifference criterion seems to appeal to one of two rather different intuitions and lines of reasoning. Let us refer to them as the “no worse off” and the “would not mind” readings of the indifference criterion.

1.1. “No Worse Off.”

The “no worse off” reading says little more than that full compensation requires that the person harmed ought to be compensated in such a way that they are “no worse off” than if they had not been harmed. This accords with Goodin’s account. For him, to compensate someone is to provide “a full and perfect equivalent” for what was lost or harmed (Goodin, 1991, p.262). Essentially, the loss of a means to a valued end is compensated for when some equally good means to the same end is offered. When such an equivalence is not possible, full compensation would require providing means to a different end that still makes the person as well off as they would otherwise have been. The point is that it is the fact that an act of compensation makes the victim as well off that gives it its reparative properties. Goodin’s account of compensation is, thus, one of a return to the previous status quo ante, a return, as far as possible, to the way things were before the loss or damage, either by offering similar means to the same end or by offering means to a different but equally valuable end (see, e.g., Chapman, 1995, p.409). Central to this account is the proportionality between loss and compensation (see, e.g., Corlett, 2002, p.148). A is a perfect equivalent for B if A is neither more nor less than B and if A restores the compensated person to the same baseline of wellbeing as before the original loss of A.

To compensate someone for something is...to provide that person with ‘a full and perfect equivalent for that thing’. If he is given more than that, we would say that he has been ‘over-compensated’; if less, ‘under-compensated’. Being bracketed as it is in between these other two notions, the notion of compensation per se clearly implies the providing of the exact equivalent – neither more nor less. (Goodin, 1991, p.262)

Could the “no worse off” version of full compensation make a convincing reading of the strong compensation claim? The idea that compensation can right
a wrong by making the victim “no worse off” only makes sense if the person suffered a loss in wellbeing (on some measure) as a result of the wronging act and that they would have been better off in terms of wellbeing had they not been so wronged. A wronging act need not, however, lead to a loss in wellbeing. It could be a case of “harmless wrongdoing” (Ripstein, 2006). Such cases of wrong could not be repaired morally via compensation since there was no loss in wellbeing to compensate for. More worrying are cases of wrongful harms that seriously harm the victim but that also lead to an increase in wellbeing overall (for the victim). In such cases, where the harmed and wronged person is made as well off, or better off, by the resulting set of events, it would not be possible to right the wrong by making the person “as well off”. The problem stems from the fact that wellbeing, as a measure, tends to be aggregative when applied to one and the same person. Furthermore, serious rights violations (that do not leave the person as well off) could be made right by simply making that person better off (in some respect). Such cases would allow for exploitations of those who are already vulnerable. If someone lacks basic needs, then providing those needs seems to make the person better off even if they are exploited. We do not want to say that if the victim is poor enough then one can do whatever one wants to them as long as their education is paid for, food is bought, shelter is provided, etc. To sum up, the “no worse off” version of successful compensation rests on the assumption that a wrong is put right because the resulting loss is repaid. But this leaves the moral wrong, over and above the harm done, basically unaddressed.

However, this objection only holds if we read the indifference criterion as a return to status quo criterion and view wellbeing as an aggregative measure. Although Goodin does suggest “a return to the same baseline of wellbeing” as a criterion, that is primarily in the context of valued ends, and he favours compensation in the form of replacing means lost with some perfect equivalent means to the same end. Furthermore, he does not argue that compensation is sufficient to right wrongs in all cases. Nonetheless, the “no worse off” criterion seems to only support the strong compensation claim if we assume the first position mentioned in the introduction: that it is sufficient to only repair the damages in order to right the wrong. Nozick partly avoids this worry by making indifference the righting criterion since a person in such a case would not be
indifferent between being wronged and not being wronged. This brings us to the “would not mind” version of full compensation.

1.2. “Would Not Mind.”

In the second reading of the indifference criterion, preferences seem to provide something of an equivalent to consent, expressing the will of the person subjected to the harm or loss. The argument goes something like this:

1. Person X is indifferent between i) no harm and ii) harm and compensation.
2. If person X is indifferent between i) no harm and ii) harm and compensation, then person X does not mind ii) harm and compensation.
3. From (1) and (2): Person X does not mind ii).
4. If person X does not mind ii), then ii) is permissible.
5. It is permissible for Y to harm X if Y fully compensates X (i.e., option ii).

The initial problem is how to make sense of premise 1. What does it mean to say that X is indifferent between “no harm” and “harm and compensation”? Nozick admits: “Shamelessly, I ignore general problems about the counterfactual ‘as well off (on as high an indifference curve) as X would have been if Y’s action hadn’t occurred’” (Nozick, 2006, p.57).

The most problematic step in the above argument is the one from “would not mind” to moral permissibility (that is, premise 4). What is it in “X would not mind being harmed and compensated” that could make harmful action morally permissible? I think we have a general tendency to think that if X “would not mind” then they would also consent to it (if asked). This seems to be what it means in everyday language. When leaving the house in the pouring rain, we might take our partner’s umbrella knowing that they “would not mind”. The reason for this, perhaps, is that we know that they have no issues about lending things or have done so willingly in the past.

On what grounds can Y presume that X would consent to be harmed if compensated? First, we can look at it “materialistically”, as in the “no worse off” argument. If the situation is objectively as good as before, then they have no reason not to consent to it. Consent is then presumed by some reference to what
is rational for X to consent to in an impersonal way, which might miss quite a few aspects that X takes into account but Y does not. Secondly, we can consider the actual preferences of X and conclude that, given those preferences, they ought to consent to A because they would prefer A. Thirdly, we can consider the preferences of a larger group, and, presuming that X’s preferences are the same as those of that group, conclude that, given the preferences of that group, they are likely to consent to A. I will address the first two in turn. The third will not be addressed separately since it is merely a weaker form of the second option.

What about the first idea that, given that the person is “no worse off”, they have no reason not to consent? This idea relies on an objective comparison of states of affairs and a judgment that they are equally good for X. The assumption is that if they are equally good then there is no reason to prefer one to the other. This might be right if the two states were precisely equally good on all relevant measures. However, usually such claims are based on more narrow measures that take some aspects into account and not others (as in the case of cost–benefit analysis, for example). If so, then there could always be extra reasons to prefer one to the other. Even between two states that are equally good in all respects, if one is current and the other is not, this could provide grounds to prefer one to the other, depending on whether the person prefers change or actual values. Thus, if we assume status quo bias (Samuelsson and Zeckhauser, 1988) or a conservative preference of existing value (Cohen, 2012), the victim could have reasons not to be indifferent between two otherwise equal states. (Furthermore, if harm and compensation were to improve the situation for the victim compared to not having been harmed at all, this would—besides being a case of “over compensation”—open up to the same kind of concerns about exploitation that I raised against the “no worse” off reading.)

I shall argue that, even if we had perfect knowledge about the preferences of X, this would not constitute sufficient grounds for inferring their consent. There are several problems with the “would not mind” argument if understood as inferred consent from preferences. First, consent is a form of permission given before the violation has occurred. But if what makes the violation permissible is an indifference between compensated harm and no harm, it is hard to see how there could be an indifference between the two that the violator could know of
before the harm (and compensation) has occurred. My point is that “indifference” is a particularly tricky sort of preference to infer because it is a comparative judgement over two states that cannot both be instantiated at the same time.

Secondly, preferences might be very weak and reflect an almost arbitrary preference in any given moment, or else they might have been swayed and influenced by biased and/or insufficient information. Suppose that a voter can choose between two candidates, P and Q. However, at a particular point in time, she has only informed herself about the policies of candidate P and not yet those of Q. Being swayed by P’s policies, she might at that point in time prefer P. This does not make it right for us to take away her right to vote and, on the grounds of her preference for P over Q, infer her consent to P. The problem, as I see it, is that inferred consent robs the agents of the choice to decide when they are sufficiently informed to consent and when to postpone a decision. Presumed consent has replaced consent proper.¹

The indifference criterion is helpful in that it takes into greater account the victim and their subjective experience of harm and compensation. But much of that is greatly weakened if the person is not consulted at all, and her preferences are inferred from an average understanding of which preferences she rationally ought to have or which preferences she is most likely to have from a sociological point of view, when in fact none of these might reflect their actual preferences. It seems, then, that compensation, if based on a hypothetical indifference, gives too much liberty to the violator to both harm the victim and to determine when they think this violation has been compensated for.

1.3. Harms and wrongs.

There are two related objections that could be raised against both versions of the indifference account. The first objection is that although the harm might be repaired, the wrong is ignored. This objection presumes a distinction between

¹ Railton makes a similar point against Nozick’s idea of compensation as a way to legitimise risk impositions without consent. Railton writes about the victim saying that there is a “preemption of his actual will and of his sovereignty, his entitlement to decide certain matters himself.” See Railton (1985, p.115). The point I am making above is that this same objection can also apply to cases where the preferences at a particular time are known.
the harm² of a rights violation (the loss of a piece of stolen property, the bruises from a physical attack, etc.) and the wrong of a rights violation (the violation of the right not to be stolen from, harmed, etc.). Olsson steals Anderson’s bicycle. In doing so, he does two things to Anderson: he causes a loss of his bicycle, and he wrongs him in violating his right to his bike. The loss could have occurred in other ways than through a crime. Replacing the bicycle with a new one would replace the loss, but not the theft. This view of compensation treats thefts no differently from losses, and, therefore, cannot claim to be a theory of moral compensation.

The second objection is related to the first. If it is sufficient to compensate another in contexts of indifference, then the very same wronging act could be repeated as long as it is compensated for. The wrongdoer might keep violating the rights of another but be ready to compensate for the act each time. Moral impermissibility is basically reduced to a transaction cost. Or, as Railton puts it in his critique of Nozick’s idea of compensation as legitimising risk impositions: “compensation is nothing but a price attached to the pursuit of one’s own ends” (Railton, 1985, p.115). We could imagine a bully that routinely violates the rights of another but also fully compensates them each time. In such a case, these exchanges of compensations seem less to communicate a wish to restore violated rights than an acceptance of a price for violating them and a lack of respect for them.

It might be the case that the “would not mind” version fares slightly better against this objection. It could perhaps be argued that a particular individual’s indifference would eventually wear off; that, after a number of compensated trespasses, indifference could no longer be achieved and such trespasses would need to stop. However, it could just as well be argued the other way around. Radzik (2004, p.147), for example, suggests that being wronged can cause the victim to doubt their value. If a victim can come to doubt their own value, then they might not be in a position to decide when compensation is sufficient. This is particularly worrying if the harm is repeated (even if compensated for). They might, as a result, lower their expectations of compensation and thereby lower

---

² “Material harm” is here intended to make it clear that I have negative physical changes in mind. I am referring to damages, injuries, losses, and the like, and not to more abstract ideas of harm such as “set-back to interests” or the like.
the “cost” of harming them. In other words, there is a risk that they might “not mind” a lower and lower level of compensation.

2. The negation criterion.

The main problem with the above accounts is that they fail to deal with the wrongfulness of the previous action. This is a general complaint often directed against compensation as a means to make amends (see, e.g., Marshall, 2001; Roberts, 2002). The question is whether accounts such as those discussed above exhaust the possibilities. According to the strong compensation claim, compensation can put right what would otherwise be held as a wrongful act. But restoring the damage done, or returning the victim to the same level of wellbeing as before, or even making them indifferent to the harm done, is insufficient to achieve this. If the strong compensation claim holds, then we need to say something rather Hegelian in nature: that the wrongful act can be annulled or negated in a way that not only restores the material harm but also restores the rights of the victims. This point has, however, mostly been addressed with regards to theories of punishment and not of compensation (see, e.g., Griseri, 1985). I will refer to this idea as the negation criterion of compensation: compensation rights a wrong when it negates that wrong.

2.1 Actions as Expressive of Meaning.

The negation criterion presumes that actions convey meaning, and Hampton suggests that all actions carry conventionally understood meanings that are accessible to all (Hampton, 2007, pp.118-120; see also Murphy, 1988, p.25). Following this idea, we have a different analysis of wrongdoing: the problem lies in what is expressed in the wrongful act. According to Hampton: “It is because behavior can carry meaning with regard to human value that it can be wrongful” (Hampton, 2007, p.119). Hegel expressed a similar thought: “His [the criminal’s] action is the action of a rational being and this implies that it is something universal and that by doing it the criminal has laid down a law which he has explicitly recognized in his action...” (Hegel, 1973, p.70, §100). In contrast to the accounts discussed above, the violation of rights of another does not consist in merely harming another but also in making a statement about the permissibility to do so. It is this message that is negated by the state in doling out punishment according to expressive theories about punishment. In Hampton’s
version, the criminal makes out that they have greater moral worth than they, in fact, do, and that the victim has lower moral worth than they, in fact, have.

A person behaves wrongfully in a way that effects a moral injury to another when she treats that person in a way that is precluded by that person’s value, and/or by representing him as worth far less that his actual value; or, in other words, when the meaning of her action is such that she diminishes him, and by doing so, represents herself as elevated with respect to him, thereby according herself a value that she does not have. (Hampton, 2007, p.126)

A similar testimonial idea is expressed by Radzik:

The wrongful act functions as a kind of testimony that this sort of treatment of the victim is acceptable. If the victim believes the testimony, if any spectators to the wrong believe it, or if the wrongdoer himself is encouraged by the apparent acceptance of his claim to superiority, then further wrongs and further harms become more likely. (Radzik, 2004, p.142)

Given this view, we can treat wrongful harms as actions that, on the one hand, cause injury or loss, and on the other, express a disregard for the rights of the other. Compensation must then seek to both cover the losses and negate what was expressed in order to annul the harm.

There is, however, an additional complication that follows from an expressive analysis of wrongdoing: the impact of such messages. In the quotation above, Radzik mentions that this message could affect the victim, spectators, and the wrongdoer himself and have effects that last long after the original harm. We have already mentioned in section 1.3 the possibility of rights violations affecting the self–esteem of the victim and that repeated harms might cause the victim to doubt their own worth and moral status. To this we can add the fear of being treated the same way in the future. The threat from the criminal is still there if the violator still condones his crime. Radzik again: “To wrong another person is to insult and threaten him. To do nothing (or fail to do enough) to correct that action is to allow the insult and the threat to stand” (Radzik, 2004, p.142). The wrongdoer might even take the lack of protest as an expression of support and make them respect such rights even less in the future. Then there is the impact on spectators who might be influenced by the criminal and take the victim’s rights less seriously. They might become more passive in the future when others are treated in similar ways. There are thus three dimensions to wrongful harms and to moral rectification: direct harm done to the victim (self-worth, fear and risk of future harms, and actual harm); effects on the wider moral community (respect for rights of others generally and respect for the rights of the victim in particular);
and the conscience and will of the violator (condoning or opposing his criminal act).

2.2 Restating what is right.

In the retributivist version, what needs to be done in order to reinstate the status of rights is to punish the criminal. And this punishment must presumably be such that it communicates to the criminal, to the victim, and to the wider community, that what has been expressed in the criminal act does not hold. Outside the criminal context, what is required is often some kind of apology. Compensation seems to fall somewhere between punishment and the informal repair of apologies.

According to Hegel, the will of the violator is not only the cause of the violation of the right but also where the wrong “resides”, and thus it is the necessary object for punishment in order to achieve a negation of the crime (see Brown, 2001, for discussion).

The sole positive existence which the injury possesses is that it is the particular will of the criminal. Hence to injure [or penalize] this particular will as a will determinately existent is to annul the crime, which otherwise would have been held valid, and to restore the right. (Hegel, 1973: 69, §99)

For Hegel, punishment is meant to annul the crime insofar as the locus of the crime lies in the will of the criminal (Hegel, 1973, pp.69-70, §§98, 99). Punishment thus denies what was expressed in the wrongful action and re-establishes what is right. Similar thoughts run through later expressive theories of punishment (see, e.g., Feinberg, 1970b; Duff, 2001, esp. 27ff): what is right is communicated to the criminal through an act of punishment to “reconnect him with correct values” (Nozick, 1981, 374ff) or to make him comply with the values of the community (Duff, 2001).

Where does this bring us with regard to compensation and the claim that compensation can right wrongs? Just as punishment is an action with a different message than the criminal act, so too can compensation offer a different message than the original wrongdoing. But there is an interesting difference. Punishment can communicate to the community and to the criminal the values of the state. It can defend the rights of the victim and communicate such rights to the community, to the criminal, and to the victim. But punishment is imposed on the
wrongdoer externally. Compensation, by contrast, can be offered by the wrongdoer himself. It can, of course, be something externally required of the wrongdoer. But it need not be. And even when forced, it can be offered out of genuine contrition. In short, compensation can communicate a change of heart where the wrongdoer now supports the rights of the victim. As an example, we can think of compensation to victims of earlier government policies that violated the rights of less protected groups in the past. The mere act of compensation signals a change of values: we no longer condone forced sterilisations, we ought not to have let insufficiently tested medical drugs be prescribed to pregnant women, etc. However, in such cases it would often be a later government rather than the original one that committed the violation.

We may now begin to address one of the objections to the strong compensation claim. The claim that compensation can only restore the harm done and not the wrong seems to overlook the expressive capacity of actions. If criminal acts and punishing acts can communicate then presumably so can compensatory acts. If an act of compensation could address both material harm and the moral message then we would have something approaching a supportive case for the strong claim of compensation.

3.0 Apologetic and non-apologetic compensation.

The preceding discussion raises a distinction between what we might refer to as apologetic compensation and non-apologetic compensation. To compensate someone non-apologetically is to offer money, goods, or services to repair damages or losses incurred by a previous act but not to admit to any wrong or to express any regret. It is often pointed out that this kind of compensation is insufficient to make amends for wrongful harms. Marshall (2001), Roberts (2002), and Lazar (2008) argue that compensation is insufficient and needs to be accompanied by an apology. Smith (2008) takes an even more pessimistic view of compensation. In a discussion of what he refers to as “compensatory apologies”, he claims that even when compensation is accompanied by an apology, such an apology “accepts no causal responsibility, admits no wrong-doing, and expresses no opinion of the underlying value.” He continues: “Apologetic language may accompany such exchanges [of goods to the injured
party], but will typically commit to little more than expressing sympathy” (Smith, 2008, p.149).

In contrast, I would like to introduce the idea of a genuine apologetic type of compensation. *Apologetic compensation* is compensation offered by the wrongdoer as a gesture expressing regret or remorse over the harming act and a desire to put things right. The regret and the wish to put things right are the very reasons for compensation. But rather than expressing this in a verbal apology, this remorse is expressed in the compensatory offering of goods, money, or services. It is acknowledging wrongdoing and expressing regret over and above any physical damages or losses.

To illustrate the difference between the two, we can compare the following two examples. First, take the case of the Bullingdon Club, the infamous Cambridge dining club known to select restaurants to vandalise after the club’s annual dinners (Ronay, 2008). The club members supposedly offered very generous payments to cover all such damages after having trashed their eating venue. Considering that this practice was repeated over many years, the compensation cannot be said to express genuine remorse. This would be a case of non-apologetic compensation. (This would be so even if the compensatory act was accompanied with some verbal apology.) Compare this to a case where a child has accidentally broken an antique gramophone that they knew they should not have played with and offers their entire collection of seashells to the owner to make up for it. In such a case, the damages are far from repaired. However, we could easily accept the gesture as one expressive of regret, of responsibility for the harm caused, and of a genuine wish to never have done it in the first place. This would be a case of apologetic compensation.

To the extent that the above examples have any intuitive support, they imply the following: 1) that there is a case for restorative apologetic compensation, that is, a case for compensation that repairs by expressing apology rather than by restoring loss; and 2) that full material compensation can be morally unsatisfactory when unapologetic. None of this should be all that controversial: in the private sphere of morality we offer flowers to make up for minor harms, as consumers we are often generously reimbursed for faulty products or mistakes, etc., and there is often a demand for apology when compensation is offered, as in
cases where wrongs toward certain groups in society have been committed by policies enacted in the past. Typically 2) would be taken to suggest a need for compensation *plus apology*. What I am claiming here is that there is an alternative to compensation-plus-apology that follows from the expressive potential of compensation: apologetic compensation. That a verbal *apology* can communicate an apologetic message is obvious. Here the claim is that compensation can be apologetic.

At this point, it might be objected that it is far from obvious that there is any important difference between *apologetic compensation* and *compensation-plus-apology*. If not, the defence for the strong compensation claim would fail, since compensation could not suffice to repair wrongs unless also accompanied by an apology either in the same act or as a separate act. We will return to this objection (section 4.1), but in order to respond to it we need to say something more about what a sincere and complete apology requires (section 3.1), how an act of compensation could be apologetic in that sense (section 3.1), and what the possible criteria for righting wrongs by means of compensation would require (section 4). I hope that, once we fit all these pieces together, the position that apologetic compensation amounts to something more and other than compensation-plus-apology becomes clearer. I also hope to clarify how apologetic compensation can repair moral wrongs *qua compensation*.

### 3.1 Meaning of apologies: regret and acknowledgement.

We will now need to say something about how an act of compensation could express an apology. It might be illuminating to say a few things about apologies in general. Smith (2008, p.80), for example, refers to the promise never to repeat the offense as one of the most important aspects of an apology.\(^3\) This is of interest

---

\(^3\) Smith’s (2008) account of “categorical apology” offers too thick a notion of apology to be of much help in illuminating the idea of apologetic compensation. He argues that categorical compensation, among many other things, requires both parties to acknowledge the same underlying values, the norms transgressed, and a shared narrative over the event. This seems to be too demanding for a general account of apologies. In most cases, it seems to suffice to acknowledge the reasonableness and entitlement of the other person’s perspective without sharing it. It must be possible to be apologetic across cultural barriers, across ethical barriers, and across language barriers. One of the beauties of apologetic compensation is that it can be expressed, and recognised as apologetic, over such barriers. I can realise that I have insulted someone by overstepping cultural norms, such as showing the soles of my feet to someone, and duly apologize for the insult and wish I had not done it without sharing those norms or the narrative. Out of Smith’s listed twelve criteria for categorical apology, several (1, 4–7) to some degree implicitly presume a common spoken language and apology as a discursive practice. See Smith (2008, 140–142) for a summary of all twelve points.
to us because one of the major objections to compensation as a way to make wrongs right was, as we saw earlier in the discussion about Nozick and Goodin, that it does not express any intention to refrain from similar acts in the future. The notion of apologetic compensation challenges this objection and rests on the assumption that, in a compensatory act, it is possible to demonstrate a will not to repeat the same offense.

The second aspect of an apology stressed in the literature is acknowledgement, in particular acknowledgement of responsibility and of some kind of culpability (see, e.g., Gill, 2000; Govier and Verwoerd, 2002). Gill (2000) lists the following five aspects of an apology:

1. An acknowledgment that the incident did in fact occur;
2. An acknowledgment that the incident was inappropriate in some way;
3. An acknowledgment of responsibility for the act;
4. The expression of an attitude of regret and a feeling of remorse; and
5. The expression of an intention to refrain from similar acts in the future.

(Gill, 2000, p.12)

A good place to start would be to assume something like the following: to the extent that an act of compensation can convey aspects 1 through 5 above, it is an apologetic act of compensation. Interestingly enough, it seems that all of these aspects can apply, to a greater or lesser extent, to compensation. Even a non–apologetic act of compensation necessarily admits to some kind of responsibility for what happened and that it did happen (1 and 3), or else there would be nothing to compensate for. What non–apologetic compensation fails to communicate is an intention to act differently in the future, acknowledgement of culpability, and remorse (2, 4, and 5). Apologetic compensation should communicate all five (with the possible exception of 3 in cases where we compensate on someone else’s behalf, such as when one government apologises on behalf of an earlier one). This seems intuitively right, but on what grounds would we say this?

As a starting point, I will propose three different aspects that could be expected to affect the apologetic impact of compensation either jointly or independently: a) the value or generosity of what is offered; b) the degree of
sacrifice or effort involved; and c) the degree of sympathy and understanding for
the nature of the harm as perceived by the victim. There might well be other
aspects, but I will focus on these three.

a) **Value or Generosity.** In monetary terms, the Bullingdon compensation was more generous than the seashell offer. Furthermore, the Bullingdon compensation probably covered all the actual costs of the damages, and possibly more. Yet it does not convey nearly as much of an apologetic message as the child’s gesture. It could be argued that the child’s offer was, in fact, more generous in that it represented a much larger proportion of all their valued possessions. One could argue that, by a gesture of compensation, the wrongdoer is putting a fictive price on a moral trespass. If this is set too low, relative to the wrongdoer’s total wealth, then he has basically given himself an allowance to repeat the same harm as long as he compensates what he himself has decided is due. Thus, it is only when what is offered is so expensive to him that he cannot afford to pay it again that true regret is expressed. The child, according to this idea, would, when offering their only seashell collection, set a price on their trespass that they would never be able to pay again and thereby be taken to be expressing a genuine regret.

This might be successful in communicating the message in one–off cases in the interpersonal case, but when it comes to a government or a large enterprise that might wrong others more than once, or more than one person, this way of expressing regret becomes impractical. A government might well offer such generous compensation as to never be able to make the same mistake again and still be able to compensate for it. The very idea of “total wealth” when applied to governments with debts and an income based partly on taxing its citizens is problematic in itself. This approach becomes even more problematic if we assume that any solutions for preventing similar problems in the future must be financed from the same source that compensation would be due from (i.e., the taxpayers).

But perhaps the idea could be much simpler: an act of compensation can convey regret if perceived as generous by the victim (for whatever reason). In this way, we could leave it open whether the generosity is to be understood in monetary terms, or symbolically, or relative to the wealth of the compensator, or
relative to the wealth of the compensated. This is too vague to be of much
guidance, but it is not obviously false.

b) Sacrifice or Effort. An alternative solution would be to sidestep the
idea of value and go straight to the cost in a more general sense, including costs
in terms of time and effort. A common objection against compensation is that the
wrongdoer might insure herself against it and be let off too easily (Marshall,
2001). This could be interpreted as wrongs being permitted at too low a cost. But
it could also be read in a different way: that compensation is emotionally or
morally too cheap. If harm has already been insured against, then there is no extra
cost involved in harming another, and the violator could remain detached from
the results of his or her actions. In contrast, it could be argued that compensation
can express apology when it involves some form of penance or sacrifice. Radzik
suggests something along those lines: “A self-imposed punishment can
contribute to reconciliation by serving as a further sign to the victim and the
community that the wrong-doer takes his error seriously and is trying to
recommit himself to a moral life” (Radzik, 2004, 149). She writes:

Giving a gift to the person one has wronged can be a form of penance,
especially if the gift requires a significant amount of effort or sacrifice from the
wrongdoer. In other cases, the wrongdoers perform some service for the person
they have wronged as a mode of penance (for example, after a quarrel with his
mother a son may dedicate his weekend to cleaning her garage). (Radzik, 2004,
144)

However, suffering by itself does not necessarily make an act of
compensation more apologetic. Adding a good portion of flagellation would not
necessarily make monetary compensation more successful in righting a wrong.
This could instead be perceived as something rather self-centred and morally
demonstrative, directing attention away from the actual victim to the wrongdoer.
On the other hand, there seems to be something morally disturbing about first
committing a wrong and then helping oneself to great privileges. Thus refraining
from gains one would otherwise have enjoyed might have something to do with
sincere regret. (This seems particularly important in cases where there might be
a link between gains for the rights violator and the harm done to the victim.)
Putting more rather than less effort into the repair could make the act more apologetic. Replacing the loss with money or what is close at hand is less convincing as an act of apology than going to the trouble to find something that is more significant to the victim, to find the “perfect thing” to heal the harm done.

c) Sympathy or Understanding of Harm. The third option for conveying apology through compensation is to compensate in a way that demonstrates sympathy for the victim (Feinberg, 1970a, p.76). This would solve the problem of compensation being emotionally too cheap and yet avoid the charge of being too focused on the violator. The act of compensation would, according to this idea, express sympathy with the losses or injuries as the victim perceived them and try to repair these harms. This plays into the acknowledgement part of apologetic meaning: to acknowledge what happened to the victim as a result of the decisions or actions of the wrongdoer. The problem with monetary compensation is that in many cases it can be handed over without any such acknowledgement, especially if it has been insured for before the harm has occurred (Marshall, 2001). Money might be essential to repair what was done, but then the acknowledgement part could still be wanting. Perhaps this is the reason for the common requirement for an apology and compensation. But compensation can also be sensitive to the harm as perceived by the victim. In particular, replacement in kind leaves a lot of room for sensitivity to subjective harm. In many cases, things lost or broken carry more than one kind of value. Perhaps the broken thing reminds the victim of the place where he proposed to his wife. If such losses are to be compensated for, then it would seem that, even if all values lost cannot be replaced, they can all be acknowledged. If the place where the victim proposed no longer exists, perhaps some other memory from the wedding can be found. Even if the compensator were to be unsuccessful in trying to find the “perfect thing”, merely having put the thought and effort into it would convey an acknowledgement of the values that have been lost.

4. Can Apologetic Compensation Right Wrongs?

---

4 Smith (2008) goes further: a genuine apology requires the violator and victim to share the same description of what happened. But this seems too strong. We can sincerely apologise for wrongs we commit against a value standard that is not our own. Perhaps the most common example would be overstepping the norms of a different culture and unintentionally insulting someone.
This paper began with what I referred to as the strong compensation claim: that full compensation is sufficient to right wrongs and to fully restore the moral situation for the victim such that it would be as if the original harm had never occurred. The most obvious problem for the traditional understanding of compensation is that compensation seems to only repair the material harm but not the moral aspects of the harm. This has commonly led to the idea that compensation is morally incomplete and needs to be supplemented with an apology. I have argued that compensation can, in fact, be apologetic and, in that way, also repair the wrong.

What are these moral aspects of harm? We borrowed Hampton's idea from the retributivist justice literature that in some (criminal) acts the wrong consists of a message about the moral value or worth of the other. Or, to put it in Hegel's cruder terms, the criminal or wrongdoer "lays down a law" in their act by suggesting a more permanent impact and influence on others. What follows from this is that in order for such a wrong to be rectified we have four aspects to take into account: 1) the message itself as an expression of the values or intentions of the criminal; 2) the moral impact of the message for the victim; 3) the moral impact of the message on the wider community; and 4) any material consequences of the act itself apart from the impact of its message. The first three would then constitute the wrong, and the fourth aspect would constitute the harm.

To what extent does an apologetic act of compensation need to remove the impact of the original message, and to what extent does it need also to repair the material harm brought about by the wrongdoing? Compare these three positions:

The Full Compensation Hypothesis: an act of compensation (provided by the wrongdoer to the victim for a previous wrongful act) can right a wrong if the previous wrongdoer compensates for all damages and fully apologizes for the whole of the wrongful act.

The Forgiveness Hypothesis: an act of compensation (provided by the wrongdoer to the victim for a previous wrongful act) can right a wrong if the compensatory act suffices to make the victim fully forgive the wrongdoer.
The Expressive Compensation Hypothesis: an act of compensation (provided by the wrongdoer to the victim for a previous wrongful act) can right a wrong if the message expressed by the compensatory act fully overwrites the message expressed by the wrongdoing act.

The underlying idea of compensation is one of replacement, or, rather, repair by replacement. It would then seem reasonable to assume that, whatever negative impact a particular act had, that impact must be “replaced” or removed in order to right the wrong. Yet the full compensation hypothesis is too demanding. It would require not only full moral repair but also that all material loss was fully replaced—with all the problems that follow. In many cases, this cannot be achieved. (There is also the additional worry over the need for double compensation: one for loss and one for wrong). For something like the full compensation hypothesis to hold, we need a weaker criterion for success than full repair of all harm. One way to do this would be to lessen the requirements for full compensation to a requirement to try to compensate as fully as possible. We could require that the compensator acknowledge all aspects of the harm and wrong committed and try as far as possible to rectify all aspects of those harms and wrongs. In such a case, it would be the trying that would right the wrong. This approach would still leave the question open as to when an attempt to rectify the impacts is sufficient to right the wrong.

Alternatively, we could start with forgiveness and argue that a wrong has been righted when forgiven by the victim. Such an account would provide a precise criterion for when a wrong has been righted, whether compensated for in full or apologetically expressed in compensation. But forgiveness does not seem to be the criterion we are looking for either. First of all, the victim might not be able to forgive the violator no matter how apologetic the compensator is or how thoughtful and generous the compensatory act. Secondly, it would render all cases where the victim did not think herself wronged, or did not know herself to be wronged, problematic. Thirdly, the most serious worry about the forgiveness criterion is that there might be a complete mismatch between when a person actually forgives another and our moral intuitions about when they ought to do so. One person might simply never forgive another even for the slightest mistake, or else they might exploit the opportunity to demand unreasonably high
compensation. Another person might easily forgive even the gravest wrongs without the wrongdoer offering much by way of either compensation or apology.

Instead, I think we need something more inter-subjectively accessible, something like a concept of “forgiveness-worthy” akin to “trustworthy”. I suggest that we assess compensation with the same criterion with which we assessed the wrong: in the message expressed. Does an act of compensation roughly convey an apologetic message equivalent to Gill’s list? Or is it a clear expression of respect for the moral rights of the other with an acknowledgement that a moral boundary was violated in the wronging act and that this boundary will be respected in the future? Is such a message one that could be understood by all that could see the wrong in the previous act? If so, then we have an act of apologetic compensation that could be recognised as “forgiveness-worthy”. Presumably, some acts of compensation will appear to us as expressive of genuine regret, a genuine will to make things right, a genuine acknowledgement of the rights of the other, of the wrong done, etc. I think we could reasonably claim that those acts also succeed in righting wrongs. Perhaps to make it more precise we could say something like this: an apologetic act of compensation rights a wrong when a sufficiently informed person (that is, one knowing about the wronging act and the compensatory act) would recognise it as a genuine act of repair that seems apologetic enough to be worthy of forgiveness by the wronged party.

What then would such an act require? Essentially, compensation is about either a) making it as good as it was or b) making it as good as it ought to have been. We have mostly been occupied with the first part: making it as good as it was. This idea is rather straightforward when it comes to losses or injuries. I replace like with like when possible but in a way that is more generous: I put in more effort to make it right, and I am very sympathetic in understanding which values were, in fact, at stake. The object that was lost might mean very different things to different individuals. A broken cup might not just be a cup but also the only memory of a beloved grandfather. A random new cup would only replace one value. But, should the compensator find a rare cup precisely to the victim’s taste and seek high and low for some other memory from this grandfather to bring back, then these two acts would jointly provide a “forgiveness-worthy” case of compensation.
I now want to say something about “a return to how things ought to have been”. This applies in particular when policies or cultures have committed systematic wrongdoings against a particular group in society. Take cases like forced sterilisations, slavery, and all kinds of hospitalisations or medical interventions that are not medically motivated. Compensation for such instances of wrongs cannot make things right by offering a return to how things were before a particular act if those acts were not a deviation from the norm at the time. Here we must draw another distinction between historic cases and current cases of systematic rights violations and rights denials. These cases must be treated differently. In demands for compensation for historic cases of wrong, part of the moral work of repair has often already begun in the form of changes of policy, values, and legal rights. There might be no need for an act of compensation to express “an intention to refrain from similar acts in the future” because the enabling values have already changed. The demands for compensation must instead be understood as compensation in the form of acknowledgement of previous wrongs, acknowledgement of equal rights, and acknowledgement of the lingering impact such previous wrongs might have had and continue to have. One such lingering impact might, for example, be the “invisibility” of a certain group. Such “invisibility” could be compensated for by greater representation, greater inclusion in history curricula, in honorary gestures like the naming of important institutions, events, and places, etc.

When it comes to compensation for wrongs that express the existing culture of values, we might have come to the far limit of the strong compensation claim. We might even state the following hypothesis: only when a moral right is publicly acknowledged in society, whether formally or informally, can an act of compensation, when sufficiently apologetic and restorative, right a wrong between victim and wrongdoer. From this, it would follow that, when a moral right is not publicly acknowledged, then any act of compensation will be incomplete and that it can only right a wrong when such rights have been acknowledged on a larger scale. This hypothesis would limit the strong compensation claim to the realm of established rights. I think this might be a plausible restriction, but I will not argue for it here.

4.1 Apologetic Compensation vs. Compensation-plus-Apology.
Let us now return to the objection raised at the end of section 3.0. This was the worry that there might not be any relevant difference between apologetic compensation and compensation-plus-apology, and that if apologetic compensation can repair moral wrongs *qua compensation* then so could compensation-plus-apology. I will address these two objections in turn.

What motivates the category of *apologetic compensation* is that non–apologetic compensation does not suffice on its own to repair moral wrongs since it does not address the wronging aspect (as argued in section 1). Furthermore, *mere apology* does not suffice to repair moral wrongs since verbal apologies could be found “too cheap” and without consequence. Both of these positions seem to be supported in the literature and intuitively plausible. What I also suggest is that *the combination of the above two*—a verbal *apology* plus a non–apologetic compensation—would not rule out each of these worries in the joint case. An apology could still be perceived as “cheap” even if there was a monetary compensation added to it but there was no action to bring about any change to the moral situation. An act of monetary compensation could still be perceived as unapologetic even if the right apologetic words were spoken. In fact, in the Bullingdon case, it seems that both verbal apology and full monetary compensation was offered—but without acknowledgement of wrong. I think this “cheapness” is not so much a case of the absence of money to back the apology up as it is an absence of *an action* that expresses a genuine regret and wish for things to be different strong enough to motivate action that makes that believable. If so, then it does not seem to be *the combination of two things*—something towards the material reparation and something towards the apology—that suffices to repair wrongs.

What I have argued is that it suffices to overwrite thewronging message with a compensatory act. This is then largely an expressive reparation. It could then be argued that it is the apologetic part that does the righting (and not the compensation). I think this is right. However, that righting could not occur by verbal apology. Rather, the righting must be expressed in an act of apologetic compensation such that one offers some service or good that expresses the intention that things be otherwise than they are as a result of the wrongful act. It is thus *an apology that takes the form of compensation* and alters the world in a
particular way that does the righting. A verbal apology could be added to such an act to make the message more explicit and less prone to be misunderstood; however, this apology would not, strictly speaking, be necessary in order to right the wrong. Is this still an act of compensation? I would think so, since what is offered is not so much a verbal statement as an act offering a service or good to the victim of wrong, which offering is deliberately aimed at addressing the impacts of a previous wrong (and which thus, from an objective point of view, would look like things changing hands, services performed, and so on).

5. Risks, Rights, and Compensation.

Should compensation be sufficient to repair moral wrongs, then this could have interesting implications for risk ethics (see Hayenhjelm & Wolff, 2012 for overview; for recent contributions see Kumar, 2015; Steigleder, 2016; Oberdiek, 2008; and Ferretti, 2016). A problem that has received a fair amount of attention in the literature is “the problem of paralysis” that arises for rights-based ethics (Hayenhjelm & Wolff, 2012, for an overview, see also, e.g., Holm, 2017; McCarthy, 1997; Teuber, 1990). If there are individual rights not to have risks imposed upon one, then any action that imposes a risk on another would be impermissible. The problem is that many everyday activities impose a risk of some kind and would thus be impermissible (unless consented to). Furthermore, many risky activities affect anonymous others whose identity will not be known until harm occurs and whose consent cannot, therefore, be attained. Nozick (2006, pp.57-84) addressed this problem in Anarchy, State and Utopia. What he suggests is something like the following:

1. Person Y wants to do A, where A is an activity that imposes a risk R on X.

2. Given that X has a right not to have risks imposed (no matter its probability or degree of potential harm), Y is not permitted to do A (unless X consents).

3. The compensation solution is that Y is permitted to do A if Y ex ante commits to fully compensate X post ante for any harm that might result from R.

Nozick’s suggestion seems congenial: we only correct harm when harm is done. Actions are not prohibited on mere grounds of what could happen. Instead,
it is the agent’s commitment to post ante repair that determines an action’s permissibility. For the tricky cases, such as those of small risks to large populations, this seems to be a promising way forward. However, this solution depends on the assumption that compensation can achieve full moral repair. I have argued above that compensation can right wrongs after they have occurred if they are sincerely regretted and come with an intention not to be repeated. This kind of compensation is not very helpful in terms of justifying risks for future harm. The only kind of compensation that can be offered before harm has occurred is a non–apologetic one. If the activity is controversial, this does not seem to offer a way out of consent. Non–apologetic compensation can perhaps offer a “good deal” that “tempts” a person to agree to a risk they otherwise would not have agreed to. But, on its own, it does not seem to justify risks. Thus, if we want to claim that impositions of risk constitute rights violations, then compensation cannot solve the problem of justification.

There is another point worth stating explicitly here. “Imposing a risk” does not typically refer to an activity with the end goal of increasing the probability of harm (the often–discussed cases of Russian roulette being, perhaps, the exception.) It is, rather, another name for the pursuit of some good that also has the negative side effect of increasing the probability of some harm. The question of justification for a risk depends on whether that good can be pursued in a less risky way or whether it can be abandoned without increasing some other risk. We have presumed that “imposing a risk” would violate a right to not be exposed to an avoidable risk of harm. But we need to say something more about the nature of this wrong. It might be helpful to think of the variety of ways in which we could wrong someone with regard to risk in a more specific manner. We could introduce a policy that imposed a risk that was far higher than it needed to be in order to save money on safety measures. We could introduce a new activity that proved very useful but would bring about irreversible damage and severe negative effects in the event of an accident—and so on. Note that such wrongs would apply to everyone exposed to the risks rather than just those that were also harmed by the risk.

Some such wrongs could be compensated for—apologetically—afterwards quite independently of any compensation of actual harm. In other words, what
seemed like a simple wrong of “imposing a risk” could be divided into two kinds of wrong-doing: 1) exposing people to a new or greater risk; and 2) bringing about actual harm for a few as a result. Both of these wrongs could be compensated for apologetically afterwards, but again, compensation could not justify them in the first place.

But this is not the only function of compensation. Compensation can also redistribute the burdens and benefits of risks. Imagine a case where everyone has consented to a particular kind of activity as being worth the risks, but that some will inevitably be more exposed to the risks than others. Some individuals might, for example, live right next to a nuclear power plant and others very far from it. In such a case, we could compensate those that are more exposed and genuinely regret that we had to expose those individuals to more risks, but without us regretting the risky activity as such. To make such a partial regret sincere, we might have to ensure that other risky policies will not similarly expose those same individuals yet again and that there is an intention to let this higher risk be an exception.

It seems that the primary role for compensation is to redistribute benefits and burdens of risk impositions: either to make the impositions of risk such that individuals would consent to them (if compensated), or to make sure that the distribution of risks can be considered fair (and those at greater risk due to greater vulnerability are duly compensated).

This, however, is compensation in a rather different sense than that of righting wrongs. In other words, compensation cannot provide a general justification for risk imposition sufficient to solve the problem of reconciling risk impositions and individual rights. That is not to say that compensation cannot play any role with regard to the ethics of risk: it is only that such a role is more limited.

References

5 See Lenman (2008) and Hayenhjelm (2012) for discussion about various distributions of risks over populations.


Compensation as Moral Repair and as Moral Justification for Risks


http://www.journals.uchicago.edu/doi/abs/10.1086/233718?journalCode=et


https://doi.org/10.1111/j.1088-4963.2006.00066.x


https://doi.org/10.1007/BF00055564


EFFORT AS RESPONSIBILITY
O ESFORÇO COMO RESPONSABILIDADE

David Jenkins*
d.jenkins.1@warwick.ac.uk

Abstract. John Roemer has created a model by which the luck egalitarian distinction between choice and luck can be used to motivate real policy decisions. By dividing society into ‘types’, Roemer suggests we are able to limit comparisons made between different people to that which they are able to control. In so doing, responsible individual action becomes the sole means by which inequalities can be justified and far more transformative redistributive legislation can be motivated. However, the model relies on two types of comparison – both within and between types – that ultimately flaw Roemer’s claims to be measuring responsible action. The model assumes that it is unproblematic to compare effort across individuals who are situated in radically unequal circumstances; it also assumes that the type can control for circumstances in a way that ignores the enormous contingency that constitutes human life. As a consequence, Roemer’s ambitious proposal fails to practically apply the choice-luck distinction.

Keywords: Luck, effort, responsibility, control, circumstance.

Sumário. John Roemer criou um modelo através do qual a distinção que o igualitarismo da sorte estabelece entre escolha e sorte pode ser usada para motivar decisões reais acerca de políticas públicas. Através de uma divisão da sociedade em “tipos”, Roemer sugere que é possível limitar as comparações entre diferentes pessoas aquilo que essas pessoas são capazes de controlar. Ao fazer isto, a acção individual responsável torna-se o único meio pelo qual as desigualdades podem ser justificadas e, simultaneamente, um meio de defender legislação redistributiva muito mais transformadora. No entanto, o modelo repousa sobre dois tipos de comparação - tanto dentro de cada tipo como entre tipos - que, em última instância, põem em causa a pretensão de Roemer de estar a medir a acção responsável. O modelo presume que não é problemático comparar o esforço entre indivíduos situados em circunstâncias radicalmente desiguais; e também presume que o tipo é capaz de verificar circunstâncias de uma forma que ignora a enorme contingência que constitui a vida humana. Em consequência disto, a ambiciosa proposta de Roemer não é capaz de aplicar na prática a distinção entre escolha e sorte.

Palavras-chave: sorte, esforço, responsabilidade, controlo, circunstância.

* University of Warwick, Department of Philosophy.
1. Introduction.

John Roemer’s innovative method of instituting equality of opportunity (hereafter, EOp) is the most practicable application (thus far) of luck-egalitarianism’s distributional principle: Namely, that people should be held responsible for that which they can control, and denied benefits or indemnified against losses over which they have no control (Roemer, 1995). In Roemer’s model, it is effort, and only effort, for which we can be held responsible. It is thus by identifying the effort different people make that issues of distributive justice are resolved. Roemer’s policy proposals are essentially a means of finding that part of our activity and constitution that is truly ours, and then to use that information to resolve distributional questions of who gets what.

The model Roemer employs does an impressive job of shifting our understanding away from merely formal understandings of EOp. This formal version attempts to exhaustively describe EOp’s content by reference to the legal or other institutional barriers that prevent individuals from pursuing careers and other competitive positions. Roemer seeks the means by which to both properly theorise the great variety of obstacles that interfere with the opportunities people face and offers concrete guidance on how to mitigate them.

A crucial component in Roemer’s methodology is the type. A type describes a group of individuals whose backgrounds are stipulated as being fundamentally identical. Put crudely, where two individual’s life chances are shaped by all the same features of some shared social background, then they can be said to belong to the same type. It is the comparisons between and within types that form the basis of Roemer’s distributive proposal and my criticisms focus on the legitimacy of these comparisons.

I begin by arguing that Roemer’s method is unable to adequately account for the effects of injustice on people’s experiences of the world around them. In particular, it cannot capture how injustices variously affect individuals’ perceptions of the possibilities or opportunities that might be available to them in whatever attenuated a form. This problematizes comparisons between the achievements of individuals who suffer from profound disadvantage and those who enjoy social, cultural and familial climates of support. However, I accept that this response remains – at this point – one of practical application: Should the
means be available to incorporate these experiences into our understanding of circumstance, then this objection is satisfied. I conclude this section with a question as to the legitimacy of allowing these experiences to be described as circumstances in the first place: Should we be describing a person’s ability to handle injustice as a part of the background?

Second, there are problems affecting the possibility of comparisons within types. Given the myriad influences that shape our lives, very few of which can be totally reduced to actions taken by the agent, the notion of control with which Roemer motivates his account seems too strong a notion to describe what we do, what we deserve and who we are. Increased effort, even from people suffering injustice, cannot necessarily be reduced to the ‘autonomous’ choices of individuals but might still be the consequence of happenstance and good fortune. I end by posing two questions – one practical, one moral – that pull my challenges together.

2. Types and Effort

Roemer’s methodology is best understood in the context of the wider substantive literature of which it forms a significant part. This literature falls under the heading ‘luck egalitarianism’, a label accorded to it by Elizabeth Anderson, a critic (Anderson, 1999). Other important exponents of luck egalitarianism include Ronald Dworkin, Gerald Cohen and Richard Arneson (Dworkin, 2000, pp. 65 – 119; Arneson, 1989, pp. 77-93; Cohen, 1989, pp. 906 – 944; Roemer, 1996, pp. 237 – 279). The founding belief of this version of egalitarianism is that individuals should not be held responsible for that which they cannot control but should be held responsible for that which lies within their power to do or not do. Luck egalitarianism therefore, generally, posits some part of the person which cannot – and should not – be assigned to the vagaries of genetic or social fortune. There is a remainder which properly belongs to the person.

Whatever people can be held responsible for, i.e. that which is not the result of circumstance, can also justify interpersonal inequality, whether understood in terms of opportunities, welfare, resources or utility. Conversely, where there is no responsibility, no inequality can be justified. On this view then, disabilities, however they be construed and insofar as they are not the fault of the person, can never be used to justify inequality. The projects we attempt and the successes we accumulate thus have a portion that is down to what we do and another
portion that belong to our circumstances. It is this issue of distinguishing the ‘parts’ of life that individuals can/cannot control which is fundamental to Roemer’s approach.

Roemer constructs a model that attempts to encapsulate that division between a person’s choices and their circumstances, between what is arbitrary and what is non-arbitrary from a moral point of view. Effort, it turns out, is the ultimate manifestation of the ‘choice’ side of that division, and comes to bear all the weight in the justification of inequalities.¹ The efforts people make, and their relationship to the circumstances within which such efforts are made, are obviously complex matters: where does circumstance end and effort/choice begin?

The background conditions individuals confront in the course of their lives are crucial in determining the choices they make and the efforts they expend. How can the seemingly infinite variety of circumstances and contingency that make up a person’s life be sufficiently manipulated and then eradicated (or at least significantly reduced) for the purposes of assigning responsibility? There needs to be a way to equalize or control for differences in background – both genetic and social – so that what remains (by way of inequality) is down to individual efforts alone. In a nutshell, this is the task Roemer sets himself.

Types are one of five key concepts in Roemer’s methodology. The others are circumstances, effort, objective and policy. Circumstances are those aspects of a situation beyond a person’s control and for which she should not therefore be made responsible – whether this be to her advantage or otherwise; effort is that ‘constellation of behaviours’ for which society does hold the individual responsible (Roemer, 2003, p 262)²; the objective is that ‘opportunity equalisandum’, the thing we wish to equalize opportunities for, whether it be welfare, longevity of life or acquisition of wage-earning capacity; and finally, the policy is the social intervention by which opportunities are equalised ‘for acquisition of the objective’ (Roemer, 2001, p 449).

Types are established via a process of collective deliberation in which citizens, drawing on relevant available empirical data, organise themselves into groups

¹ It is not effort as such but the degrees of effort that do the work; effectively, the relative amount of effort that is expended. This will be more comprehensively attended to in the section dedicated to the type and its construction.

² On another occasion these are referred to as a ‘constellation of choices’.
Effort as Responsibility

according to a list of circumstantial factors that mark out the background conditions against which they and their fellow citizens make their choices.³ So, in the case of proposals for education, relevant data might refer to parental education levels as one possible factor indicative of the environment within which children grow-up. Another could be parental income. Yet another could be IQ, used as an indicator of native intellectual potential.

After each member of a type receives bundles identical to every other member of the same type, where they end up in the distribution of that type will, so the methodology goes, be a result of choice. Roemer argues that variations of effort within the type will produce a normal distribution: On the x-axis is the effort (‘constellation of behaviours’) being measured and on the y-axis is the number of people performing at each level of effort. The majority of people will occupy the median of the distribution with minorities of outliers at either end. This median level of effort within the type describes a set of behaviours that are considered reasonable and achievable for all people confronting that set of circumstances. Whatever the objective, those minorities at the positive-end of the distribution have expended a great deal of effort. The negative-end minority has expended less effort than most in their type. Where a person appears within the distribution of their type can now be compared with other members of other types: Where people appear in the same part (centile) of their respective distributions, they have tried equally hard (cf. Risse, 2002). So, if I am one of only a handful in my type to get an A in an exam then I am trying harder than another recipient who is in a type where high grades are ten a penny: that is, are more easily available to those expending median levels of effort.

Perhaps the most interesting application of this theory is a policy designed to bring about equality of opportunity in the acquisition of wage-earning capacity (Roemer, 2002). The aim of the policy is to ‘use educational finance to equalize opportunities for wage-earning capacities among young men in the United States’. (Roemer, 2002, 464). The population is divided into four types according

---

³ Whether this should be something done by ‘the people’ or by some team of statisticians/metaphysical experts is considered by Carl Knight (Knight, 2009 187-188). The matter of Type-creative processes is problematic in a way that Roemer never addresses. This links up interestingly with my concern to address social suffering. For example, consider Iris Marion Young’s aversion to idealised forms of deliberative democracy: Societies characterised by injustice will complicate the processes that inform whatever decisions get made with respects to the type. Young, 2001. I think this issue is an important one but I bracket it for the purposes of my paper.
to the level of education enjoyed by the individual’s most educated parent. This is defined by ‘years of education’ amounting to less than 8 years for the most disadvantaged type and, for the most advantaged type, includes some experience of tertiary education (12 years and up). The outcome in this instance is taken to be the individual’s wage at age 30.

Roemer advocates a distributional difference between types such that those disadvantaged by the circumstances of their upbringing, i.e. whose parents did not receive as much education as others, benefit from additional funding from their state’s educational finances. Consequently, those whose most educated parent has less than eight years’ education would receive about five times as much funding as those whose most educated parent experienced some tertiary level education (in 1989 dollars, $5,360 compared to $1,110) (Roemer, 2002, p 464). In this example, money is used to equalise the ability of the least advantaged to compete with the most advantaged in the job market, thereby levelling the playing field.

What is particularly interesting about this account is the fact that it moves beyond a strictly compensatory paradigm. It still utilises money’s inherent fungibility but more as a restructuring device than as a compensatory mechanism: This is not a policy that motivates redistribution as a form of reparation or apology for injustice (Barry, 2006). Roemer’s proposal, at least in this case, is better characterised as attempting a preventative measure. His model uses money alongside information pertaining to the effects of disadvantage to restructure and equalise opportunities. Money is not a means to reconcile individuals to their lack of opportunity. In employing models like this, Roemer is able to escape those critics of luck-egalitarianism who claim that it focuses only on – and takes as given – after-the-fact inequalities (Anderson, 1999; Young, 2011; Scheffler, 2003.; Scheffler, 2005).

Roemer’s model can thus offer a response to Samuel Scheffler’s point that ‘equality, as it is more commonly understood, is not, in the first instance, a distributive ideal, and its aim is not to motivate compensation for misfortune. It is, instead, a moral ideal governing the relations in which people stand to one

---

4 The population is taken from a late 1960s sample in which both the educational spending of the individual’s residential district at age 16 and his wage at age 30 are available.
another’ (Scheffler, 2003, p 21). The education proposal attempts to give form to and quantify these relations, enabling assessment of the effects that differences in people's circumstances have on their lives. It also takes the further step of attempting to make these relations more just by introducing substantial differences to the level of resources allocated to differently advantaged sections of society. This could still be framed in terms of compensation: types are organised according to the ability to convert resources, and money is supplied to make up for disparity in this ability. But it is not compensation as motivated by a sense of pity or commiseration, or the recognition of a failing on the part of the individual. The different amounts of resources provided to differently advantaged individuals are intended as a way of equalising starting points so that circumstances do not unfairly prevent individuals from realising their ambitions and projects.

3. Problems with Methodology

Roemer’s model makes two comparisons using types. The first comparison is between members of the same type: circumstances having been controlled for, individuals are compared according to the effort they have expended. A person’s ability to convert resources into achievement is deemed equal within types. One’s position within the type’s distribution of achievements is therefore, according to the model, determined by one’s efforts and choices alone. Moreover, it is not just one’s position that is decided by effort but also the differences in effort and achievement between members of the same type. Compared to members of other types, a person’s trying hard might not look that impressive: For example, casual smokers might not look like they are trying hard when amongst non-smokers. But within the type, if you are trying hard, it is in terms of trying harder than others in the type.

The second comparison is between types. Where one appears in one’s own type can be compared to where other individuals appear in their type because this is the measure of the residual (choice) and not of circumstances. Two individuals appearing in the same percentile of their type’s respective distributions are deemed effective equivalents: they have expended the same degree of effort despite what might have seemed to be the case absent the comparisons enabled by the typology methodology. By making the first comparison amongst peers
where conversion ability is the same, the second comparison is also supposed to accurately capture effort alone.

Differences between types are the result of differing circumstance, whereas within the type variations occur as the result of individuals’ different expenditures of effort: The former are equalised, the latter preserved (Roemer, 1998, p 8). Take education for instance: If the median effort level of type A is ‘2’ and the median effort level for type B is ‘5’ (where this could be measuring years in post-16 education for example) then individuals within each type, operating at that median level, deserve equal reward (measured as wage after graduation for instance). This equality of reward is because they have tried equally hard. An individual of type A who expends effort level ‘5’ (perhaps attending and graduating from university) deserves more than a median individual from type B because reaching that level within type A is a more demanding achievement, i.e. type A has exercised the more responsible choice and tried harder.

Both these comparisons are problematic. Firstly, because they fail to attend to the subjectivities of individuals who occupy different ends of social hierarchies and relations of domination, and secondly because they rely on an insufficient understanding of the role of contingencies in a person’s biography. I shall argue that making the first comparison across types fails to adequately capture important differences between the efforts of differently situated individuals. In other words, people in very different types facing very different possibilities are likely to be doing very different things when they either try hard or fail to. There is something troubling about judging and comparing people across type when the ‘available behaviours’ and opportunities they face within their type vary so widely.

Extending this idea of the subjective experience of injustice, I consider the problems of comparisons made within the type. First, the extent that people’s lives are filled out in deep and substantive ways by the workings of chance, throws the usefulness of the division between effort and circumstance into doubt. Secondly, the way in which certain people respond to their being at the wrong end of social hierarchies and relations of domination can also draw on contingent elements of their personality and biography that troubles the idea that what they are expending can be labelled “autonomously chosen efforts”.

72
4. Comparisons across Types

This concern with the subjective experience of people confronting different circumstances, how they understand and try to make their way through an unjust world, does not yet issue in an obvious problem. It might, in fact, be considered a virtue of the model that it is able to evade these tricky subjective issues regarding the precise nature of what people are up to and why when they expend effort. By doing so the model motivates comparisons between phenomena that at first blush seem utterly incomparable, whilst also retaining the language of responsibility and deservingness.

From the perspective of positive policy proposals, focusing explicitly on the more objective and quantitative dimensions of effort is perhaps far more useful than entanglement with the subjective experiences and understandings of effort (Björklund, Markus & Roemer, 2012, 692).\(^5\) What individuals do when they convert resources into achievement becomes a question we avoid having to answer, precisely because of the tools Roemer supplies and the radically egalitarian policies they justify. In other words, the meaning of effort for the person who is expending it is simply too complex, too personal an issue to do much work in an account of distributive justice.

However, the effects of injustice at the level of subjective experience, of people’s dispositions and attitudes, can alter our sense of what it is appropriate to hold them responsible for. In light of this, we have a reason to question the virtue of the techniques by which Roemer attempts objectivity. Consider the following example: We compare two students, one who attends school in a poor area (type A) and another who attends a top private school (type B). Both students operate at the same low percentile within their respective types. Students in type A live in socio-economically deprived areas where opportunities both for future employment and personal development are few and far between. Students of Type B on the other hand come from wealth, privilege and a cultural climate of

\(^5\) They certainly make the case for a very progressive distributive regime in Sweden without recourse to any acknowledgment of these differences of conversion ability across types (70% of the income inequality can be explained by the ‘sterilised residual’ i.e. effort. However, perhaps this is a consequence of Sweden being an already relatively equal society (Gini Coefficient of 0.426 in late 2000s). As a result, differences in conversion-ability and the circumstances they reflect are not really such great issues. More unequal societies might suffer from this problem more easily and might produce differences between people that more readily reflect these differences in conversion ability.
support and encouragement where lots of opportunities are available for all kinds of valuable activity.

Can membership of equivalent centiles capture the effects of such differences in circumstance on the efforts made both by these students, and the people they will become? Put differently, is this arrangement of different students according to their different achievements an accurate reflection of responsible action? It would be hard to argue that their low positions within their respective types are reflective of the same kinds of choices and attendant behaviours. Rather, we might plausibly argue that where the latter could be accused of a variety of idleness for which we are not so willing to deny responsibility, something else is going on in the case of that other child, something which might even require dropping the language of responsibility – at least temporarily – and looking deeper into this student’s experiences of her everyday life.

For an example of what this ‘something else’ might consist in, imagine a child who views her overall life-situation – and those of her type – as altogether hopeless, in a way that some/many/most others of her type might not. She attends a school in a town with a plethora of socio-economic problems. There is high-unemployment, high crime, high every-rate-indicative-of-social-malaise. There are also the stereotypes that come with living amongst these conditions, that might affect her especially deeply (Steele, 2010, p 4; Shelby, 2017, p 47). This puts a dampener on her educational effort, even relative to others who might belong to the same type. Her especially acute sensitivity to what she regards as her lack of prospects for a good life after graduation, uncoupled to any especially profound academic, athletic or entrepreneurial talent, mean she cannot sum up the will to perform at an even median level – why try when mediocrity is, at best, all that awaits her? Here, acute sensitivity and (maybe) an overly pessimistic disposition are being collapsed into a description of effort’s absence. As a consequence, what describes her relation to the median-level of effort is not properly described as ‘effort cleansed of circumstance’. The question emerges: What is or is not being controlled in this instance?

---

* The effects of stereotypes, which relate to general features of a person, or an ‘identity contingency’ as Steele calls it, will be internalised in different ways. Having to deal with a stereotype of a particular kind, produces different reactions, choices and, of course, efforts.

74
It is not, of course, only acute sensitivity and/or pessimism that produces these affects. Shame, self-destructive defiance, ambient rage, boredom, loneliness, feelings of isolation or of not fitting in, an inchoate lack of esteem, the absence of identifiable role models in one’s milieu, etc. can all interact with one another and various other aspects of suffering to affect a person’s sense of agency and her performances at school and in the wider social environment.\(^7\)

To be clear, we need not accept at face value or excuse this student’s lack of effort. It is perfectly consistent to both recognise the unjust specifics of her experience and the dispositions that emerge from it, and still take her to task for them. I am not suggesting we substitute ambient pity for policy prescriptions. But we should recognise that the roots of that lack are substantially different to the roots of whatever absence more advantaged students represent and cannot, more importantly, be put down simply to a failure to try as hard as others, \textit{even those who are similarly positioned}. The focus of any alternative prescription would be to attend first and foremost to those dispositions and the environments toward which they are directed and by which they are shaped, rather than to focus on the presence or absence of something like ‘choice’.

The experience of injustice itself then precipitates behaviours, disposition and reactions we would do well not to conflate with the idea of some ‘sterilised’ effort, and which cannot be predicted according to the characteristics of a type (McNay, 2014, p 29).\(^8\) Some of these difficulties might, potentially, be handled by more fine-grained understandings of circumstance, which could then be included as a relevant aspect of the type. But some difficulties will remain dispositional in a way that cannot be described either as an absence of effort or as amenable to being factored in as a circumstance. The fact that the average/median student on the type is able to perform better is irrelevant: These are problems with which she \textit{as an individual} has difficulty contending with. The effects of injustice thus

---

\(^7\) Think, for instance, of James Baldwin’s description of African-Americans ‘living with rage’ in Harlem, his sense that his ‘life, (his) real life was in danger, and not from anything other people might do but from the hatred I carried in my own heart’. This is dramatized by Baldwin’s explosive reaction to being refused service at a restaurant, where he felt himself close to committing murder. (Baldwin, 1998, 72). It is all well and good to control for race, as Roemer proposes, and even for neighbourhood, but it seems strange to divide within such groups between those who try hard to confront those challenges described by Baldwin, and those who are, or are likely to be, defeated by them.

\(^8\) McNay advocates for an experiential model of agency that is ‘explicitly attuned to the ordinary violences of everyday life’. From this perspective, these experiences should not be relegated to the background to be controlled for across persons who are assumed to experience these in equivalent ways. Rather, they should be directly attended to as a constitutive and fundamental starting point of political theory.
interfere with the equivalence necessary to motivate the comparisons across types that are crucial to Roemer’s methodology.

Or consider the following: Imagine two individuals performing at the 90th percentile within each of their types. In type B, we again have an individual blessed with stellar educational opportunities, a supportive background at both home and school, and a wide variety of role-models and exemplars of success in her peer group. In the disadvantaged type A, by contrast, we have an individual who confronts a situation lacking support, opportunities or role-models. Nevertheless, and not necessarily perversely, the student of type A might conceivably *thrive* in just those circumstances given certain aspects of her dispositional and attitudinal constitution, which would lay dormant in less challenging environments. It is *only* when the primary motivation in the struggle to achieve academic excellence is the *escape* of the formative environments that constitute her type that this student can sufficiently gear herself toward excellence (Wright 2000).

Even with this turn to intra-type comparisons then – to which I return in more detail below – there still might be something importantly different across individuals’ experiences. This is true whether they perform well or badly within their types, precisely because those performances cannot be so easily described in terms of either the presence or absence of responsibility-producing effort. It is therefore important, given these variations, to challenge the assumption that the equivalence necessary for comparability across types is being achieved. In other words, distance from a median set of behaviours can have many reasons, only some of which have anything to do with effort or choice.

We can of course say that, for instance, both the poorly-performing students from the above example are lazy. The problem is just that lazy people in type B will achieve a great deal more in life, including a larger wage-earning capacity. But the different reasons for their idleness are relevant material for assessing the absence of efforts and the making of poor choices, especially when much of the attractiveness of the model derives from its ability to retain the moral content of concepts like desert and responsibility. By considering the concrete ways in which opportunities are not only structured but experienced, we are able to better understand the *reasons* people act the way they do and what they as individuals
understand themselves to be doing when they either succeed or fail to try hard. Many people operating at the same centiles both within and across different types are qualitatively not doing the same thing whatever the methodology of the ‘type’ might have us believe.

The assumptions of equivalence that are required to motivate comparisons both between and within types, also trouble Roemer’s discussion of an ‘equal-opportunity-for-health ethic’: specifically, people’s entitlement to treatment for diseases related to smoking (Roemer, 1995). He imagines two types derived from the factors that we are to suppose influence a person’s propensity to smoke (race, gender, occupation). The first type is a black, male, steelworker and the second a white, female, college professor. A decision to smoke will be in part influenced by these circumstances. Within these types then, the number of cigarettes smoked (the ‘achievement’ on the type’s x-axis by which effort is measured) will be a consequence of the choices individuals make. The environmental pressures supply a certain range of ‘cigarettes-smoked’ within which it is reasonable to hold members of a type responsible.

However, what a black steelworker does when he smokes and what a college professor does when she smokes is not necessarily the same, even though they might smoke equivalent numbers of cigarettes, i.e. appear in the same centile. Before comparing them directly it is necessary to look at the cultural contexts within which they smoke their respective cigarettes. Imagine two steelworkers who occupy the same type: one enjoys socialising while the other is a bit of a loner going his own way home after work. The person who likes to socialise is surrounded by smokers and, exercising restraint, only smokes a couple a night where others, again in the same type, go through a pack each. The steelworker lacking the preference for socialising has no need to exercise restraint. His preference has saved him from the pressure and thus from the need to try hard to avoid smoking.

For the white, female, college professor we can imagine a reverse of this situation. The college professor who likes to socialise does so in a way that avoids any communal pressure to smoke: Her co-socialisers do not smoke, either. In this case her preference for company ‘saves’ her from the cigarette. On the other hand, there is another professor who, like the second steelworker, prefers not to
socialise. Her method of relieving stress is to have a couple of cigarettes. The other professor who avoids smoking is not trying harder to do so. If she abandoned her preference for socialising, she too might enjoy a cigarette. It is thus her preference – although this time, for socialising – that saves her from the need to try hard.

As a result, it is not just effort that is being measured when we count the cigarettes a person smokes. Effort might still be part of it to some extent, the stress relief strategy can be resisted perhaps, but it is not as a phenomenon isolated from people’s preferences. Our efforts make sense from within the preferences we already have. We pursue our preferences in worlds to a large extent beyond our choosing: people do not invent the communal habits of either their type or their immediate work environment. Where the non-smoking loner steelworker/socialising college professor are commended for being particularly restrained in their behaviour, all that can really be said is that a preference, combined with how that preference manifests itself within a given cultural context, helps them avoid the need to restrain themselves or ‘try hard’ not to smoke (Roemer, 1996, p 246-247).

Going a little deeper into the minutiae of this (currently overly simplified) example, it is not at all obvious that socialising can be uncomplicatedly controlled for, and then added to the relevant circumstances. The white college professor, for example, might have enjoyed the company of her colleagues at a previous place of employment – perhaps it was full of pleasant and progressive luck-egalitarians. In such a situation, her ‘circumstances’ (a preference for this kind of company) saved her from smoking. In her new environment, surrounded by (let’s say) hard-nosed, committed Nozickians, she turns to her own company and cigarettes. The actions that follow on from this preference for a certain kind of company go beyond a mere ‘preference for socialising’, since that preference is always embedded in circumstances that cannot be so easily controlled for across persons of different types. ⁹ Any reference to circumstances would thus have to factor the immense variation that can occur depending on this interaction between context, preference and effort.

⁹ There is also an issue of the order in which she inhabits such environments – if the Nozickians came first, she might have started earlier and with more frequency. If the Luck-egalitarians, perhaps she’d never have started at all.
Recognition of the significant role preferences have in guiding our actions thus provides an important counter to the suggestion of equivalence which underlies the validity of Roemer’s comparisons. The smokers are supposed to be exercising responsibility toward their smoking habits: It is this which the methodology is supposed to capture. But they are trying to do different things: Cigarettes are the same means by which different ends are realised. The means only make sense given the presence of certain preferences that are in turn played out within environments over which the individual has little or no control. Returning to the education example: differences between people who in trying hard at school can express the mode of their type are not trying to do the same thing as others who use educational opportunities to escape from their type. To treat inter-type efforts as straightforwardly comparable is to miss the respective meanings that ‘trying’ hard has for the two students and thus the extent to which it classifies as trying hard at all.

Effort is not some uniform substance used in uniform ways. It is influenced by our understandings of what we are doing and why we do it, and what we bring to the answering of such questions with our dispositions and the traits of our personalities. We expend effort with certain goals in mind, reasons as to why we do what we do. The origins of our preferences are hard to pin down. Roemer’s model is an attempt to isolate just that preference to expend effort, unsullied by circumstance. However, given that our efforts are invariably influenced by how preferences take form in the unchosen social contexts in which we move, it seems there is always something outside of effort doing important constitutive work in determining whatever efforts we are capable of making.

Circumstances are the background against which people choose the projects, relationships and commitments that will matter to them and give their lives meaning. The design and pursuit of these things are only ever partially connected with ‘a choice’. We choose from within sets of constraints and opportunities over which we have very little control. Indeed, most of the time our choices are better described as dependent on prior assent to the way the world is: we develop our talents, follow our inclinations and pursue specific opportunities depending on how we perceive the world around us (Mills, 1998, p 164). The socialising, steelworker smoker could not invent the ways in which sociality was expressed in
his environment any more than could the socialising female professor non-smoker invent hers. The ways of sociality are, to some extent at least, accepted as given after which choices are made accordingly.

Time for an important caveat: Given that Roemer’s approach does not attempt an exhaustive approach to egalitarian justice there is room for him to manoeuvre. Clare Chambers criticises radical liberal egalitarianism for being insufficiently concerned with the ongoing competition that occurs after the moment when egalitarian opportunities have been initiated. But this rests on an inadequate account of what we might call the ‘just stakes’ of competition within a just society (Chambers, 2009). When, for example, so much hangs on attending a particular elite university in order to reach one’s professional potential – in Chambers’ example, Oxford – one is left to wonder about the justice of the practices within which egalitarian opportunities are to be initiated. There is nothing within Roemer’s account that explicitly precludes attendance to these other issues. However, at the very least, these issues demand attention in order to properly position the value of any EOp vis-à-vis other strategies required by justice.

There is thus the possibility that all I have done up to this point is demand from Roemer – and those convinced by his model – a more nuanced understanding of circumstances. The formulating of types is based on three suppositions: First, that we know which circumstances jointly determine a person’s ability to process resources into a given kind of achievement. For example, based on existing evidence we may know that if a child’s parents achieved some tertiary level of education, they will have a greater chance of educational success compared to children whose parents did not graduate from high school. Secondly, we assume that these circumstances can be broken down into a given number of components (I.Q., parental education levels, race, income etc.) which then take on a given vector; finally, that it is possible for this vector to take on a number of values within the population, so that we can reorganise that population into a final set of types ‘where a type consists of all individuals whose value of this vector is the same’(Roemer, 1998, p7).

So, at this stage at least, the matter might remain a practical-epistemological one: If the context within which socialising takes place, i.e. the relevant circumstance, can satisfy those three assumptions and thus be incorporated within the
algorithm of Roemer’s model than this form of the objection would have to be withdrawn. Similarly, if the experiences of injustice described above could be factored into the circumstances defining types then perhaps that idea of choosing where one ends up within the type gets effectively recaptured. In the event that such practical matters can be answered, this first set of objections loses some of its force (though I shall return to these issues with what I believe is an important reason for continued scepticism). However, the second set of comparisons, which deal with more fundamental matters of agency, cannot be so easily responded to at the ‘merely’ practical level.

5. Comparisons within Types
The last section considered the ways in which people from different types do different things with their efforts depending on the circumstances they confront. These differences complicate the comparisons Roemer’s model makes between types: what is it we are actually comparing if the performances under consideration are potentially very different? In particular, the different possible responses individuals make to the injustices they confront are not easily – either practically or morally – incorporated into such comparisons.

For comparisons made within types the notion of control carries a large part of the burden of justification. To restate once more: control is understood in contrast to luck. That which we cannot control is called ‘luck’. This collapses a great deal of different kinds of circumstance into the same category. People’s nationalities, the colour of their skin, whether or not they live in an area with a good state school are all treated as part of the general ‘jumble of lotteries that constitutes human life as we know it’(Arneson, 2008, p 80). A person’s effort is measured by removing all the effects of luck – ultimately synonymous with that for which we are not responsible – so that what remains is only that which the individual controls.

Contrasted with this account, in ordinary moral practice we do not give effort such a central a role in our understandings of responsibility. Indeed, as Hurley notes, in quotidian use of the terms, effort and responsibility do not matchup: ‘To give people what they deserve because they make more effort is not to give them what they are responsible for’. (Hurley, 2003, p 185). What people end up bringing about is not the result of effort alone: it is always mixed up with the
circumstantial aspects both of our selves and our situations (Nagel, 1982, p. 182; Miller, 1999, p 149). Roemer does not ask us to change this habit. His is an approach that uses statistical methods to glean the information he needs from actions already performed and so it need have no impact at the level of everyday practice. The quotidian can, from the perspective of his methodology, remain as is.

Nevertheless, something potentially disturbing happens to our view of individuals when responsibility is derived from the methodology of the type. Since the distribution of effort is a characteristic of the type it thus stands beyond the control of the individual. Individuals should not therefore be held responsible for that distribution but only for their place within that distribution. However, as Mathias Risse has argued, since the individual is not responsible for the behaviours of other members of their type, according to Roemer’s own methodology they should not then be held responsible for that behaviour. However, ranking them comparatively with the behaviour of one’s fellow type-members is precisely holding them responsible for choices beyond their control (Risse, 2002, p733; Hurley, 2001, p 184). By focusing on control, the individual becomes abstracted from everything which lies beyond it – including the behaviour of one’s peers – to such an extent that responsibility – as conceived by Roemer himself – seems to reduce to a vanishing point.

Roemer has responded to this argument by conceding that the effort of others within the type is out of one’s control. Nevertheless, one’s rank within the range provided by the type remains a subject of control because what the type organises is precisely the range of actions that are supposed to be reasonably available to its members: ‘My circumstances, the circumstances of others, the distribution of efforts of other types, and, if my type is large, the distribution of effort of my type, are all morally arbitrary for me, while my own effort level is not morally arbitrary for me’ (Roemer, 2003, 264). By dint of my effort I can appear anywhere I like within my type. This is what it means for effort to be under my control and this is why I can be held responsible for it.

In addition, there are resources here that can help Roemer respond to challenges others have made against the moral status of the individual at work within his methodology. For example, if, as Susan Hurley asks, ‘most of a person’s basic life
circumstances, or some very influential aspect of them, are a matter of luck’, then ‘what particular choices would he have made in the absence of luck?’ Or, put differently, ‘what choices would someone make if his life had been a very different life?’ (Hurley, 2001, p 187). Circumstantial pressures are always going to exist to some extent: we cannot escape circumstance tout court and be left with nothing but our efforts: Effort is always ‘on’ or ‘for’ something other than effort itself and a large part of this ‘something’ is what we recognise as luck.

But it is not entirely clear why this should trouble Roemer’s account. While it is true that we cannot tell what people would do were they to face entirely different circumstances, the advantage of the intra-type comparison is precisely to measure reactions to shared circumstances, the median reaction of which is supposed to be accessible to all those similarly positioned through reasonably expectable decisions and choices. Effort and control, as Roemer understands them, are always circumscribed by the range of behaviours our types, and more broadly our society, presents to us. There is nothing other than this reaction to circumstances people actually do face, which can act as a measure of a person’s deserts. And that is all his account needs.

In addition, Anne Phillips overstates the case when she suggests that, after we construct the type, what we are left with is supposed to be the aspects of the person for which we can hold her morally responsible. It is this part that is supposed to come close to some moral centre, as if ‘only these last are ‘really’ ours’ (Phillips, 2006, p 21). But in fact, Roemer really need not weigh in on these matters. An individual can regard her circumstances as providing a large part of her identity and that which has meaning for her. All Roemer argues is that these circumstances should not entitle/condemn her to better/worse outcomes.

Despite the failure of these critical accounts to fully find their mark, there remains something troubling about the sociology of the person at the centre of Roemer’s account. While it is true that collecting all forms of luck together and bundling into a package called circumstance should not blind us to the variations within what is being called luck, I do not want to criticise Roemer’s position because it collapses social and institutional injustices into the same category as ‘injustices of nature’ (Anderson, 1999, p 309). Instead, I focus on other contingencies that although affected by social injustices are not reducible to them. These are the
kinds of flukes and circumstances that have the potential to massively influence what a person is able to achieve, but over which that person has very little control.

But, it seems reasonable to ask, is it not enough that the theory can approach the idea of control? Am I not asking too much of any theory if I attach these kinds of demands to the concept of control? What Roemer’s type tries to capture are the conversion abilities of individuals similarly situated. Circumstances are held constant across its members and are then used to measure the effort people (really) make. Admittedly, this is made more complicated by the presence of those ‘incidentals’ that are constitutive of persons and their biographies. But perhaps proximity to this measure is the best we can hope for: the model goes as far as possible in determining where control begins and ends and therefore gets us as close as we can hope to the ideal of distributive justice. It is essentially as free as we can get effort to be, as free as ‘morally relevant choice’ can possibly be realised.

But this idea of ‘approaching’ control, of getting close to a part of the person that is being controlled, is already to beg the question – it is to assume something real, out (or rather in) there, waiting to be discovered by just the right application of the right algorithm. This is to accept the assumptions of the model too quickly. This move to criticise control in Roemer’s account does not commit me to saying that control does not exist in any form at all. My claim is – and need only be – that control and the functions it performs cannot be captured as this quantifiably determinate feature of agency, which can in turn be used to determine systems of distribution and reward.

Control as this approachable thing-in-the-world is crucial for Roemer’s model, conforming as that model does to what Shlomi Segall calls ‘the responsibility view’, i.e. ‘that the absence of responsibility triggers a concern with equality’ rather than the absence of equality triggering concern with responsibility (Segall, 2013, p 42). Roemer is explicit on this: ‘strictly speaking, the EOp view is not one whose fundamental primitive is equality: deservingness is fundamental, together with the normative thesis that justified inequality tracks deservingness’ (Roemer, 2003, p 279). The model therefore requires that this deserving part of the person be both real and approachable.
The child who stumbles across an author who inspires them toward their own literary aspirations, or who is motivated into learning about the cosmos after catching Carl Sagan on TV, or who in being the focus of a particular relative’s attentions is encouraged to strategically confront the limits that define their type, might all be deemed ‘lucky’. They have an additional resource that can instigate possibly remarkable achievement. But how do we factor this luck into (or out of) Roemer’s account of responsible control? Even when the decisions concerning what are to count as circumstances are made by society (or a philosopher committee) (Knight, 2009, pp. 187-190) there is simply ‘too much going on’ beyond the question of one’s own ‘hook’ for the concept of autonomously-chosen effort to hold much weight, especially the kind of weight that is supposed to determine the morally relevant part of a person’s actions (Roemer, 2003, p 270).

Luck (in its more cosmic forms) plays too large and constitutive a role in shaping and determining our life. The attempt to neutralise this luck is not to leave us with a residue of effort or control, it is to leave us with nothing. The bedrock of the person that Roemer hopes to have reached through his methodology seems more than a little chimerical.¹⁰

Roemer’s methodology seems to rely on an intuitively plausible assumption, which could conceivably discount for these kinds of luck: individuals with similar backgrounds (as defined by the type) will face similar amounts of this constitutive good and bad luck as well. The accidental boons enjoyed by some and the chance misfortunes suffered by others within one type will be of a similar, though never identical, quality. The effects of the additional disabilities/resources that are the consequence of this variety of luck will thus cancel one another out in the long run. Everybody within the same type will get a more or less equal allotment of good and bad luck. Yet this only seems plausible if we abstract both from the nature and depth of the unintended occurrences that fill out what we understand by “circumstance”, and also from the order in which they occur: Luck does not manifest itself in merely singular incidences that are contained and finite, enjoyed today by one member of Type-A and tomorrow by another member. They can often be deeply constitutive of a person’s identity, informing not just one particular moment but the narrative course of an entire life. This kind of luck is

¹⁰ This is similar to Bernard Williams’ accusation of ‘meaningless privacy’ directed at those accounts which seek too readily to extricate the controlled from the uncontrolled part of what a person is from what he most fundamentally is. (Williams, 1981, p 38).
not something that occurs equally – or even measurably – across all different members of all different types (Barry, 2007, p 42).

The examples of Sagan, inspirational neighbours and the diligence of relatives cannot be factored into Roemer’s model without rendering each type n=1. There is no such thing within us which we – or anyone else – can refer to as our own, privately-smelted ‘hook’. All such hooks are forged both within the foundry of the individual and the blaze of her surrounding environments, along with the depth of contingency that entails.

Moreover, while Roemer can effectively respond to Risse’s objection and some of the normative issues surrounding agency, once we factor in the effects of the different experiences that people have when confronting injustice this idea of appearing wherever one wants within the confines of the type becomes harder to maintain at both a practical and normative level. This free-floating, equally capable agent, even if that equality is confined to being intra-type, does not allow for the different ways in which a person might respond to her experience of injustice and the plethora of ways in which these can be said to impact on her ability to expend effort. There is no reason to assume that injustice has a singular, uniform way of being experienced. This remains true even when confined to the typical experiences of people within a single type.

To be clear, I am not committing myself to the position that control as a feature of human agency is impossible or worthless to understandings of responsibility. I am not, that is, arguing for the truth of determinism: various notions of responsibility that place control as central to issues of normativity survive my arguments. For example, Charles Taylor’s notion of ‘responsibility for self’, anchored in developing a practical ability to take responsibility for our various preferences, dispositions and senses of what is important to us, retains understandings of control deriving from an understanding of agency that is very

---

11 Barry describes the good ‘luck’ of being a library card owner or brought up in a household with an abundance of newspapers or magazines and the difference these can make to the individual. It is just such flukes that while beyond our control can make all the difference, especially for those to whom such opportunities might be few and far between. For my own part, a copy of Kerouac’s On the Road found at aged fifteen had enormous consequences that would have been denied me if I had picked it up at twenty, i.e. after the improved literary awareness granted by those five years.

12 This is as true as much for those children who imbibe an academic environment in the home as it is for those Asian children pressured into grinding away at their work without that environment as alluded to in Barry’s response to Roemer (Roemer, 1998 p 21). Both of these children have a ‘hook’ that is being forged beyond their control.
different to that offered by Roemer (Taylor, 1984). These go altogether untouched by my criticisms.

Very differently situated individuals are able to operate with some level of control, even those agents who suffer, perhaps substantially, from the injustices of their society. With this in mind, it is important to be alive to the question, non-existent within Roemer’s overarching methodology, as to whether, for example, the smoking college professor/steel worker actively embrace their preferences to smoke, or whether the pessimistic student endorses her overall attitudes to study. In addition, and importantly, these stances are themselves open to challenge: Their identification – or otherwise – with such a preference is also not the complete story.

My point is thus the simpler, less dramatic one that if it is possible to talk about control as an important aspect of agency, it is not something that can be got at by paying close attention to circumstances, followed by the elimination of the effects of those same circumstances on what is supposed to be some sterilised, purely moral place within the person.

This difficulty is altogether compounded when the precise ways in which people respond to injustices are unable to be properly countenanced and incorporated into an understanding of circumstances. Yet even if such responses to injustice – or some appropriate proxy for those responses – can be incorporated into Roemer’s model, the issue then becomes whether it is morally appropriate to divide people according to their various competencies in dealing with injustices, and then to handle that newly formed type with distributive measures. Imagine someone who, absent finding their way toward Carl Sagan or whatever other contingent saving grace their circumstance might have presented them with, develops instead a chronic, inchoate experience of marginalisation and dispossession, combined with a sense of disempowerment, all compounded by the feeling that very little can be done to save it. Should failings in this kind of competency be folded into the background against which we then judge the responsible actions such a person might still be able to perform?

*
From the above discussion there emerge two final questions – the first practical, the second normative – that advocates of Roemer’s methodology must answer. Question one: Is it possible to properly integrate the following considerations into a broader account of circumstances that could define a type: First, the general depth and breadth of luck constitutive of life; second, peoples’ experiences of injustice? Question two: Does it strike us as legitimate that people’s abilities to deal with injustice should be considered part of their circumstances, and contribute to the formation of types that then inform practical distributive outcomes? We await whatever answers might be given on behalf of Roemer’s account.

The issues regarding the chimerical normative status of the controlling person do pose a deeper problem for Roemer – and luck egalitarians more generally – whose project would seem to collapse absent the ability to effectively separate choice from circumstance. There remains something undeniably attractive about a distributive proposal that takes seriously the fact that the conceivability of certain actions is deeply dependent on one’s circumstances. Moreover, the kinds of sensitivities to subjective experiences of injustice I advocate as a necessary complement to Roemer’s methodology could, for certain policy interventions, even help guide the spending of the different amounts of distribution, targeting the various competencies individuals require to handle and confront injustices. And types, at first blush at least, do provide an intuitively useful way of thinking about parameters of conceivability.

The problem is that the constructive moves necessary for the formation of the type are not matched by an appropriately nuanced understanding of responsible agency, one which avoids placing exclusive weight on the supposed moral content that is expressed by individuals’ behaviours. As a consequence, even if the experiences associated with being on the wrong end of inequality could be more carefully considered and incorporated within the model, and assuming the wider concerns to do with ‘social suffering’ could be even partially integrated into a deeper critique of societal injustice, the overarching attempt to approach some moral content within the person is too problematic to ground those attractive proposals.
6. Conclusion

Roemer’s account is undoubtedly ambitious and performs the impressive feat of combining a notion of responsibility that is both extensive and broadly endorsed across all parts of the political spectrum, with aggressively egalitarian proposals. In particular, the way it can justify different levels of spending on policies like education before the competition in the marketplace gets going, means it is able to escape the compensatory paradigm and use redistributive measures to generate far more transformative legislation.

However, it relies on a pair of comparisons between and within types that currently fail to capture the complexities in the ways we use effort. The comparison between types neglects the profoundly different impacts that experiences of injustice can have on individuals. The questions then become whether we can or should be treating these experiences of injustice as features of a person’s circumstance.

The second comparison within types relies on an account of responsibility which is tied closely to issues of control. Where we line up within our type is said to be down to something called ‘autonomously chosen effort’. Yet even where the type is supposed to account for and eliminate the effects of background conditions, control is simply too strong a term to describe a great deal of what we bring about in the world and how we do it. Too much of who we are as people are the result of things over which we have no control and which we are also unable to systematically account for without stretching the concept of the type toward (almost) infinity.

References


LYING AND HYPOCRISY
IN POLITICS AND MORALITY,
WITH RUTH GRANT
INTRODUCTION

This dossier includes a selection of papers based on some of the presentations delivered at the second edition of the *Braga Colloquium in the History of Moral and Political Philosophy*, which is an international conference organized every year by the Centre for Ethics, Politics and Society at the University of Minho. The purpose of the Colloquium is to promote the study of the tradition of political and moral philosophy, with a special interest in understanding how its legacy has contributed to shape our institutions, culture and beliefs and thus how it can contribute to tackle the challenges our societies are facing today. To this end, every year a specific theme is chosen for the conference by taking in consideration the current political, economic and social situation in Europe (and beyond).

Taking into account the great attention that has been recently given, in the media and public sphere, to the so called issue of “Post-Truth”, the second edition of the Colloquium (which took place on the 12th and 13th of January 2016) was dedicated to the theme of “Lying and Hypocrisy in Politics and Morality.” It is a topic that, touching upon such philosophically significant concepts as truth, verisimilitude, authenticity, intention, etc., has been extensively discussed in the

---

* Integrated Researcher at CEPS (Centre for Ethics, Politics and Society), University of Minho.
** Integrated Researcher at CEPS (Centre for Ethics, Politics and Society), University of Minho.
history of Western thought, from a variety of ethical and political perspectives. Furthermore, it is a topic specially relevant today as it points to the core of some key issues and problems in contemporary political science and philosophy, as they are faced by the challenge of dealing with the deep transformations our democratic regimes are undergoing, such as their transmutations into ‘audience democracies’, the crisis of representativity, or the advent of populism. Among them we can mention: the ethical value of citizenship, the meaning of integrity in the public sphere, the uses and abuses of hypocrisy in political life, the role and nuances of political deceit and the need for compromises, or the role and scope of political representation.

Having chosen the question lying and hypocrisy in morality and politics as the main theme of the conference, we decided to invite as keynote speaker Professor Ruth Grant, who published in 1987 a path-breaking work on this topic titled *Hypocrisy and Integrity: Machiavelli, Rousseau, and the Ethics of Politics* (University of Chicago Press). Besides presenting innovative insights into the moral and political thought of Machiavelli and Rousseau, this book provided a significant contribution in placing the role of hypocrisy in politics under a new perspective: namely, it shows how political hypocrisy, under certain circumstances, does not necessarily compromise the value of integrity and can even lead to better outcomes. On the other hand, blind allegiance to an ideal of absolute moral purity and integrity can lead to disastrous political consequences (as the radicalism of the French revolutionaries like Robespierre, just to mention an obvious case, shows). In this respect, as the lucid analysis of Prof. Grant reveals, not even Rousseau completely ruled out the use of hypocrisy, despite the fact that he has been traditionally considered the prototype of the purist, the rigorous defender of the ideal of integrity and authenticity.

For this and other reasons, Prof. Grant’s seminal work has inspired and influenced many scholars in political theory and several related fields, some of whom have participated in this conference.

In her keynote address, which we publish here as the first contribution to our edited section of the journal, Prof. Grant has retaken the topic of lying and hypocrisy in politics and morality and revisited it in the light of the current political and cultural context. If in the above mentioned book she defended the
idea that a certain kind of hypocrisy can have positive effects in politics, now her attention moves on the “dangers that arise from cynicism and indifference to the truth”, on what she sees as an “attack on the value of truthfulness itself” (p. 99 of this issue). Today in effect we witness the consolidation of a new and dangerous attitude towards lying—what she calls a ‘new lying’ (p. 105 of this issue)—which seems to make the same distinction between truth and lies, fact and fiction, useless, or better irrelevant. All that matters today, in the context of an increasing polarization, is to hold a position, accepting all that comes from one’s own part and disregarding all that comes from the opposite part. But, as she points out, lying, even if under certain circumstances can be necessary in politics, always comes at a high cost, since it undermines three basic conditions that make democracy possible: trust, accountability, and a shared sense of reality.

The remaining papers of this special section of the journal deal with the question of lying and hypocrisy from a variety of angles and theoretical approaches. They can be divided into two groups: those more historically oriented, which provide new readings of canonical theorists, mostly but not only from the western tradition (Kautilya, Niccolò Machiavelli, David Hume and Jeremy Bentham) and those with a more contemporary perspective, which focus on the question of publicity and hypocrisy in public deliberation. The last paper of the first group (by Kathrin Bouvot) provides a comprehensive overview on the positions of a large number of Western thinkers (both from the past and the present), who have provided key insights on lying and hypocrisy in politics and morality. Because of that it works as a link between the two groups of papers.

In what follows, we provide a brief overview of these papers:

In “Political Pragmatism and Ethics in Kautyla’s Arthashastra: Contradiction or Complementarity?”, Kankana Saikia and Shreya Bhattacharya reflect on the possibilities to justify political pragmatism on ethical grounds relying on the renowned treatise the Arthashastra, written by Kautilya, advisor to Emperor Chandragupta Maurya (c. 317-293 B.C.E). Although much of the Western political tradition has remained oblivious to Kautilya’s contribution to realpolitik, this article clarifies to what extent his treatise provides an original and fruitful perspective on the dilemmas between ethical goals and the imperatives of political pragmatism.
The next paper, “A contingência da política e a necessidade da hipocrisia em O Príncipe de Maquiavel” (The contingency of politics and the necessity of hypocrisy in Machiavelli’s The Prince) by Albano Pina, analyses Machiavelli’s use of hypocrisy in order to bring a new light on his much discussed distinction between moral and political ‘virtù’. Pina tries to demonstrate that Machiavelli’s endorsement of hypocrisy is a part of a broader strategy he devised to allow the Prince to cope with the unpredictable effects of time, symbolized by the metaphor of ‘Fortuna’.

Alexandra Abranches’ paper “Can Hypocrisy Be a Virtue? Hume on the Morality of Princes” deals with the question of hypocrisy in the moral and political philosophy of David Hume. Through an extensive analysis of Hume’s major works, Abranches shows that Hume is able to provide a moral justification of the use of hypocrisy in the public sphere. Because of that, she concludes, his defence of hypocrisy is superior to that provided by classical realists such as Thucydides, or Machiavelli, as it is not predicated on a neat separation between morals and politics.

Benjamin Bourcier’s “Rescuing Politics from Lying and Hypocrisy: Utility and Truth in Jeremy Bentham’s Thought” explains Bentham’s opposition to lying and hypocrisy from the perspective of its incompatibility with the utilitarian principle of utility and the value of truth, which are the two pillars on which his utilitarian philosophy rests. Bourcier’s analysis proceeds in two steps: first, he reconstructs how, according to Bentham, public officials need to internalize the utilitarian code and comply with it out of sincerity; second, he reconstructs Bentham’s argument in favour of the liberty of press (despite its possible abuses), which he understands as a basic condition for a liberal government, since it permits to scrutinize and criticize the activity of public officials.

The next paper is Kathrin Bouvot’s “The Role of Lying in Politics”, which we have already mentioned above. Based on a rich variety of authors from different epochs of the tradition of western political though, Bouvot shows in which sense lying can be considered a virtuous skill, and why a strict prohibition of lying, notably defended by Kant, can be in many situations very dangerous.

Giovanni Damele’s “Argumentative Hypocrisy and Constituent Debates: the Italian Case” is one of the two papers dedicated to the question of lying and
hypocrisy in public deliberation. In this work, he discusses Jon Elster’s case for argumentative hypocrisy and applies it to the debates of the travaux preparatoires of the Italian Constituent Assembly (1946/48). Damele shows that the analysis of such debate confirms Elster’s idea that hypocrisy in deliberation can have a kind of civilizing effect and that “even if used hypocritically, impartial arguments can lead to agreements based on general interest” (p. 203 of this issue).

Finally, in his “Why the Deliberative Ideal Justifies Publicity: Even if Publicity May Undermine Deliberation”, John Pitseys discusses the epistemic arguments against publicity in public deliberation and argues that, even if they reveal situations in which closed-door deliberation can be more efficient, at the end of the day they are not strong enough to counteract its fundamental value as a legitimizing principle for democracy. Pitseys bases his argument not only on a theoretical discussion but also on the analysis of the 2010-2011 Belgian Sixth State Reform.
LYING AND HYPOCRISY IN MORALITY AND POLITICS
MENTIRA E HIPOCRISIA NA MORALIDADE E NA POLÍTICA

Ruth W. Grant
grant@duke.edu

Abstract. Hypocrisy is necessary in politics, especially in democracies, but while hypocrisy can facilitate democratic cooperation, lying tends to undermine it. There are two basic alternative possibilities for how to think about political ethics. The first begins with universal moral principles that are then applied to politics as well as to other areas of life. In the second approach, instead, each activity or type of relationship has its own moral requirements. What is it about politics that makes hypocrisy and lying either morally legitimate or morally illegitimate? For the first approach, lying and hypocrisy are vices, whereas for the second, they may be considered as virtuous under certain circumstances. Hypocrisy is necessary because political relationships are relationships of dependence among people whose interests do not exactly coincide. To secure supporters and coalition partners requires a certain amount of pretense. The case of lying, however, is quite different due to three additional characteristics of political relationships: cooperation over time requires trust; accountability requires transparency; and consensus requires a shared sense of reality. Lying undermines all three. Thus, truthfulness is among the political virtues even if exceptions sometimes must be made. Today, “post-truth” politics (“New Lying”), threatens to create a dangerous indifference to the truth and a cynical, wholesale acceptance of political lying.

Keywords: Lying, Hypocrisy, Politics, Ethics, “Dirty Hands”, “Post-Truth”.

Sumário. A hipocrisia é necessária na política, especialmente nas democracias; mas enquanto a hipocrisia pode facilitar a cooperação democrática, a mentira tende a miná-la. Há duas alternativas básicas possíveis sobre como pensar acerca da ética política. A primeira parte de princípios morais universais que são depois aplicados à política assim como a outros domínios da vida. Ao invés, de acordo com a segunda abordagem, cada tipo de atividade ou relação apresenta requisitos morais próprios. Quais as características da política que fazem com que a hipocrisia e a mentira sejam, respetivamente, moralmente legítimas ou ilegítimas? De acordo com a primeira abordagem, a mentira e a hipocrisia são vícios, enquanto que para a segunda podem ser consideradas virtuosas em certas circunstâncias. A hipocrisia é necessária porque as relações políticas baseiam-se na dependência entre pessoas cujos interesses não coincidem exatamente. Para garantir apoiantes e parceiros de coligação é necessária uma certa dose de dissimulação. O caso da mentira, contudo, é bastante diferente devido a três

* Professor of Political Science and Philosophy, Duke University.
In 1987, I published a book called Hypocrisy and Integrity: Machiavelli, Rousseau and the Ethics of Politics. That was over twenty years ago. So, when I was invited to revisit this subject, I began by asking myself whether or not I would revise my opinions today, considering that the times have certainly changed. At that time, it seemed to me important to recognize the value of a certain kind of hypocrisy, because the political dangers, in my view, emanated primarily from moralistic, absolutist and self-righteous anti-hypocrites. Hypocrisy seemed in need of defense. Then, I argued that hypocrisy is necessary in politics, especially in democracies, and that it often has positive effects. I still think that is true and that it is likely to remain true, and I will try to make the case for these claims shortly.

But today, we face a very different set of political dangers, the dangers that arise from cynicism and indifference to the truth. There is, not only an awful lot of lying going on, but also an attack on the value of truthfulness itself, along with an emphasis on manipulating appearances regardless of the realities. These developments are corrosive of politics, particularly democratic politics. As Hannah Arendt put it:

The ideal subject of totalitarian rule is not the convinced Nazi or the convinced Communist, but people for whom the distinction between fact and fiction (i.e., the reality of experience) and the distinction between true and false (i.e., the standards of thought) no longer exist (Arendt, 1973, p. 474).

While I argue that hypocrisy can facilitate democratic cooperation, lying tends to undermine it. And so, though it may seem somewhat paradoxical, hypocrisy, on the one hand, and lying on the other, do not have the same relationship to political morality.

But before turning to hypocrisy and lying in particular, let us consider how we might think about the relation between morality and politics generally. There
are two basic alternative possibilities for how to think about political ethics. The first begins with general, universal moral principles that define what makes a good human being, and these principles are then applied to different areas of life. For example, if a good person is compassionate, a good soldier should be also. And, if a good person is honest, a good politician should be also. If it turns out that a soldier must be ruthless to be an effective soldier or a politician must lie in order to be an effective politician, then the person who cares about leading a moral life must avoid war and politics. This approach is familiar in Christian pacifism, for example, or in Socrates’ claim in the Apology that the just man must avoid politics since politics requires injustice. Generally speaking, taking this first approach leads to classifying hypocrisy and lying among the vices.

The second approach is represented by Machiavelli. He is not asking how one can be a Christian prince. Rather, he makes a famous remark in chapter 15 of The Prince that a prince has to learn “how not to be good” (Machiavelli, 1979, p. 127). If there are general rules of morality, they do not apply to princes. What is good for a prince is what is necessary in politics. According to this second approach, each activity or type of relationship has its own moral requirements given by the character of the activity. Doctors should be compassionate in order to be good doctors, though soldiers may need to be ruthless. Family members should be loyal in order to be good parents, children and siblings. Instead of a single set of ethical principles applied to our different roles and relationships, each distinct domain has varying moral requirements. We have medical ethics, political ethics, legal ethics, business ethics, and so on, and not just ethics per se.

This doesn’t mean that there are no commonalities. Activities can be grouped in various ways. Competitive endeavors share certain characteristics - for example, sports, politics, business, and law. And cooperative endeavors share certain characteristics that require different sorts of virtues as well – for example, families, teams, education and politics. Note that politics is included in both categories, as a competitive and a cooperative endeavor. The complex character of political relationships is part of what makes the subject of morality in politics so difficult.

If we take the second approach to the question of the relation of politics and morality, our question now becomes: “What is it about politics that makes
hypocrisy and lying either morally legitimate or morally illegitimate?” This question has been asked before. Max Weber, in *Politics as a Vocation*, attributes the complexities of political morality to the fact that violence is the characteristic means of political power (2004, p. 33). Michael Walzer, in “The Problem of Dirty Hands” (1973), takes a similar view. Political relationships are relationships of power backed up by violence.

A very different perspective emerges by viewing political relationships, not so much as relationships of power, but as relationships of dependence. All politicians need supporters and coalition partners. Moreover, they must win their voluntary cooperation. It cannot be forced. While political actors are competitors who do not share the same interests and aims, they depend nonetheless on one another’s cooperation. Political relations are unlike friendships: friends need one another and want to benefit each other. And political relationships are unlike enemies: enemies do not need one another and wish each other harm. Politics is something in between: friendly relations among people who are not friends – people who need one another, but seek their own benefit. It is this dependency among people with conflicting interests that is the unique characteristic of politics.

This characteristic dependence explains where hypocrisy comes from in political life. With friends, you can be honest. With enemies, you can be honest. But with supporters and allies, you might need to be hypocritical. Politicians need to flatter supporters. To secure the cooperation of allies, we use diplomacy. And to act diplomatically means not being entirely frank. Democratic politicians are especially dependent on supporters and coalition partners. They can do very little alone. One should expect, therefore, to find quite a lot of hypocrisy in democracies.

The link between dependence and hypocrisy is one lesson of Machiavelli’s discussion of political morality in *The Prince*. Machiavelli begins the section of the book about lying, hypocrisy and other aspects of political morality (beginning with chapter 15) by telling the reader that he is now turning from discussing war (enemies) to discussing relationships with “subjects and allies,” that is to say, political relationships. Rousseau teaches the same lesson in the *Discourse on the Origins of Inequality*. People become hypocritical when society has developed to
the point where we are all dependent on one another: “... [Men] must therefore incessantly seek to interest [others] in their fate, and to make them find their own profit, in fact or in appearance, in working for his (Rousseau, 1992, p. 52). We need one another, but our interests do not coincide. Hypocrisy is what allows us to cooperate in these circumstances.

A simple example will illustrate the point. Imagine a meeting of a local school board where redistricting plans are under consideration. The room is full of parents advocating one plan over another. In fact, all of the parents are there because they do not want their own child moved. But not one of them says that. If everyone made claims on the basis of their own personal interest, the meeting would get nowhere. Instead, they argue that one plan will make more efficient use of the school buses or another plan will help meet diversity goals, and so on. These claims are not lies, and they are all acceptable public justifications. But the parents articulating these positions adopt a pretense of concern for the public good. They are hypocrites. This is a kind of hypocrisy that is necessary for democratic cooperation and, indeed, for any kind of political cooperation.

This was one of the basic arguments of the book published 20 years ago. Hypocrisy is necessary due to the character of political relationships as relationships of dependence. For this reason, it makes no sense to condemn all hypocrisy in the name of some kind of moral absolute. Sometimes, as in the case of the school board deliberation, hypocrisy is perfectly legitimate and sometimes not. The question is how to distinguish legitimate from illegitimate kinds of hypocrisy.

The case with respect to lying is quite different as is the contemporary political context. The news confronts us daily with the political dangers associated with cynicism and indifference to the truth. The case for truthfulness in politics begins with the same approach that we took when considering hypocrisy. Beginning with the nature of political relationships, the question is: what does the character of politics tell us about the legitimacy or illegitimacy of lying? There are three additional characteristics of politics that are important here. Politics requires trust so that people can cooperate over time. Politics requires accountability in order to check abuses of power. And politics requires a shared sense of reality in order for some kind of consensus to be achieved and for
deliberation to be possible. Without common ground, there can be no deliberation. It should be clear that all three of these requirements depend upon truthfulness. Trust cannot be maintained unless promises are truthfully entered into. Accountability depends upon transparency. And there can be no deliberation without some agreement that standards of truth will govern the discussion. Lying corrodes the conditions that make politics, and especially democratic politics, possible.

It could be objected that we see lying in politics all of the time. Far from rendering politics impossible, lying is endemic to politics. Here is Hannah Arendt again:

No one has ever doubted that truth and politics are on rather bad terms with each other, and no one, as far as I know, has ever counted truthfulness as among the political virtues. Lies have always been regarded as necessary and justifiable tools not only of the politician’s or the demagogue’s but also of the statesman’s trade (Arendt, 2006, p.223).

Her metaphor of lying as a “tool” of the politician’s “trade” suggests that lying is a morally neutral technique. The ethical question is whether lying is used to good or bad ends; by a statesman or a demagogue.

I take issue with this view. Lying in politics may sometimes be necessary, useful, or justifiable, but it always comes at a cost for the reasons I just gave – it undermines the conditions for politics. Truthfulness is among the political virtues even if exceptions must sometimes be made.

This may seem somewhat paradoxical. Lying can be the right choice, but it is wrong nonetheless. Politicians need to get their hands dirty sometimes. When they do, they need to recognize that they have done something morally wrong, even if it is politically justifiable. Michael Walzer calls this the “problem of dirty hands.” As Machiavelli says, the Prince must learn how “not to be good” – but the standard for goodness remains. In other words, when politicians lie, they must recognize that they have breached a fundamental value. It is important to care about the truth in politics.

This is what we are losing in the United States today. There has been a dramatic change in the quantity and quality of lying in politics. Formerly, politicians would lie to serve particular purposes and try to conceal their lies. Their lying was instrumental, and if they could achieve the same goal truthfully,
they would. Now, lying is brazen, fortuitous, and it seems directed at undermining the very distinction between fact and fiction, truth and lies. One of its key characteristics is that, when the truth is exposed and the facts are publicized, nothing changes. This is what I will call the “new lying.” And its political consequences could be very serious. Remember Arendt’s warning quoted at the beginning of this article: the ideal subject of a totalitarian regime is the person “for whom the distinction between fact and fiction (i.e., the reality of experience) and the distinction between true and false (i.e., the standards of thought) no longer exist.”

One indication of this shift is that the Oxford English Dictionary chose “post-truth” as the word of the year for 2017. They define “post-truth” as “relating to or denoting circumstances in which objective facts are less influential in shaping public opinion than appeals to emotion and personal belief.” The Dictionary found two spikes in searches for this word; one during the Brexit campaign and the other when Donald Trump was nominated for the Presidency. It is important to recognize that this is not just an American phenomenon. Nonetheless, Trump is clearly the prime example of the “new lying.” There are hundreds of examples of his falsehoods, with new ones added almost daily. The constant barrage of claims and counterclaims normalizes the lying and confuses the public. Moreover, Trump’s advisors defend his lying, not by claiming that he is telling the truth, but precisely by claiming that his lying doesn’t matter. They say publicly without shame, “Don’t take him literally,” “Many people make statements without facts to back them up,” “There are no facts anymore, what counts is what people will believe.” Politics becomes about making images, not confronting reality. Politics comes to resemble televised professional wrestling (one of Trump’s involvements) or the tabloid press. The audience knows the show is fake and the readers know the stories are largely false, but they don’t care.

Why has there been this dramatic change in the character of lying? One plausible explanation would be that it is an effect of the increasing polarization in American politics that began to intensify about ten years ago. Polarization means that parties have less and less common ground; less and less sense of a shared reality. And because the media has also become partisan and polarized, there is no shared authoritative source of facts and news trusted as reliable by all parties.
Moreover, political parties thoroughly distrust each other and see their political opponents as enemies. Under these circumstances, lying to discredit the opposition or to win at all costs seems justified. Politics has become warfare where everything seems to be at stake in every conflict over every issue.

However, the problem with attributing the “new lying” to political polarization is that it would lead one to expect that both parties would be equally prone to “post-truth” politics. Why, then, do we find the phenomenon more pronounced on the right than on the left? This is a question for which I do not have an adequate response. But whatever the explanation, undermining truth as a common public standard is particularly threatening in democracies, because democracies require a common language for deliberation; truth-telling for accountability; and trust for bipartisan action.

What I have called the “new lying” seems to have come out of nowhere. But it is not unprecedented in history. The government of the Soviet Union was well known for rewriting history to serve its own political purposes, for example. George Orwell makes clear in his fictional account of an authoritarian regime, 1984, that controlling the “truth” is an essential aspect of modern tyranny (Orwell, 1971). Going further back, Thucydides told how the breakdown of a common language was tied to polarization in times of civil war. He described, not lying exactly, but the re-description of reality.

So the condition of the cities was civil war ... men inverted the usual evaluation of actions ... Irrational recklessness was now considered courageous commitment ... Moderation was a cover for lack of manhood ... while senseless anger now helped to define a true man ... the man of violent temper was always credible, anyone opposing him was suspect (Thucydides, 1982).

These precedents are not encouraging – authoritarianism and civil war. It would be well worth examining the link between lying and the breakdown of democratic politics. Lying always comes at a cost because healthy politics depends on politicians – and citizens – who value the truth and recognize that there is an objective reality, a world of facts we can appeal to in order to settle our differences.

I have argued that we should judge political ethics by considering the distinctive characteristics of political relationships. One of these is that political relations are relations of dependence. Dependence fosters a kind of hypocrisy
which can contribute to cooperation. Lying, on the other hand, threatens political cooperation by undermining the conditions for it: viz. trust, deliberation and accountability. And so, all forms of deception in politics are not alike. We live in a world where the quality and quantity of political deception is changing and growing. At this particular moment, there couldn’t be a better subject for a journal volume than: “Lying and Hypocrisy in Politics and Morality.”

References


POLITICAL PRAGMATISM AND ETHICS IN KAUTILYA’S ARTHASHASTRA: CONTRADICTION OR COMPLEMENTARITY

Abstract. Is there always an inherent contradiction between political pragmatism and ethics? This paper explores this through the classical Indian realist political treatise, the Arthashastra, written by Kautilya. He was the key advisor to Emperor Chandragupta Maurya (c. 317-293 B.C.E), credited with first consolidating the Indian subcontinent into an empire under centralized control. Kautilya is widely agreed to be a pioneer in developing a ‘science of politics’ - a great departure from the hitherto existing moralistic, spiritual, ascetic preoccupations and, religious speculations of Indian political philosophy. Kautilya blurred the distinction between morality and political utility. Much of political theory tradition remains oblivious to Kautilya’s contribution to morality in realpolitik and his continuing influence in non-Western societies. His treatise is a significant point of entry into exploring the dilemmas attached to active pursuance of ethical goals while also balancing the imperatives underpinning political pragmatism- a concern which has carried on to this day.

Keywords: political pragmatism, morality, means, ends, Arthashastra.

Sumário. Haverá sempre uma contradição intrínseca entre o pragmatismo político e a ética? Este artigo explora este tema através do realismo do tratado político clássico indiano, o Arthashastra, escrito por Kautilya. Ele era o conselheiro principal do Imperador Chandragupta Maurya (c. 317-293 B.C.E), tendo sido reconhecido como o primeiro a consolidar o subcontinente Indiano como um império submetido a um controlo centralizado. É amplamente consensual que Kautilya foi pioneiro em desenvolver uma “ciência política”, a qual representava um grande avanço, quer em relação às preocupações moralistas, espírituais e ascéticas da Filosofia Política indiana, quer em relação às suas especulações religiosas, as quais tinham sido até então dominantes. Kautilya esbateu a distinção...
entre moralidade e utilidade política. Muita da teoria política tradicional continua a ignorar o contributo de Kautyla relativamente à a moralidade na realpolitik e a sua persistente influência em sociedades não ocidentais. O seu tratado representa uma via significativa para explorar os dilemas relacionados com a procura ativa de finalidades éticas, sem perder de vista os imperativos atinentes ao pragmatismo político – uma preocupação que se manteve até aos dias de hoje.

**Palavras-chave:** Kautilya, Arthashastra, não-ocidente, ciência política, pragmatismo político, ética.

### 0. Introduction

The realm of political philosophizing and practice is inherently replete with certain conundrums. One of the fundamental dilemmas which has puzzled thinkers across centuries has been the perceived dichotomy between political pragmatism on the one hand and ethical considerations on the other. In the past, these two aspects have usually been viewed as mutually exclusive, that is, one has to be/can be obtained at the expense of the other. In practice, this took the form of creating a hierarchy. Opposing sides would either try to elevate political pragmatism over morality; or portray ethics as a higher order of operation thereby relegating practical politics to a lower level of consideration. This often results in straitjacketing both into non-permeable moulds. What happens in such a case is that a relationship and interaction between the two is not only constricted, rather it is disregarded as a possibility. In Western political philosophical tradition for instance, Plato’s concept of politics rests on a thoroughly ethico-moral foundation. The Philosopher King was to embody the highest virtues that would set him apart from the ordinary masses. Hence, to think about and to practise politics would put high moral demands on the thinker/practitioner. The evolution of western political philosophy embarked upon a new track which emerged during a period of flux, as embodied by the Renaissance. Niccolò Macchiavelli controversially revolutionised the idea of morality by demarcating an amoral (distinct from immoral) practice of politics to which universal standards of morality are not seen to be applicable.

### 1. Political Philosophy And The Non-West

Beyond the West, debates regarding the relationship between morality and politics have been an integral part of philosophical deliberations. However, given the Eurocentrism of political philosophy, where mostly Western thinkers
and thought traditions have been accorded primacy, such non-Western deliberations and interventions have often been ignored or sidelined. This omission has significant implications for the theory and practice of politics. Postcolonial scholars have been committed to engaging with “universals” such as a universal conception of man and the universal ideals of morality that are seen to emerge from there. This concept of universal becomes problematic because such “universals” are mostly based on European Enlightenment traditions that do not necessarily reflect the notions of the individual and structures of society and polity prevalent in the non-West. Also, to assume that questions on morality in politics as well as the explanations put forward, have originated in Europe is to fall into the trap of a form of thinking which can be simplistically summed up as “first in the West, then elsewhere” (Chakravarty, 2007). In fact, according to Arnold Toynbee, most of the world’s civilisations, have been non-Western. Three of the five surviving civilizations - the Islamic, the Hindu, and the Far Eastern - are non-Western, and another - the Orthodox Christian - has absorbed many non-Western characteristics (Palmer, 1955). India holds a significant place in the enduring struggle between the East and the West, the reason being that India has been a meeting ground for many different cultures and civilizations. While the Western impact on Indian political thought has been great, India’s contribution to Western political thinking is much greater than is generally realized. In a different but relevant context, John A. Lukacs wrote, "For the West can interpret the West; it can only describe the East (Lukacs, 1953). Hence, it is important to critically move beyond such a Eurocentric bias and incorporate these voices within the ambit of political and moral philosophizing because they not only open up new avenues of deliberation but also provide the possibility of alternative answers to existing perennial questions. Indian philosophical thought has often been characterized as being preoccupied with spiritual, metaphysical and theological concerns. This is a misconception; because to have survived this long as a civilization, practical issues and concerns necessarily have had to be very much a part of the key Indian texts and discourses. This is evidenced by the texts of the Arthashastra tradition that this paper seeks to discuss. This tradition more or less borrowed its epistemology from the philosophy of Lokayata. Lokayata philosophy did not believe in concepts like rebirth, Karma, existence of God, caste system and divinity of the Vedas because their efficacy could not stand the
test of reason and they were not amenable to sense perception. They argue that this world is developed due to combinations of different human actions and efforts; and there is no omnipotent divine agency influencing it. This argument giving primacy to human Reason over divine predetermination was one of the most prominent features of the European Enlightenment. This example illustrates the prejudice that guides the first-in-the-West-then-the-rest ethnocentric understanding of certain core tenets and assumptions underlining political philosophy.

This argument is also important to take into cognizance because when we talk about morality, we have to recognize that morality does not take shape in a vacuum. Morality is essentially derived from socio-cultural, historical norms and ethos, and is very context-specific. This argument is given shape in the theory of ‘ethical relativism’ which holds morality to the norms of one’s culture; that is, whether an action is considered right or wrong depends upon the moral norms of the society in which it is practiced (Brecht, 1959). What is often portrayed as universal is in fact reflective of one specific socio-cultural context which excludes many others. Within the Indian philosophical tradition itself, different strands of views pertaining to the relationship between ethics and politics as well as the role of morality in politics have been articulated across centuries.

2. Kautilya And The Contingency Of Morality

Kautilya’s Arthashastra is considered a classic in the illustrious Arthashastra tradition of Indian political philosophy. Kautilya was the key adviser to and the chief architect of the strategy undertaken by the Indian king Chandragupta Maurya (c. 317–293 B.C.E.). Chandragupta Maurya defeated the tyrannical Nanda kingdom and also stopped the advance of Alexander the Great’s successors, thereby uniting most of the Indian subcontinent in an empire for the first time. Kautilya - sometimes called chancellor or prime minister to Chandragupta -, composed his Arthashastra to show a wise king how to defeat his enemies and rule on behalf of the general good. R. P. Kangle translates the word Arthashastra as “science of politics,” a treatise to help a king in “the acquisition and protection of the earth.” Others translate Arthasastra in slightly different ways: A. L. Basham says it is a “treatise on polity,” Kosambi emphasizes the economic importance of the word in calling it a “science of material gain,” and
G. P. Singh labels it a “science of polity.” Kautilya claimed to be putting forth what Heinrich Zimmer rightly calls “timeless laws of politics, economy, diplomacy, and war” (Zimmer, 1967).

According to Kautilya, the state, which he sees as being embodied in the position of the King, has both a negative and positive function. The negative and positive functions of the state have generally been classified within Hindu philosophy as Danda and Dharma. Danda incorporates the punitive measures adopted and meted out by the King for the orderly maintenance and preservation of worldly life in an otherwise violent state of nature characterised by the concept of Matsya-nyaya (law of the fish where in a situation of anarchy and lawlessness, the bigger fish prey upon the smaller ones). The positive function of the state is to ensure the fullest welfare of the people by upholding the principle of Dharma. Dharma here is seen to mean righteous action and performance of ethical duties. Interestingly, in the 'Shantiparva', the relativistic nature of dharma is clearly brought out wherein it is stated that when it is dangerous to speak the truth, untruth should be spoken. One should first understand circumstances under which he has to act- he should act unrighteously if it helps the course of dharma. Therefore, to protect Dharma, we should behave unrighteously with unjust people and righteously with just people. “Ethics had not assumed in Hindu thought the theoretically independent character that it did at the hands of western writers” (Verma, 1959). In his account, Kautilya shows not the least awareness of an existing dichotomy between politics and ethics, and in fact, takes the whole gamut of life within the political system into his account. In reading his Arthashastra, we find no moral considerations other than a king doing what is right for his own people. For instance, Kautilya’s ideal King would have no moral qualms about breaking obligations or trust when it came to alliances in foreign policy: “That ally who might do harm or who, though capable, would not help in times of trouble, he should exterminate him, when trustingly, he comes within his reach.” Kautilya also sought to take a nation trying to remain neutral or “indifferent” and secretly provoke war between that nation and a neighbouring kingdom, until the neutral nation sought his help. Then Kautilya’s king could “place him under (his) obligations.” Kautilya’s ideal King in pursuit of preserving and consolidating the welfare of his own state and people, would take steps such as- commanding secret agents to declare that the Regent of the king is about to
take power, while the agents kill leaders at night and blame the murders on the Regent of the enemy king; use secret agents in the countryside to sow the seeds of discontent and instigate protests against the oppression of the enemy king’s bureaucracy and kill agents of the king hoping to start a revolt. Even when talking about a King’s commitment to the welfare of his people, Kautilya appealed to justifications which have essentially pragmatic underpinnings. Kautilya urged the king to be popular with the people and rely on the countryside. “If weak in might, [a king] should endeavour to secure the welfare of his subjects. The countryside is the source of all undertakings; from them comes might.” The king’s power, for Kautilya, is in the end tied to the power and popular energy of the people, without which a king can be conquered, for “not being rooted among his subjects, [a king] becomes easy to uproot.” A righteous king should intervene in the kingdom of a tyrannical king who has weak roots in his populace. This intervention is not only to liberate oppressed masses but here morality is something advantageous and in one’s self-interest.

Kautilya made a distinction between morality of the ruler and the ruled, and put the ruler above the ordinary canons of morality. He avoided discriminating between good and evil, and writing from the standpoint of the ruling elite, he confined his interests almost entirely to the affairs of State and man’s political behaviour: that is, maintenance and enlargement of power. In this quest for power, Kautilya prescribes various tactics for the king to adopt, which, while ostensibly might seem ‘unethical’, when perceived from the larger end-goal, can be justified for its ethical political ends.

Kautilya stated that the king should be ready to use almost any means of violence in fighting a war; however, the violence should be directed towards the leaders of the opposing kingdom and not towards the ordinary people. In the pursuit of this power struggle, he discussed at length different means to be adopted by the king to attain victory against an enemy king, which included detailed accounts on how to employ poison, how to assassinate a king, how to arrange for spies and secret agents, how to “sow discord” and to foment and inflame “mutual enmity, hatred, and strife”, how to spread disinformation or propaganda, et al.
For example, on the topic of employing poison, Kautilya advised that while “giving unadulterated wine to the army chiefs, [the secret agent] should give them (wine) mixed with poison when they are in a state of intoxication.” Whereas Kautilya did suggest that an army laying siege to a fort try to “defile the water,” this measure seems designed to make those in the fort surrender from illness, not to kill everyone in the fort.

While addressing the question of how to assassinate a king Kautilya suggests hiding “inside the image of a deity or a hollow wall” and emerging at night, making something heavy fall on the king, or using women as secret agents to “drop on him serpents or poisonous fire and smoke.” In short, Kautilya prescribed the use of any possible means to assassinate an enemy king—drown him, burn him with fire, suffocate him with smoke, or even use crocodiles as assassins, not to mention employing women and children as poison-givers. By prescribing such tactics, Kautilya’s Arthashastra contrasted the prescriptions laid down in the Dharmasutras where the authors directed that a king should not “strike with barbed or poisoned weapons”.

Aside from assassination, another method used to defeat an enemy without full-scale battle was to arrange for the enemy to quarrel and fight among itself. Kautilya believed that the promise of power has the potential to ignite quarrels among high officers or officials. Therefore, as a means to acquire greater power he discusses the need and way to employ spies and secret agents. For example, an elaborate account is provided in one of the passages of the Arthashastra, in which Kautilya details on how one should arrange for a secret agent, disguised as an astrologer, to tell a high officer that he has all the marks of a king, and similarly arrange for a female secret agent, the wife of this officer, to complain that the king wants to keep her in his harem. A third secret agent who is a cook or a waiter should lie, saying that the king has ordered him or her to poison the high officer. “Thus with one or two or three means,” according to Kautilya, the king “should incite the high officers one by one to fight or desert” the enemy king.

---

1 The Dharmasutras is a genre of Sanskrit texts, and refers to the treatises of Hinduism on Dharma, dating back to 1st millennium BCE. The texts include discussion of ashrama (stages of life), varna (social classes), purushartha (proper goals of life), personal virtues and duties such as ahimsa (non-violence) against all living beings, rules of just war, and other topics.
In a discussion about sowing dissensions among oligarchies, Kautilya suggested that “assassins should start quarrels by injuring objects, cattle or men at night,” “should stir up princelings enjoying low comforts with (a longing for) superior comforts,” and “should start quarrels among the followers of the chiefs in the oligarchy by praising the opponents in brothels and taverns.” The goals were constantly to “sow discord” and to foment and inflame “mutual hatred, enmity and strife.” In the same manner, he praised the use of disinformation, propaganda and lies to gain political ends. Disinformation also made the use of religion to strengthen its appeal.

Such advice as stated by Kautilya in clear terms violated the tacit code of war found in the great Indian epics. The assassination of envoys and the use of poison were considered to be against the rules of warfare and thus not honourable. *The Laws of Manu* says that “Fighting in battle, [the king] should not kill his enemies with weapons that are concealed, barbed, or smeared with poison or whose points blaze with fire.” Spies were common in Indian history, but not spies who assassinated enemy officials and started quarrels among enemy leaders. Kautilya overlooked such earlier accounts while formulating his own code of conducting the Arthashastra. He judged the means by the result, and the result he sought was the general good of his kingdom and the welfare of its people. The Arthashastra, in this sense, can be said to be one of the earliest texts to have put into question the unmarked distinction between ethics and politics.

For Kautilya, the end “good” justifies the usage of “bad”, if not “evil”, means. He, therefore, is highly cautious of the activities to be adopted by a king once he emerges victorious. A conquering king should reassure a defeated people that not much, except their rulers, has changed. The king who has triumphed “should adopt a similar character, dress, language and behaviour (as the subjects). And he should show the same devotion in festivals in honour of deities of the country, festive gatherings and sportive amusements.” The conquering king should show his goodwill toward the defeated by instituting “a righteous custom, not initiated before.” While the victorious king is reassuring the general population with generous policies, he must continue to kill anyone who is dangerous and those who are disgruntled: “He should put down by silent punishment those capable of injuring [him] or those brooding on the master’s
destruction.” The importance attached by Kautilya towards keeping the population satisfied, to the extent of appeasement of the vanquished people, and swiftly rooting out the disgruntled elements of the society is a political goal in itself through which the king can seek to further his power. Thus, Kautilya puts forth a unique combination of both the positive and negative aspects of political practice, that is, Dharma and Danda respectively, for a just and powerful rule by the king over his people.

3. Conclusion

In this paper, we do not claim that the solution Kautilya offered in navigating through the conundrum between morality and politics is the only correct or right approach. What we have intended to highlight is that Kautilya started the conversation by acknowledging the tensions that might crop up while trying to locate the role of morality (if there was any) within politics. His position was a significant break from the earlier traditions of Indian philosophy which put forward predominantly moralistic and normative injunctions. Kautilya was writing at a time of great political churning - the existing ruling order was crumbling from within while conquerors were hounding at the doorsteps. In such a situation, Kautilya responded to these challenges by not just discarding morality as irrelevant to practical political needs, but reconfigured morality to strengthen practical politics itself.

Morality in politics is often seen as a liability - something which hinders effective results in the practical world of real life politics. What such an assumption fails to take into account is that many of the issues that politics faces and has to respond to, are inherently moral in nature. To try and evade them by putting forward so-called “objective” solutions does not do justice to the gravity of those problems. The goal should not be to put morality and politics within separate air-tight containers to be brought out together as and when convenient. There is rather, a need to recognize that morality and politics coexist, even if we close our eyes to the fact. This is not to advocate a homogenous, static idea of morality to be imposed by the majority on the others. What this paper has sought to demonstrate is that what is morality and what is politics themselves need to be

---

2 Kautilya, Artha's-asta, 13.5.17: 492.
constantly examined, keeping in mind contemporary contingencies and differential contexts while we talk about the role of morality in politics.

References


Abstract. In his most famous work, Machiavelli showed how the prince is never assessed by what he is, but according to what he appears to be. Therefore, the politician needs to act as a hypocrite, as a ‘great pretender and dissembler’, to maintain the support of the people on which his power relies. However, hypocrisy does not simply mean deception of the ruled by the rulers. Given the social division which underlies the state, it actually plays a fundamental role in the imaginary constitution of a sovereign power capable of imposing unity and stability on collective life. Our aim is to question the use of hypocrisy in order to emphasize the distinction introduced by Machiavelli between moral virtue and political ‘virtù’. Furthermore, we will try to demonstrate how it is presented as part of a broader strategy against the unpredictable effects of time, symbolized by the metaphor of ‘Fortuna’.

Keywords: Machiavelli, prince, hypocrisy, virtue, fortune.

Resumo. Na sua obra mais conhecida, Maquiavel mostrou como o príncipe nunca é avaliado por aquilo que é, mas segundo o que parece ser. O político deve, portanto, agir como um hipócrita, como um ‘grande simulador e dissimulador’, para manter o apoio do povo sobre o qual o seu poder assenta. Contudo, a hipocrisia não equivale simplesmente ao engano dos governados pelos governantes. Face à divisão social subjacente ao estado, ela desempenha na verdade um papel fundamental na constituição imaginária de um poder soberano capaz de impor a unidade e a estabilidade sobre a vida coletiva. O nosso objectivo é indagar o uso da hipocrisia por forma a enfatizar a distinção introduzida por Maquiavel entre virtude moral e ‘virtù’ política. Além disso, tentaremos demonstrar de que maneira a hipocrisia é apresentada enquanto parte integrante de uma estratégia mais ampla contra os efeitos imprevisíveis do tempo, simbolizados através da metáfora da ‘Fortuna’.

Palavras-chave: Maquiavel, príncipe, hipocrisia, virtude, fortuna.

* Universidade da Beira Interior (UBI); PRAXIS – Centro de Filosofia, Política e Cultura; Bolseiro de Doutoramento FCT.
0. Introdução

O Príncipe foi considerado durante muito tempo um manual de conselhos práticos para governantes, com vista à captura e manutenção do poder. De acordo com o antimaquiavelismo tradicional — que assumiu diversas formas ao longo dos séculos, e continua, ainda hoje, a exercer uma poderosa influência —, um dos ensinamentos mais infames do secretário consistiria no apelo à “manipulação” da imagem pública do príncipe, sendo que o emprego do termo “manipulação” pressupõe desde logo que há um manipulador e um manipulado, alguém que engana e alguém que é enganado. Tentaremos mostrar, antes de mais, que a classificação desta obra como um simples “guia político”, composto por um conjunto de princípios universais, se baseia num mal-entendido. É o próprio Maquiavel quem assevera que “todas as ciências requerem prática, de modo a serem convenientemente dominadas” (Maquiavel, 2010, p. 350). Como veremos, isto não significa que as hipóteses teóricas devam ser verificadas através de um confronto empírico, mas que a teoria, pela sua abertura constitutiva à contingência, se articula necessariamente numa referência contínua à prática.

Já vários autores ressaltaram as dificuldades interpretativas criadas pelo nexo que a obra maquiaveliana institui entre teoria e prática. A este propósito, observa Merleau-Ponty o seguinte: “o que faz com que não compreendamos Maquiavel é o facto de ele unir o sentimento mais agudo da contingência ou do irracional no mundo com o apreço pela consciência ou pela liberdade no homem” (Merleau-Ponty, 1960, p. 215). Segue-se daqui a impossibilidade de reduzir a política a uma planificação ou a um cálculo que prevê todos os resultados possíveis, fixando de antemão os modos de proceder em cada situação. No Príncipe, o ponto de vista do teórico está inextricavelmente ligado com o ponto de vista do actor. Como elucida Althusser, “o texto implica-nos e envolve-nos praticamente, politicamente. Ele [Maquiavel] interpela-nos de um lugar que nos convoca a ocupar como ‘sujeitos’ (agentes) potenciais de uma prática política potencial” (Althusser, 1999, p. 32), pelo que o leitor não pode permanecer indiferente ao conteúdo das teseis aí enunciadas.

Além disso, pretendemos questionar a ideia, formulada no seio da tradição Nietzsche-Foucault, segundo a qual a verdade corresponde àquilo que foi imposto unilateralmente pelos vencedores (cf. Frosini, 2001, p. 10). Através da análise do
Príncipe, e em particular dos capítulos XV, XVIII e XXV, colocaremos em evidência que o político deve ser de facto alguém que representa — um verdadeiro “hipócrita”, como se designam em grego os actores —, sem esquecer que nem a mais subtil hipocrisia é capaz de suprimir a “natureza intrinsecamente incerta e contingente do político” (Aurélio, 2012, p. 93). Longe de ser a expressão de uma pura relação de forças previamente dada, a “verdade” revelar-se-á o produto de uma relação dinâmica e de interdependência em que nenhum dos termos (príncipe e povo) se sobrepõe totalmente ao outro.

1. A verdade efectiva da política

Maquiavel é amiúde elevado a fundador da ciência política e equiparado a figuras precursoras do positivismo como Galileu e Descartes, arautos de uma nova forma de racionalidade depurada dos preconceitos antigos. Com efeito, é fácil identificar no Príncipe inúmeras afinidades com o método cartesiano, como sejam a listagem exaustiva dos diferentes tipos de principado (que antecipa, em certa medida, a regra das enumerações completas) ou a necessidade de abandonar a “imaginação” para “ir atrás da verdade efectiva da coisa” (Maquiavel, 2008, p. 185) — exigência que parece partilhar da suspeita típica da ciência moderna em relação às “aparências” (cf. Althusser, 1999, p. 126). Apesar dos paralelismos que podemos encontrar, até que ponto é lícito ver a obra que notabilizou o secretário florentino como um tratado puramente teórico, organizado de acordo com as normas do positivismo nascente? Mais do que uma questão formal de estilo ou método, está aqui em causa algo que tem directamente a ver com a acção política.

Conquanto realizada no quadro de uma investigação original em torno da genealogia e uso do conceito de “governamentalidade”, a leitura foucauldiana do Príncipe repete muitos dos preconceitos da leitura convencional, chegando mesmo a acentuar o carácter positivista e técnico do tratado (daí o interesse em considerá-la aqui):

Para Maquiavel, o príncipe está numa relação de singularidade e de exterioridade, isto é, de transcendência em relação ao seu principado. O príncipe de Maquiavel recebe o seu principado por herança, aquisição ou conquista; seja como for, ele não faz parte do principado: é exterior a ele. O vínculo que o liga ao seu principado é um vínculo de violência, tradição ou ainda um vínculo que foi estabelecido pelo acomodamento de tratados e pela cumplicidade ou concordância de outros príncipes, pouco importa. Em todo o caso, é um vínculo puramente sintético: não há
associação fundamental, essencial, natural e jurídica entre o príncipe e o seu principado. (Foucault, 2004, p. 95)

Posto que a relação do príncipe com o seu principado (aqui entendido como o conjunto objectivo composto pelo território e os seus habitantes) é uma relação de “exterioridade” e “transcendência”, ela permanece sempre frágil e ameaçada. Assim, a “arte de governar” descrita por Maquiavel visa sobretudo fortalecer esse vínculo pessoal ao principado contra os ataques constantes dos inimigos externos e internos. Ora, o saber político não passa somente por identificar os perigos, apurando a sua origem, natureza e intensidade relativa; mas também pela «arte de manipular as relações de força», no sentido de legitimar o imperium do príncipe e manter a obediência dos súditos (cf. Foucault, 2005, p. 95). Em suma, na perspectiva de Foucault, o príncipe maquiaveliano “governa”, por assim dizer, à distância, relacionando-se com o estado (que inclui o povo por ele governado) de um modo estritamente material, como se de uma propriedade sua se tratasse.

Contra a tradição interpretativa dominante (em que Foucault se insere), para a qual o Príncipe não passa de um manual destinado a governantes, Gramsci declarou que a obra de Maquiavel constitui um “manifesto político” com um escopo histórico concreto: a unificação da Itália do Cinquecento sob a liderança de um indivíduo dotado de uma virtù excepcional. Na esteira de Gramsci, Althusser inclui o Príncipe no “mundo da literatura ideológica” e assinala que o texto “toma uma posição nesse mundo”, utiliza “todos os recursos da retórica” com o intuito de “ganhar partidários para a sua causa” (Althusser, 1999, p. 23). O secretário não pode, por isso, ser equiparado ao cientista moderno que investiga a natureza desinteressadamente. Além de elaborar a “teoria dos meios ao dispor do Príncipe para salvar Itália, [Maquiavel] trata o seu próprio texto (…)
simultaneamente, como um desses meios”, transformando a escrita num verdadeiro “acto político” (Althusser, 1999, p. 23).

O problema abordado no tratado maquiaveliano excede, por conseguinte, o campo da teoria e emerge na facticidade da prática política (cf. Althusser, 1999, p. 17). Pese embora o carácter aparentemente universal das máximas que pontuam as obras de Maquiavel, já não podemos verdadeiramente falar de uma “objectividade pura mítica das leis da história ou da política” (Althusser, 1999, p. 19). As regras ou conselhos propostos apenas são válidos no quadro da “conjuntura” particular que ocasionou o problema práctico-político sob exame (e cuja configuração pode ser instantaneamente modificada pela intervenção do sujeito, obrigando a uma adaptação, ou mesmo inversão, do modo de procedimento originalmente adequado). Eis, talvez, o motivo pelo qual o secretário é frequentemente acusado de incoerência, o que revela não tanto a falta de rigorismo do autor, mas a inépcia dos seus detractores.

Temos que avançar até ao capítulo XV do Príncipe para encontrar, finalmente, o esboço de um “método” (no sentido moderno do termo) ou critério de avaliação do político:

Resta agora ver quais devem ser os modos e governos de um príncipe, quer com os súbditos, quer com os amigos. E porque sei que muitos escreveram sobre isto, receio que, escrevendo também eu, seja tido por presunçoso, mormente por me afastar, ao debater esta matéria, das ordens dos outros. Mas, tendo a minha intenção sido escrever coisa que seja útil a quem a escute, pareceu-me mais conveniente andare dietro alla verità effettuale della cosa do que da sua imaginação. Muitos imaginaram repúblicas e principados que não foram jamais vistos nem se soube se existiram de verdade. Porque é tanta a distância de como se vive a como se deveria viver, que aquele que deixa o que se faz por aquilo que se deveria fazer, mais depressa conhece a sua ruína do que a sua preservação. (Maquiavel, 2008, p. 185, sublinhado nosso)

A que é que Maquiavel se refere quando fala da “verdade efectiva da coisa”? Rejeitar a “imaginação” e as concepções utópicas da política implica, simplesmente, substituir a ideia de bem pela de utilidade? Na realidade, o príncipe não se pode orientar segundo valores morais nem segundo as regras da técnica política (cf. Lefort, 1986, p. 403). O capítulo XV marca um ponto de viragem no opúsculo não só por consumar a ruptura do autor com os seus antecessores e com o género literário dos specula principum (caracterizado pelo confronto da acção do soberano com os preceitos universais da moral), mas sobretudo por encetar a análise da relação entre príncipe e súbditos que se vai
prolongar até ao fim da obra. Ora, a “verdade efectiva” dos governantes consiste precisamente na representação que deles fazem os súbditos (cf. Lefort, 1986, p. 404). Fora do espaço aberto da opinião pública, onde as suas acções são julgadas e adquirem um significado particular, as qualidades privadas do príncipe tornam-se politicamente irrelevantes. Ainda no capítulo XV podemos ler:

Deixando, pois, de lado as coisas imaginadas acerca de um príncipe, e discorrendo sobre aquelas que são verdadeiras, digo que todos os homens, quando se fala deles, mormente os príncipes, por estarem postos mais alto, são referidos por algumas destas qualidades que lhes acarretam censura ou louvor. E assim é que um é tido por liberal, outro por mesquinho (...); um é tido por dador, o outro por rapace (...). Eu sei que qualquer um admitirá que seria coisa muito louvável encontrar-se um príncipe com todas as supracitadas qualidades que são tidas por boas. Mas porque não se pode tê-las todas, (...) é necessário ele ser tão prudente que saiba fugir à má fama dos vícios que lhe tirariam o estado. (Maquiavel, 2008, p. 186, sublinhado nosso)

Enquanto os tratadistas escolásticos condenavam as acções que não decorressem da virtude pessoal e autêntica do príncipe, independentemente dos benefícios daí resultantes para os súbditos, agora a bondade do príncipe passa a ser indiferente, ou mesmo evitável, sempre que conduza à perda do Estado. Mas se o objectivo fosse excluir os valores da moral e introduzir, sic et simpliciter, uma arte de governar utilitarista, porque é que o autor dedicou quatro capítulos (XV-XVIII) ao exame crítico das compatibilidades e incompatibilidades entre as qualidades morais e políticas? Ao afirmar que o príncipe deve saber “fugir à má fama dos vícios que lhe tirariam o estado”, Maquiavel chama atenção para o facto de a prática política ser indissociável do ethos próprio de cada sociedade (cf. Lefort, 1986, pp. 403-404). Por outras palavras, o agente político tem que começar por respeitar as leis morais, costumes e opiniões que compõem, de acordo com Althusser, “o que pode ser anacronicamente chamado de ideologias” (Althusser, 1999, p. 97).

Consequentemente, o que define a “verdade efectiva” não é a sua factualidade, mas o seu carácter ideológico (cf. Frosini, 2006, p. 56). Ao contrário da “verdade das ciências naturais”, a “verdade efectiva” da política não remete para um dado primário que vale por si mesmo, mas apenas existe por intermédio da opinião de outrem. Não admira, pois, a importância dada pelo secretário à noção de “reputação” [reputazione], termo empregue sucessivamente nas suas obras. Todavia, essa “verdade” permanece ambígua, sendo impossível de fixar de um modo definitivo e unívoco, visto que a opinião comum resulta de um processo aberto onde se cruzam pontos de vista, discursos e práticas conflitantes que
estão, outrossim, em constante modificação. É em função deste contexto fluido — irreduzível a um sistema moral genérico — que o príncipe deve conduzir a sua ação e construir a sua imagem pública (cf. Frosini, 2006, p. 60; Visentin, 2006, pp. 235-236).

A principal novidade do conceito de “verità effettuale” reside na elisão da distância entre “ser” e “aparência”, uma vez que “a generalidade dos homens aceita do mesmo modo o que parece e o que é, ou melhor, muitas vezes se impressionam mais pelas coisas que parecem do que por aquelas que efectivamente o são” (Maquiavel, 2010, p. 94) — conforme se lê numa passagem de Discorsi, I, 25. Que os indivíduos sejam (para outrem), de facto, aquilo que parecem, constitui uma condição geral das relações humanas, condição essa que se torna particularmente evidente no âmbito da política, em que a actuação dos príncipes, “por estarem postos mais alto”, é continuamente observada e escrutinada. Ficar preso à “imaginação” da coisa, significa, então, esquecer que a verdade de cada um se encontra fora de si, no olhar do outro, e que a existência do príncipe coincide absolutamente com as acções que lhe dão visibilidade (cf. Visentin, 2006, p. 231).

Em resumo, a identificação entre “ser” e “aparecer” impede-nos de separar o príncipe, enquanto “objecto” da visão de outrem, do príncipe, “sujeito” de uma prática (cf. Frosini, 2006, p. 61). O político nunca determina inteiramente os efeitos da sua ação. São os súbditos que, vendo-a e sentindo-a, a qualificam positiva ou negativamente. Gera-se assim uma tensão entre a necessidade de responder às exigências da sociedade, de um lado, e os limites impostos pela contingência própria da política, do outro, que nos permite definir o saber da “verdade efectiva” já não como um conhecimento técnico ou arte de manipulação do povo, mas como a capacidade de distinguir aquilo que depende da iniciativa do príncipe daquilo que escapa por inteiro ao seu controlo. Dito de modo mais concreto, trata-se de discriminar entre as medidas tomadas e a forma como estas se repercutem no meio flutuante da opinião pública.

2. A variação da Fortuna e a versatilidade do Príncipe

Como se sabe, Maquiavel recorre à imagem da Fortuna para descrever os eventos imprevisíveis que se furtam a uma explicação racional, assemelhando-a a aqueles “rios ruinosos que, quando se irritam, alagam as planícies, derrubam as
árvores e os edifícios, levam terreno de um lado, põem-no no outro” (Maquiavel, 2008, pp. 231-232), num esforço por nos levar a reconhecer os limites da acção. O secretário recusa, porém, qualquer forma de fatalismo, posto que a “fortuna” só é “árbitra de metade das nossas acções” e nos deixa “a nós governar a outra metade” (Maquiavel, 2008, p. 231). A virtù joga-se nesta margem indeterminada, aberta à intervenção humana, e exprime o êxito do príncipe face aos desafios colocados pela história. Cabe agora averiguar, em primeiro lugar, a validade da interpretação canónica segundo a qual Maquiavel desenvolve uma teoria da virtù enquanto simples capacidade de adaptação contínua à mudança; e, em segundo, como é que a virtù maquiaveliana está intrinsecamente ligada à noção de “verdade efectiva”.

No capítulo XV, o florentino afiança que um príncipe deve, para se manter, “aprender a poder ser não bom e usá-lo e não usá-lo consoante a necessidade”. Note-se que o autor emprega a expressão “não bom” [non buono] para frisar a diferença relativamente à pura maldade do tirano. Como vimos, o príncipe não se pode guiar por um código moral previamente estabelecido (nem mesmo aquele que ele escolheu livremente seguir); contudo, a sua actuação continua a estar condicionada por aquilo que Maquiavel chama “necessidade”. Para ser domada, esta força irracional — também apelidada de “acaso” ou “fortuna” noutros passos da obra — requer, por vezes, acções extraordinárias, à revelia das leis e costumes populares. Atendendo ao facto de o príncipe ser um “indivíduo político” cuja existência, em contraste com o “indivíduo privado”, é completamente definida pela sua “função política”, Althusser clarifica:

O Príncipe pertence a um domínio diferente (...). Não é a satisfação das suas necessidades que o motiva, e tão pouco o apaziguamento das suas paixões. Ele está além das categorias morais de vício e virtude (...). Virtù não é o oposto de virtude moral: ela é de uma ordem diferente. Ela não exclui a virtude moral; a virtù está posicionada de tal forma em relação à virtude que ela a pode incluir, e, não obstante, excedê-la simultaneamente. Assim, a virtù pode tomar a forma da virtude moral. Mas tem que ser dito que o Príncipe é moralmente virtuoso através da virtù política, e Maquiavel gostaria que isso se verificasse tanto quanto possível. (Althusser, 1999, pp. 92-93)

O encontro entre a virtù e a virtude é meramente acidental. Se fosse aplicada rigorosamente no domínio político, a virtude moral teria consequências nefastas, “porque o tempo varre adiante de si todas as coisas e pode levar consigo bem como mal e mal como bem” (Maquiavel, 2008, p. 122). Na verdade, a virtù é mera “reflexão, tão consciente e responsável quanto possível, das condições objectivas.
para o cumprimento da tarefa histórica da hora” (Althusser, 1999, p. 94). Seria, portanto, um absurdo pretender fixá-la de acordo com um conjunto de características universais e imutáveis, pois ela deve acompanhar a própria variação das coisas:

Creio também que seja bem-sucedido aquele cujo modo de proceder condiz com a qualidade dos tempos e, de modo semelhante, que seja mal-sucedido aquele de cujo proceder os tempos discordam. (...) Daqui nasce aquilo que eu disse, que dois, atuando distintamente, surtem o mesmo efeito; e dois, atuando igualmente, um leva por diante o seu fim e o outro não. Disto depende também a modificação do bem; porque, se um governa com cautela e paciência e os tempos e as coisas giram de modo que o seu governo seja bom, ele vai tendo sucesso; mas, se os tempos e as coisas se mudam, arruína-se, porque ele não muda o modo de proceder. (Maquiavel, 2008, pp. 232-233)

O político incapaz de decifrar a “necessidade” das circunstâncias históricas onde a sua acção se inscreve está condenado ao fracasso. Para durar, ele deve modificar-se com a evolução do contexto, espelhar-se na multiplicidade e variedade infinitas da natureza (cf. Lucchese, 2004, p. 365). Isto implica, sempre que se imponha, tomar decisões impopulares. Como ressalta Merleau-Ponty,

pouco importa que o poder seja condenado num caso particular: ele estabelece-se no intervalo que separa a crítica do repúdio, a discussão do descrédito. As relações do súbdito ao poder, como aquelas do eu e de outrem, (...) sobrevivem à contestação, desde que não se trate da contestação radical do desprezo. (Merleau-Ponty, 1960, p. 211)


A maioria das vezes, supõe-se que a virtù provém do encontro [riscontro] da acção com os tempos. Assim, o virtuoso seria qualificado enquanto tal porquanto teria um “ânimo disposto a virar-se consoante os ventos da fortuna e a variação das coisas” (Maquiavel, 2008, p. 197), adaptando-se continuamente. No entanto, esta ideia de virtù enquanto mudança e adaptação permanente cai por terra no fim do capítulo XV:

Eu julgo realmente isto, que seja melhor ser impetuoso que cauteloso, porque a fortuna é mulher e é necessário, querendo-a ter debaixo, vergá-la e acometê-la. E vê-
se que ela se deixa vencer mais por estes que por aqueles que procedem friamente,
e, por isso, como mulher, é sempre amiga dos jovens, porque são menos cautelosos,
mais ferozes e mandam nela com mais audácia. (Maquiavel, 2008, p. 234)

A metáfora da fortuna enquanto “rio ruinoso”, ou seja, enquanto força
arrebadora à qual nos devemos inexoravelmente submeter, é então substituída
pela metáfora da fortuna enquanto “mulher”. Com o emprego desta figura, põe-
se em relevo que só triunfa do tempo quem se impõe a ele. Daí a preferência de
Maquiavel pelo *modus operandi* do “impetuoso” — aquele que, antecipando-se à
mudança, logra surpreendê-la, virando o desenrolar dos acontecimentos a favor
de si mesmo (cf. Aurélio, 2012, p. 89). Não se trata de eliminar a contingência da
política, mas de perceber que a fortuna apenas “demonstra a sua potência onde
não está ordenada virtude para lhe resistir” (Maquiavel, 2008, p. 232). Neste
contexto, o imperativo da versatilidade adquire um novo significado: mais do que
concordar com a fortuna, o príncipe tem de aprender a agir como ela, imitando o
seu comportamento imprevisível e as suas variações abruptas, por forma a ser ele
a definir as regras do jogo.

Cumpre salientar, enfim, que apesar de os mecanismos ideológicos
(costumes, leis, religião, etc.) cimentarem certos padrões comportamentais, “a
natureza dos povos é inconstante” (Maquiavel, 2008, p. 136), pelo que a flutuação
dos ânimos está incluída na variação geral das coisas. Se os hábitos e valores
comuns mudam de geração para geração, as paixões e desejos dos homens
mudam quase diariamente, o que torna impossível ao príncipe agradar sempre
aos seus súditos. Por esta razão, como nota D. P. Aurélio,

> A virtude política (...) consiste não só em ostentar virtudes, verdadeiras ou falsas,
> mas acima de tudo em condicionar os modos de ser avaliado, em ser ator e ao mesmo
tempo encenador. (...) Dominar a cena é também dominar o tempo, determinar
aquilo em que os súditos pensam, chamando a si próprio todo o poder de inovar,
ou seja, submetendo a mudança à sua vontade, reduzindo o mais possível o fator
surpresa e, em última instância, reconhecendo que o que se chama “determinismo
da história” não é senão síntema da sua impotência. (Aurélio, 2012, pp. 96-97)

### 3. A Teoria da Fraude

No capítulo XVIII do *Príncipe*, Maquiavel giza a chamada *teoria da fraude*
(ou astúcia), segundo a qual o político precisa de “ser grande simulador e
dissimulador” para conservar a sua posição. É com base nas ideias aí expostas
que se denuncia habitualmente o secretário como fundador da técnica moderna
de manipulação das massas (que visa mascarar a verdadeira natureza das
relações de poder). Todavia, em que medida podemos reduzir o *Príncipe* a um manual de ilusionismo político? Modificar as formas de agir e a imagem pública ao longo do tempo implica, necessariamente, ocultar uma verdade objectiva sob um manto opaco de aparências? Somente através da noção de *verità effettuale* é possível, no nosso entender, descortinar o alcance das recomendações feitas neste capítulo marcante e que tanta tinta fez correr.

Antes de indagar se é ou não aconselhável respeitar a palavra dada (o que equivale a escolher entre a *honestidade* e a *astúcia*), o autor já havia cotejado as qualidades da *liberalidade* e da *parcimónia* (cap. XVI), bem como da *bondade* e da *crueldade* (cap. XVII), seguindo em ambos os casos um raciocínio semelhante, que se pode sintetizar do seguinte modo: a aparência de liberalidade dá lugar à verdade da rapacidade quando descobrirmos que, para dar, é preciso tirar; que, para continuar a ser generoso, é preciso — por exemplo — cobrar duros impostos. Por outro lado, a aparência de bondade redunda na verdade da crueldade quando descobrirmos que, em consequência de não aplicar um castigo necessário no momento adequado, consentimos que a desordem e o caos adquiram proporções incomportáveis. No fim do exame de cada uma das “virtudes”, Maquiavel inverte o juízo inicial, demonstrando como o príncipe que não cede à tentação imediata da generosidade, acaba por se poder mostrar liberal para com o seu povo; enquanto aquele que sabe “ser não bom” quando as circunstâncias o exigem, acaba por guardar os seus súbditos de males maiores no futuro (cf. Lefort, 1960, p. 406).

Se o secretário nos leva a aceitar esta argumentação, é apenas para a subverter logo em seguida (como acontece amiúde na sua obra). Ao desmontar o jogo retórico aqui subjacente, C. Lefort sublinha que “o preconceito nos engana; Maquiavel não pretende regressar do parecer ao ser; ele interroga o parecer na certeza de que o príncipe não existe senão para os outros, que o seu ser está no *exterior*. A sua crítica desenrola-se apenas na ordem das aparências” (Lefort, 1960, p. 408). Pese embora o título que lhe foi atribuído — “De que modo deve ser mantida pelos príncipes a palavra dada” —, o objectivo do capítulo XVIII não é, como nos capítulos precedentes, desvelar a natureza imoral da fidelidade à palavra dada, para restabelecer a verdadeira honestidade, mas ligar imediatamente o vício da astúcia — que, neste caso, se traduz no incumprimento
dos pactos estabelecidos — à imagem da virtude (cf. Lefort, 1960, pp. 412-413). Politicamente, a única qualidade que conta é a astúcia, ou melhor, a habilidade para exibir qualquer qualidade conforme a necessidade:

A um príncipe não é, pois, necessário ter de facto todas as supracitadas qualidades, mas e realmente necessário parecer tê-las; atrever-me-ei, aliás, a dizer isto: tendo-as e observando-as sempre, elas são danosas, e, parecendo tê-las, são úteis, como parecer piedoso, fiel, humano, íntegro, religioso, e sé-lo, mas ter o ânimo edificado de modo a que, precisando de não ser, tu possas e saibas tornar-te o contrário. (...) E por isso é preciso que ele tenha um ânimo disposto a virar-se consoante os ventos da fortuna e a variação das coisas lhe mandam. (Maquiavel, 2008, p. 197)

Lembremos que a astúcia não dimana da perversidade do príncipe, e que, além disso, o modo de proceder astucioso não é incompatível com a bondade. Porém, o encontro entre ambos resulta de uma coincidência feliz entre o ser e o parecer que nunca apaga a sua diferença. Cabe também esclarecer que a astúcia política se distingue da astúcia vulgar, dado que não visa um fim específico (o interesse pessoal do príncipe) nem constitui um meio que podemos usar ou não de acordo com as circunstâncias: a astúcia, seguindo de novo a fórmula de Lefort, apresenta-se como a “a arte de ligar cada acção particular e cada imagem que ela suscita a uma boa imagem do príncipe” (Lefort, 1960, p. 413). Efectivamente, é o próprio Maquiavel quem declara: “Deve, pois, um príncipe ter grande cuidado não lhe saia jamais da boca uma coisa que não esteja repleta das supracitadas cinco qualidades; e que, ao ouvi-lo e vê-lo, ele pareça todo piedade, todo fidelidade, todo integridade, todo humanidade, todo religião” (Maquiavel, 2008, p. 197).

Logo no começo do capítulo XVIII, o secretário afasta-se do tema sugerido pelo título e entra imediatamente no tratamento da questão da duplicidade, afirmando que “há dois géneros de combate: um com as leis, outro com a força”, sendo que “o primeiro é próprio do homem, o segundo das bestas”. Diferentemente, portanto, do que acontecia na tradição medieval, o texto procura definir não tanto as qualidades imutáveis do príncipe, mas os modos do seu combater, ou seja, os modos do seu agir, para concluir que precisa “um príncipe saber usar uma e outra natureza” (Maquiavel, 2008, p. 196). Contudo, Maquiavel introduz em seguida uma nova distinção na esfera do animal, servindo-se da imagem da raposa e do leão para ilustrar formas de actuação
diversas: enquanto o leão simboliza o uso da força pura; a raposa simboliza o uso da astúcia e da dissimulação (sem o qual a primeira permanece incompleta).

Ao glosar ironicamente Cícero, para quem a força e a astúcia do leão e da raposa constituem as duas fontes de injustiça\(^1\), o autor do *Príncipe* recusa audaciosamente a distinção clássica entre razão e paixão, homem e besta, argumentando que “uma sem a outra não dura” (Maquiavel, 2008, p. 196). O reconhecimento da fraude como um elemento fundamental da prática de governo deve ser visto, antes de mais, à luz da concepção maquiaveliana da política como uma arena de combate e disputa incessante. Neste primeiro sentido, a astúcia é sobretudo a arte de escapar às armadilhas dos adversários — duplicação que permite assumir o ponto de vista de outrem, conhecer as suas intenções, forças e planos secretos (cf. Lefort, 1960, p. 411). Mas a astúcia corresponde ainda àquela versatilidade sem a qual o príncipe não pode satisfazer os desejos e exigências variáveis dos súbditos. Althusser releva que

nele o animal se desdobra, faz-se à vez leão e raposa; e, no fim, é a raposa que governa tudo. Pois é a raposa que lhe imperará seja parecer mau, seja parecer bom, em suma, fabricar-se ou não uma imagem popular (ideológica) de si que responda aos seus interesses e aos interesses dos “pequenos”. (Althusser, 1994, pp. 545-546)

A referência ao centauro Quíron como mestre *par excellence* na fina arte da duplicidade prende-se justamente com a ambiguidade da sua constituição: este ser híbrido, misto de humano e cavalo, conjuga na sua composição anatômica o elemento humano e o elemento equino. No entanto, a figura mitológica do centauro representa também a articulação dupla (mais significativa) do “homem” e do “animal”, em que a natureza bestial, selvagem e bárbara se confunde com a do herói civilizador, pedagogo, iniciador nos campos mais variados como a caça, a música, a medicina, etc. (cf. Derrida, 2009, p. 117). De um modo geral, podemos dizer que o centauro Quíron personifica a multiplicidade individual, e é nele que o príncipe se deve inspirar para fazer face às contínuas alterações da fortuna, num esforço para superar as barreiras e limites que lhe são naturalmente impostos, multiplicando os pontos de vista e experiências que suportam a sua acção (cf. Lucchese, 2006, pp. 364-365).

---

\(^1\) Ver Cícero, 1913, pp. 45-47.
Chegados a este ponto, importa questionar: o apelo de Maquiavel ao uso da duplicidade faz do príncipe um impostor (como pretendem os seus críticos)? Há pelo menos dois momentos no cap. XVIII que atestam, *prima facie*, esta ideia: o primeiro, quando o secretário assegura: “são tão simples os homens, e obedecem tanto às necessidades presentes, que aquele que engana encontrará sempre quem se deixe enganar” (Maquiavel, 2008, p. 196); o segundo, já numa fase mais adiantada do capítulo, onde se lê: “Os homens em geral julgam mais pelos olhos que pelas mãos, porque o ver toca a todos, sentir toca a poucos: todos vêm aquilo que tu pareces, poucos sentem aquilo que tu és” (Maquiavel, 2008, p. 197). O secretário parece estar então a incitar o príncipe a ludibriar os súbditos, antes de mais porque o pode fazer — uma vez que eles têm uma inclinação natural para se deixar enganar —, e, além disso, porque a política se desenrola exclusivamente ao nível das aparências, impedindo que o povo conheça a verdade objectiva — à qual apenas alguns têm acesso.

Ora, que os homens julguem mais pelos olhos que pelas mãos, “não quer dizer”, como elucida Merleau-Ponty, “que seja necessário ou mesmo preferível enganar, mas que, na distância e grau de generalidade em que se estabelecem as relações políticas, desenha-se uma personagem lendária, feita de certos gestos e palavras, e que os homens honram ou detestam cegamente” (Merleau-Ponty, 1960, p. 214). Se “todos vêm aquilo que tu pareces” e “poucos sentem aquilo que tu és”, tal acontece porque as condições da comunicação entre príncipe e súbditos (aquilo que chamaríamos hoje comunicação política) se distinguem inteiramente das condições da comunicação privada. No espaço aberto, público, e dinâmico onde as suas acções se revestem de significado, “as qualidades do chefe [mesmo que verdadeiras] são sempre presas da lenda” (Merleau-Ponty, 1960, p. 215).

Ainda que haja, de facto, uma tendência natural no povo para se deixar mistificar, esta mistificação é sempre apenas parcial, pois “nas acções de todos os homens, e mormente dos príncipes, em que não há um tribunal para onde reclamar, olha-se é ao resultado” (Maquiavel, 2008, p. 198). No âmbito da política, como já foi observado, a única verdade que conta é a efectiva, e esta resume-se à forma como os efeitos das acções individuais do príncipe figuram na imaginação colectiva dos súbditos. Em conclusão, o verdadeiro engano é admitir que um sujeito pode manipular completamente a realidade política e o modo
como é percepcionada por via de meios coercivos ou instrumentos ideológicos. Contra este delírio, devemos lembrar que a verdade da política apenas se define no domínio da opinião pública — o qual, decerto, o príncipe deve influenciar tanto quanto possa a seu favor, sem esquecer, todavia, que persiste sempre uma margem irredutível de contingência fora do seu controlo.

**Bibliografia**


CAN HYPOCRISY BE A VIRTUE? HUME ON THE MORALITY OF PRINCES

HUME E A VIRTUDE DA HIPOCRISIA

Alexandra Abranches*
alexandra@ilch.uminho.pt

Abstract. This paper examines the moral status of hypocrisy in the moral and political philosophy of David Hume. Its aim will be to try to determine whether, according to Hume, hypocrisy has any positive moral value, or whether, not having any, Hume should therefore be placed in the same category of political realists such as Machiavelli, with his sharp distinction between moral and political values. If the latter is the case, then hypocrisy can be described as an absolute moral vice. But if the former is the case, that is, if hypocrisy has any moral value, then Hume does not support the sharp separation between what is right from a political and from a moral point of view, which means that there may even be, then, some relation between hypocrisy and moral obligation. In other words, hypocrisy may very well be virtuous.

Keywords: Hume, Machiavelli, benevolence, justice, History, human nature, virtue, vice.

Sumário. Examina-se aqui a filosofia moral e política de David Hume para determinar se a hipocrisia tem algum valor moral ou se, pelo contrário, não tem qualquer valor moral. Se este último for o caso, então é plausível colocar Hume na mesma categoria de realistas políticos a que pertencem Maquiavel, Hobbes ou Mandeville. Dada a distinção clara entre valores morais e valores políticos, a hipocrisia seria um vício moral absoluto, fosse qual fosse o seu valor político. Mas se, pelo contrário, se puder mostrar que a hipocrisia tem algum valor moral, então, não fazendo Hume a separação estrita que o realismo político exige entre o moral e o político, poderá existir uma relação entre hipocrisia e obrigação moral e, consequentemente, a hipocrisia será um vício apenas relativo ou, dito de outro modo, poderá ser virtuosa.

Palavras-chave: Hume, Maquiavel, benevolência, justiça, História, natureza humana, virtude, vício.

* Philosophy Department University of Minho and CEPS.
"(...) the world is still too young to fix many general truths in politics, which will remain true to the latest posterity. We have not as yet had experience of three thousand years; so that not only the art of reasoning is still imperfect in this science as in all others, but we even want sufficient materials upon which we can reason."

"Of Civil Liberty" (1741)

There is something about hypocrisy that makes it a particularly vicious character trait. The hypocrite is someone who deceives others about what he truly believes, desires and values, who presents a false image of himself and his intentions. This deception is, moreover, deliberate, which explains why we attach such strong moral blame to it and also the revulsion it provokes. It is seen as a substantial moral defect, since the hypocrite not only lies about himself, but he wants to have the advantages that morality furnishes without being himself moral. He pretends to be moral in order to manipulate the system of morality to his advantage. He is to blame both for deceiving others and for believing that it is good or permitted to deceive and manipulate others. His is, therefore, an absolute moral vice.

Still, there could be an argument to the effect that the viciousness of hypocrisy is not absolute. In fact, that argument has been made throughout the history of political philosophy. Venerable thinkers such as Thucydides, Machiavelli, Hobbes, or Mandeville, have defended the idea that hypocrisy is necessary in the public domain and that it is therefore justified. This political

---

1 References to Hume’s works include the following abbreviations: Hume’s *A Treatise of Human Nature* is cited with notations of the form T j.k.m.n / SBN pqr, the lower-case letters here standing for arabic numerals. Numerals immediately following T indicate book, part, section, and paragraph numbers in (Hume, *A Treatise of Human Nature*, 2005); numerals following ‘SBN’ indicate page numbers in (Hume, *A Treatise of Human Nature*, 1969). Hume’s *Enquiry concerning the Principles of Morals* and *Enquiry concerning Human Understanding* are cited similarly. EPM followed by two numbers gives the chapter and paragraph in (Hume, *An Enquiry concerning the Principles of Morals*, 1998). EHU followed by two numbers gives the chapter and paragraph in (Hume, *An Enquiry Concerning Human Understanding*, 2007). And in each case numbers paired with these, following ‘SBN’, indicate the corresponding page in (Hume, *Enquiries Concerning Human Understanding and Concerning the Principles of Morals*, 1975). References to Hume’s essays shall be given to individual essays in (Hume, *Essays: Moral, Political and Literary*, 1985) followed by page numbers. In all quotations, emphasis is original unless otherwise indicated.

2 As Judith Shklar writes in her book *Ordinary Vices* “hypocrisy remains the only unforgivable sin, perhaps especially among those who can overlook and explain almost every other vice, even cruelty” (Shklar, 1984, p. 45).
Can hypocrisy be a virtue? Hume on the morality of princes

realist defence of hypocrisy might, however, be interpreted as spelling out its manipulative and deceitful character, instead of actually being a defence and even less a justification. After all, it would still be about getting what you want through deceit, but under the guise of a public good of some sort. It would still be an abuse of morality in which morality would be treated as a manipulative means to an end the hypocrite might be hard pressed to unveil were he, the hypocrite, to be unveiled as a hypocrite. How can we, after all, trust the hypocrite and believe in his good intentions? The public good seems to require transparency. Hypocrisy is, by its very own nature, irredeemable, a vice even while presenting itself as a virtue precisely because it consists of presenting itself as a virtue.

The political realist would not be discouraged by considerations such as these, though. His is a pragmatic stance. There are times when it really is necessary to deceive in the interest of the public good. When those times arrive, one must not be squeamish. Necessity, and not justification, is what they affirm.

Our concern here will be not with the realist defence of hypocrisy, but with whether there is a case when hypocrisy might be considered a virtue. Hume’s thought about hypocrisy, and his moral and political philosophy, may give us some help to determine whether hypocrisy’s viciousness might be relative and not absolute.

The limits of self-interest

Hume devotes a lengthy appendix in his An Enquiry Concerning the Principles of Morals to a refutation of self-love and the hypocrisy that accompanies self-love. In that appendix, he writes that:

THERE is a principle ... that all benevolence is mere hypocrisy, friendship a cheat, public spirit a farce, fidelity a snare to procure trust and confidence; and that while all of us, at bottom, pursue only our private interest, we wear these fair disguises, in order to put others off their guard, and expose them the more to our wiles and machinations (EPM, App. 2.1; SBN 295) (My emphasis).

This principle, he writes, resembles another “which has been much insisted on by philosophers, and has been the foundation of many a system; that, whatever affection one may feel, or imagine he feels for others, no passion is, or can be disinterested; that the most generous friendship, however sincere, is a modification of self-love; and that, even unknown to ourselves, we seek only
our own gratification, while we appear the most deeply engaged in schemes for the liberty and happiness of mankind” (EPM App.2.2; SBN 296) (My emphasis).

These principles seem to make of us all, and always, hypocrites, no matter what the circumstances might be. In the first case, Hume is describing a principle according to which we are hypocrites in a moral sense, worthy of reproach and blame, to the extent that we wear “fair disguises”, which means it is our conscious intentions that are vicious, that we do on purpose wish to appear as something we are not so as to achieve some hidden end. We hide both our intention to deceive and the end we seek by deceiving. We manipulate others in order to benefit ourselves. We use them for our own self-interest.

The second principle is saying that we are not aware of what our real motivations are, of the causes of our actions, which means hypocrisy could be interpreted as neither a vice nor a virtue, but as a fact about our nature. Behind the experience of love or friendship lies self-love. Sincerity is unmasked by the philosopher as deceit. Nature is the hypocrite, not the individual persons whose feelings and actions it motivates. And Nature, as a system of facts, is morally indifferent.

The two principles are, then, quite different. For it is one thing to say that all men are selfish manipulators, and that they are so knowingly. It is another thing to say that all men are selfish manipulators, but they are so unknowingly. Still, both affirm self-interest as the ultimate motivation for human conduct.

One of Hume’s arguments against the “selfish system of morals” that these principles represent, a system defended by thinkers such as Hobbes or Locke, who, in these respect, are “superficial reasoners”, is that it is a bad theory of human behaviour. It is a reductionist theory, and to such an extent that it ends up being a simplistic theory. And, what is more, it is a simplistic theory which demands very complex turns of thought in order to reduce the phenomena to that

---

3 “An epicurean or a Hobbist readily allows, that there is such a thing as a friendship in the world, without hypocrisy or disguise; though he may attempt, by a philosophical chymistry, to resolve the elements of this passion, if I may so speak, into those of another, and explain every affection to be self-love, twisted and moulded, by a particular turn of imagination, into a variety of appearances” (EPM App.II.4; SBN 249-50).

4 “All attempts of this kind have hitherto proved fruitless, and seem to have proceeded entirely from that love of simplicity which has been the source of much false reasoning in philosophy” (EPM App. II.6; SBN 298).
Hume writes that “[T]he most obvious objection to the selfish hypothesis is, that, as it is contrary to common feeling and our most unprejudiced notions, there is required the highest stretch of philosophy to establish so extraordinary a paradox” (EPM, App. 2.6; SBN 298).

And then, there are the numerous examples in nature of general benevolence which make it simpler to explain human behaviour “from the combined motives of benevolence and self-enjoyments ...” (EPM, App.2.13; SBN 302).

Hume identifies benevolence as a second principle in human nature besides self-interest. This identification of benevolence as a second source of motivation for human behaviour and as a second source of feeling and experience provides an explanation for common phenomena which doesn’t call for an excess of theory. Facts make the selfish system too costly to uphold. Too much mental energy has to be spent in speculative explanations if one is to do it. And for someone who wants to do for the science of human nature what Newton did for the science of nature, facts should not be overrun by hypothesis, however simple and elegant in their simplicity they might seem. The fact that benevolence is also a motivation for human behaviour means that we are not all and always deceivers, be it in the sense described by the first of the two principles mentioned above, be it in the sense described by the second of those principles. We can be sincere in our benevolent actions, when we profess friendship, when we exhibit public spirit, when we make promises and affirm our fidelity. Nature does not compel us to pursue only our self-interest and to be, therefore, incapable of being disinterested. We must not all, and always, be hypocrites.

**Benevolence and justice**

Benevolence, then, is one of the main motives behind human behaviour. So is self-interest, of course, but the fact that benevolence is an independent factor means that not all behaviour can be reduced to self-interest. And this in turn means that not all benevolent acts are, of necessity, self-interest in disguise.

---

5 “When a philosopher, in the explication of his system, is obliged to have recourse to some very intricate and refined reflections, and to suppose them essential to the production of any passion or emotion, we have reason to be extremely on our guard against so fallacious an hypothesis” (EPM App. II; SBN 299).
Even though we are not egoists, in Hume’s speculative history of human association self-interest plays the major role. Benevolence, the natural virtue, is both too short in its reach and too weak to provide the proper resources and materials with which to build a civil society. The history of civil society is the history of self-interest pitted against itself, and from this struggle emerges justice, the artificial virtue which keeps humans under the authority of impartial and very often harsh rules. Benevolence, the natural virtue, is also a social virtue, but its scope hardly goes beyond family and friends, and its strength diminishes as its scope is enlarged. It shows that human existence has always been social and sociable, and it shows that there are sources of feeling and behaviour in human nature beyond self-interest, but it doesn’t guarantee that sociability will prevail over competition. What makes benevolence a social virtue is the fact that it contributes to general happiness by being useful and/or agreeable to oneself and/or to others. With each benevolent action, and with each action expressing some derivative form of benevolence, such as delicacy or generosity, general happiness is enhanced. It shows that humans are not lonely, independent individuals roaming the landscape and getting into scrapes with other individuals, but it doesn’t say human nature is all benevolence and no hostility.

Benevolence and justice, both social virtues, one natural the other artificial, have each a distinctive way of contributing to the common good. While benevolence is practiced in individual actions, justice is a system and application of general rules. This difference means that the two virtues can and often do come into conflict with one another. And when this happens, justice has the upper hand, which may contribute to its image of stern coldness, of indifference to personal suffering. Hunger, destitution, illness might well be considered, from a benevolent point of view, reasons to ignore or bend the rules that protect private property, for instance. But justice has not been born out of benevolence, its origin lies in the fight against the self-destructive effects of unbridled self-interest. From the point of view of justice, each action must be treated in the same way. Exceptions, even if justified by benevolence, put at risk the whole system of laws.
that keeps in check that other, and very powerful, human motivation, selfishness.⁶

**The general point of view**

In a footnote to the *Enquiry Concerning the Principles of Morals*, Hume writes:

> It is wisely ordained by nature, that private connexions should commonly prevail over universal views and considerations; otherwise our affections and actions would be dissipated and lost, for want of a proper limited object. Thus a small benefit done to ourselves, or our near friends, excites more lively sentiments of love and approbation than a great benefit done to a distant commonwealth: But still we know here, as in all the senses, to correct these inequalities by reflection, and retain a general standard of vice and virtue, founded chiefly on general usefulness (EPM, 5.42 n25; SBN 229 n1).

People commonly do not live up to what morality requires of them, *but* it is a good thing that they do not. The apology Hume offers on behalf of humankind is stated in the observation that nature ordained this *wisely*. Nature did not just ordain it, but it did so wisely. Human psychology explains this wisdom of nature: proximity and definition of the objects under consideration “excite more lively sentiments”, which otherwise would be “dissipated and lost”. Still, this doesn’t mean nature is against considerations of general usefulness. Nature is the basis upon which we have been building a general standard of vice and virtue. And even if we are not Ideal Spectators, we are capable of reflection⁷, and so of correcting and enlarging our point of view towards general usefulness. Benevolence’s narrow scope and short reach represent a moral benefit, since they give us an object that excites our sentiments, that is, they are motivationally powerful; but they also represent a moral deficit, since they are a new way of practicing egoism. It is justice, with its demand for a general and disinterested point of view that gradually corrects the “inequalities” of feeling we have for those who are near -

---

⁶ Hume has a very illuminating metaphor for the difference between the two social virtues: “The happiness and prosperity of mankind, arising from the social virtue of benevolence and its subdivisions, may be compared to a wall, built by many hands, which still rises by each stone that is heaped upon it, and receives increase proportional to the diligence and care of each workman. The same happiness, raised by the social virtue of justice and its subdivisions, may be compared to the building of a vault, where each individual stone would, of itself, fall to the ground; nor is the whole fabric supported but by the mutual assistance and combination of its corresponding parts” (EPM App. 3. 5; SBN 305).

⁷ “as in all the senses”: we also correct the external senses by reflection; they may deceive us, we resort to the testimony of others, to experience, to probability, in order to correct the information they give us, and to be “[T]he wise man [who] proportions his belief to the evidence.” EHU 10.1
and for the nearest one of all, our selves – and those who are far. By inspiring feelings such as civic virtue and honesty, of which the sensible knave described by Hume in the second Enquiry is incapable, justice ensures the possibility of political society.

The virtue of Justice brings objectivity into the human scene. To make a proper objective judgment founded on proper sentiments, to give reasons which other people can be expected to understand, we have to adopt an impersonal stance. Either we adopt the point of view of the Impartial Spectator, the point of view of humanity, the impersonal point of view, or we don’t. And if we don’t, we are no longer giving moral reasons for our judgments. As Hume writes:

When a man denominates another his enemy, his rival, his antagonist, his adversary, he is understood to speak the language of self-love, and to express sentiments, peculiar to himself, and arising from his particular circumstances and situation. But when he bestows on any man the epithets of vicious or odious or depraved, he then speaks another language, and expresses sentiments, in which he expects all his audience are to concur with him. He must here, therefore, depart from his private and particular situation, and must choose a point of view, common to him with others; he must move some universal principle of the human frame, and touch a string to which all mankind have an accord and symphony. If he mean, therefore, to express that this man possesses qualities, whose tendency is pernicious to society, he has chosen this common point of view, and has touched the principle of humanity, in which every man, in some degree, concurcs (EPM 9.10; SBN 272) (My emphasis).

Hume’s criticism of self-interest and his theory of justice as impartiality apparently steer him in the direction of the condemnation of hypocrisy as an absolute vice. But there is a topic in Hume’s writings where his less rigorous views about hypocrisy are manifested. In his opinion on the phenomenon of politeness, which he develops in his Essays, he presents hypocrisy as the condition of civilization. His defence of “manners” comes very close to a defence of hypocrisy when he writes that “among the ancients, there was not much delicacy of breeding, or that polite deference and respect, which civility obliges us either to express or counterfeit towards the persons with whom we converse” (My emphasis) (Hume, 1985a, pp. 111-137).

The ancients were too rough by being too straightforward. Hume is, in fact, one of the main figures of the eighteenth century “culture of politeness” that

---

8 An expression used by Lawrence Klein (Klein, 1994).
would be accused by the “culture of sincerity” of hiding privilege and injustice under the romantic cover of civility and chivalry.\textsuperscript{9}

**The morality of princes**

Another test on the limits of morality is politics, the actual behaviour of princes and governments. Machiavelli, a thinker who in *The Prince* represents political realism at its strongest, argues for a separation between morality and politics. This separation means that that which is called hypocrisy is simply a necessary act in order to preserve power or the common good. One might call it hypocrisy from the point of view of morality. But from the point of view of politics there is no sense in calling it that because the political realm is not the place to make moral judgments. Politics and morality have different values. You choose one or the other and judgments made from one and the other do not necessarily coincide. In Chapter 15 of *The Prince*, Machiavelli writes that “... it is necessary to a prince, if he wants to maintain himself, to learn to be able not to be good, and to use this and not use it according to necessity.”\textsuperscript{10}

Necessity is dictated by the way the world is, including human nature as part of that world. Acting according to virtue in the world is ruinous, given that most men are not virtuous\textsuperscript{11}. Morality, in fact, has not gained root in this world, except as part of perception: people “consider” virtues to be good and vices to be bad,

\textsuperscript{9} In Mary Wollstonecraft’s *A Vindication of the Rights of Men*, for instance, this attack reaches considerable proportions, with Burke, another representative of this defense of custom, being lambasted for his deference to power, and denounced as a self-serving hypocrite. Wollstonecraft writes that behind Burke’s professions of concern for liberty lie his “tyrannical principles”, that his cry for the birthrights of Englishmen does not refer to that birthright which all men have but only to those that some men happen to have. “Security of property”, writes Wollstonecraft, “is the definition of English Liberty”. Hume’s conception of justice includes protection of property at its core, so this definition suits it very well. And the revolution Wollstonecraft is defending in her attack on Burke is also a form of political action Hume would find both foolish and destructive. In his ‘Idea of a Perfect Commonwealth’ (Hume, 1985b, pp. 512-529) Hume rejects utopian plans that would destroy the present order and make a fresh start. Society, he believes, is a complex mechanism that has been and is being built throughout human history, and its screws and planks and springs have been tested by experience. Sometimes they get rusty or damaged, but they can be fixed, replaced even. Through gradual change, reform and progress are made possible. Still, this anti-revolutionary point of view is not necessarily a form of hypocrisy. It can be, when it resists those changes that would bring down unjustified privileges or abhorrent forms of oppression under the guise of something else, as Wollstonecraft’s criticism of Burke says he is doing. But being an advocate of reform is compatible with perfect sincerity, as the example of John Stuart Mill should make clear.

\textsuperscript{10} (Machiavelli, 1998)

\textsuperscript{11} “for it is so far from how one lives to how one should live that he who lets go of what is done for what should be done learns his ruin rather than his preservation. For a man who wants to make a profession of good in all regards must come to ruin among so many who are not good” (Machiavelli, 1998, Ch. 15).
and they consider some actions to be vicious and others to be virtuous; and the prince has to be “so prudent as to know how to avoid the infamy of those vices that would take his state from him and to be on guard against those that do not, if that is possible; but if one cannot, one can let them go on with less hesitation”12 (My emphasis).

Virtue and vice may look like virtue and vice yet reveal themselves to be the exact opposite if we measure them against the standard of politics, which is success, success in bringing “security and prosperity” to the state. And this means that there will never “be wanting to a prince legitimate reasons to excuse this nonobservance”13 of treaties or promises made to other princes.

Hume seems to share some ideas with Machiavelli that would make it plausible to put them in the same category of political realists: the rejection of utopian views in Machiavelli’s The Prince and in Hume’s “The Idea of a Perfect Commonwealth’ points in that direction; and so does the fact that both dedicate themselves to the study of History instead of designing an abstract theory of how the state should be.

But Hume’s stance on the relation between morality and politics is also based on a complex theory of human nature, and on the speculative, but still naturalistic, history that gradually evolves society out of the universal human passions of pleasure and pain. Hume claims he is not speculating but doing empirical science. And certainly he does not pretend to be doing normative theory. But Hume also has a substantive theory of human nature from which he has extracted a rule that, insofar as it contributes to the general good, should be adopted as a standard of behaviour and judgment. That rule states that “every quality which is useful or agreeable to ourselves or others” should be called a virtue, and the contrary a vice. Usefulness and agreeableness could be accompanied with a metric, but Hume doesn’t say how we might be able to

12 “And further more one should not care about incurring the fame of those vices without which it is difficult to save ones’ state; for if one considers everything well, one will find to be virtue, which if pursued would be ones’ ruin, and something else appears to be vice, which if pursued results in one’s security and well-being” (Machiavelli, 1998, Ch. 15).

13 “And if all men were good, this teaching would not be good; but because they are wicked and do not observe faith with you, you also do not have to observe it with them. Nor does a prince ever lack legitimate causes to color his failure to observe faith” (Machiavelli, 1998, Ch. 18).
measure them, just as he is vague about what utility consists in. But what is important in this rule is that it is not a set list of do’s and don’ts, a catalogue of virtues and vices. It is, rather, a rule that determines why qualities are virtues or vices and also when they are virtues or vices. And when properly applied, this rule may demand significant changes in our catalogue of virtues and vices:

Celibacy, fasting, penance, mortification, self-denial, humility, silence, solitude, and the whole train of monkish virtues; for what reason are they everywhere rejected by men of sense, but because they serve to no manner of purpose; neither advance a man’s fortune in the world, nor render him a more valuable member of society; neither qualify him for the entertainment of company, nor increase his power of self-enjoyment? We observe, on the contrary, that they cross all these desirable ends; stupify the understanding and harden the heart, obscure the fancy and sour the temper. We justly, therefore, transfer them to the opposite column, and place them in the catalogue of vices; nor has any superstition force sufficient among men of the world, to pervert entirely these natural sentiments. A gloomy, hair-brained enthusiast, after his death, may have a place in the calendar; but will scarcely ever be admitted, when alive, into intimacy and society, except by those who are as delirious and dismal as himself (EPM 9.3; SBN 270).

Hume is, then, very much on the side of morality and on the side of its reformation, and he is one of those who, like Rousseau and many others in the eighteenth century, defended Machiavelli from the charge of Machiavellism, of teaching ruthless amorality or even immorality. In the essay ‘Of the Study of History’, Machiavelli is even given as the example of the impartiality, that intellectual and moral virtue, of historians, who “... have been, almost without exception, the true friends of virtue, and have always represented it in its proper colours” (Hume, 1985c, pp. 563-68).

Hume considers that the study of History is part of the science of human nature inasmuch as it “extends our experience to all past ages, and to the most distant nations; making them contribute as much to our improvement in wisdom, as if they had actually lain under our observation” (Hume, 1985c, pp. 563-68). There is no contradiction or detachment in his work as a philosopher and as an historian. Through History, the past becomes part of the experience philosophy has to refer to, to learn about human nature.

---

14 See the example of chastity in T 3.2.12.

15 “MACHIavel himself discovers a true sentiment of virtue in his history of FLORENCE. When he talks as a Politician, in his general reasonings, he considers poisoning, assassination and perjury, as lawful arts of power; but when he speaks as an Historian, in his particular narrations, he shows so keen an indignation against vice, and so warm an approbation of virtue, in many passages, that I could not forbear applying to him that remark of HORACE. That if you chace away nature, tho’ with ever so great indignity, she will always return upon you” (Hume, 1985c, pp. 563-68).
But this doesn’t mean that the historian is subservient to the philosopher. And according to Hume, historians are better than philosophers in an important way. Unlike philosophers, historians look directly at agency and agents. A historian is “sufficiently interested in the characters and events, to have a lively sentiment of blame or praise”. She has a closer view of the moral dimensions of actual life being lived by actual individuals, and so she has a more lively sense of what it means to be virtuous or vicious. History is the scene in which vice and virtue unfold. And this is why the study of History is recommended by Hume as a source of moral learning for the public at large. History is not merely a collection of facts. Hume connects the study of particular historical events with his study of human nature, which includes the social virtues of benevolence and justice. And he attributes to History an important role for instruction on private and public virtue not only for philosophers but for the general reading public.

In the second Enquiry, he writes that “History, experience, reason sufficiently instruct us in this natural progress of human sentiments, and in the gradual enlargement of our regards to justice, in proportion as we become acquainted with the extensive utility of that virtue” (EPM 3.1; SBN 192). Utility, or usefulness, as we saw above, is what makes any virtue a virtue, and justice is no exception. In other words, where justice is not useful then it must be abandoned. Perfect benevolence or complete abundance, for instance, makes justice useless, so does the extreme need for self-preservation in civil war or disaster. In stable societies, respect for property, fidelity to promises, obedience to laws, are moral obligations. But they cease to be so in times of emergency or in the land of milk and honey.

But justice isn’t only a domestic virtue. The Law of Nations is the name of the new set of rules and duties which arise from the intercourse between different states16. As far as we are talking of law, and rules, and duties, we are talking about

---

16 “When civil government has been established over the greatest part of mankind, and different societies have been formed contiguous to each other, there arises a new set of duties among the neighbouring states, suitable to the nature of that commerce, which they carry on with each other” (T 3.11; SBN 376). “When a number of political societies are erected, and maintain a great intercourse together, a new set of rules are immediately discovered to be useful in that particular situation; and accordingly take place under the title of Laws of Nations. Of this kind are, the sacredness of the person of ambassadors, abstaining from poisoned arms, quarter in war, with others of that kind, which are plainly calculated for the advantage of states and kingdoms in their intercourse with each other” (EPM 4.2; SBN 205).
justice or morality, in this case, international justice. This intercourse between
nations is carried out because there is some advantage to the parties involved,
and in this sense, relations among states can be compared to the relations among
individuals that constitute each political society. The difference is that, Hume
writes in the Treatise, “though the intercourse of different states be
advantageous, and even sometimes necessary, yet it is nor so necessary nor
advantageous as that among individuals, without which it is utterly impossible
for human nature ever to subsist”17 (T 3. 2.11; SBN 377).

The utility of the intercourse among nations is not as strong as the utility of
the association among individuals. Or, as Hume puts in the second Enquiry, “... here is the difference between kingdoms and individuals. Human nature cannot
by any means subsist, without the association of individuals; and that association
never could have place, were no regard paid to the laws of equity and justice.
Disorder, confusion, the war of all against all, are the necessary consequences of
such a licentious conduct. But nations can subsist without intercourse. They may
even subsist, in some degree, under a general war” (EPM 4.3; SBN 206).

Because, then, regulated relations among states are less useful than
regulated relations among individuals, “the morality of princes ... has not the
same force as that of private persons, and may lawfully be transgressed from a
more trivial motive”18 (T 3.11; SBN 376) (My emphasis). Hume’s position is not
that morality and politics are separate realms, but that moral obligations in those
two domains have different strengths. Individual agents are bound to the virtues
and to justice in such a way that “nothing less than the most extreme necessity, it
is confessed, can justify individuals in a breach of promise, or an invasion of the
properties of others” (EPM 4; SBN 206). But “we must necessarily give a greater
indulgence to a prince or minister, who deceives another; than to a private
gentleman, who breaks his word of honour” (T 3.11; SBN 377). And we can
“reconcile us more easily to any transgression of justice among princes and
republics, than in the private commerce of one subject with another” (T 3.2.11;

17 “though the intercourse of different states be advantageous, and even sometimes necessary, yet it is nor so necessary nor
advantageous as that among individuals, without which it is utterly impossible for human nature ever to subsist” (T 3.11; SBN 377).

18 “though the intercourse of different states be advantageous, and even sometimes necessary, yet it is nor so necessary nor
advantageous as that among individuals, without which it is utterly impossible for human nature ever to subsist” (T 3.11; SBN 377).
Blame and praise are still in order as forms of assessment of the behaviour of princes, who are still objects of moral judgment and causes of moral feeling. But these aren’t as strong as those we reserve for our neighbours and fellow citizens.

Justice within the state and between states differs in terms of usefulness, but it is still justice, it is still a form of morality. Transgressions of justice among nations are still violations of moral obligations, but these are proportional to usefulness, so where usefulness is small or non-existent, moral obligation is small or non-existent. If there is some agreement whose usefulness is small, our moral obligation to abide by it is also small, which means that transgressing it is not seriously bad.

What is more, if, as Hume writes, “[T]he observance of justice, though useful among them, is not guarded by so strong a necessity as among individuals; and the moral obligation holds proportion with the usefulness” (EPM 4; SBN 206), then the prince is not morally obligated to keep his promises and being a hypocrite is not morally to blame. And if usefulness is the standard, it may even be the case that, in some individual instances, the prince is morally obligated to be a hypocrite, in case that proves useful to the state.

How, then do we turn a vice (hypocrisy in the private realm) into a sort of virtue (political hypocrisy)? If a virtue is that which satisfies the standard of morality, and if political hypocrisy satisfies the standard of morality, then political hypocrisy is a virtue.

Whereas political realists such as Machiavelli separate politics from morals and make moral judgments insignificant in the political realm, which means that there are no hypocritical princes, just effective or ineffective princes; Hume does not effect this separation, and offers instead a way of making hypocrisy compatible with morality, since “… the moral obligation holds proportion with the usefulness” (EPM 4.3; SBN 206).

So, even though his position seems to be the same as Machiavelli’s, as he expressly talks of reasons of state justifying the breach of justice, the fact is that as far as he has utility as a standard of virtue, and as far as he admits that there might be changes in our catalogue of virtues if only we reflect upon those acts and
qualities of mind which contribute to our common good, one may conclude that hypocrisy, which is an absolute vice in the realm of the relations among individuals, may be a relative virtue, and not merely a breach of justice, in the realm of the relation among nations. This insight might be made more plausible when paired with Hume’s defence of politeness as a pillar of civilization, but this theme will not be pursued here.

**Conclusion**

Hume’s treatment of the theme of the morality of princes and of the relations between the states seems to allow for the idea of a relatively virtuous hypocrisy. While hypocrisy, as a character trait of individuals, is an absolute vice, and individuals are morally obligated to keep faith, and may be legally sanctioned when they don’t, hypocrisy as the particular actions of princes in particular circumstances, even though they appear to be breaches of justice, even of that weaker form of justice which is the law of nations, may after all be considered a relative virtue if by virtue what we mean is that action or character trait which is useful. Hume admits that justice itself might cease to be a virtue when circumstances render it useless. Hypocrisy, then, may become a virtue if it becomes useful. When it comes to the relation among individuals, it is in fact a vice because it is in fact harmful to social life. But in the case of relations between states, if it is specifically useful, we may consider it a relative virtue, relative to circumstances and to the weakness of moral obligation in the international sphere.

There is another form of hypocrisy that might draw out the fact of political hypocrisy’s relative virtuousness or viciousness: religious hypocrisy. This, indeed, is a form of hypocrisy that deserves being called absolute, that one can say that has no redeeming value whatsoever, because the end it is a means to is false, and because it stirs up violent and destructive passions. As Hume says in a celebrated sentence: “Generally speaking, the errors in religion are dangerous; those in philosophy only ridiculous” (T 1.4. 7; SBN 185).

**References**


RESCUING POLITICS FROM LYING AND HYPOCRISY: UTILITY AND TRUTH IN JEREMY BENTHAM’S THOUGHT

RESGATANDO A POLÍTICA DA MENTIRA E HIPOCRISIA: UTILIDADE E VERDADE NO PENSAMENTO DE JEREMY BENTHAM

Abstract. As a utilitarian philosopher, Jeremy Bentham (1748-1832) devoted most of his career to create a new institutional setting to promote the greatest happiness of the greatest number on the basis of the principle of utility. In doing so, Bentham committed himself not only to the value of utility but also to truth. Truth meant universal interest and denouncing all abuses of power and corruptions that sacrifice the interest of the greatest number. As two opponents of truth, lies and hypocrisy are the symptoms of the sacrifice of the universal interest and directly oppose the principle of utility. Yet, what is at stake in this opposition? How did Bentham value and analyze these two phenomenon? In this paper, I will explain how Bentham's commitment against lying and hypocrisy in politics aims to support the truth of the utilitarian principle and demonstrates the wrongness of other moral principles.

Keywords: Hypocrisy, Lying, Asceticism, Principle of Utility, Utilitarianism, Jeremy Bentham.

Sumário. Como filósofo utilitarista, Jeremy Bentham (1748-1832) dedicou a maior parte de sua carreira à criação de um novo ambiente institucional que promovesse a maior felicidade do maior número de pessoas, com base no princípio da utilidade. Assim, Bentham comprometeu-se não só com o valor da utilidade mas também com a verdade. Verdade significava interesse universal e a denúncia de todos os abusos de poder e corrupções que sacrificavam o interesse do maior número de pessoas. A mentira e a hipocrisia, enquanto dois adversários da verdade, são os sintomas do sacrifício do interesse universal e opõem-se diretamente ao princípio da utilidade. No entanto, o que está em jogo nesta oposição? De que maneira Bentham valorizou e analisou esses dois fenômenos? Neste artigo, explicarei como o compromisso de Bentham contra a mentira e a hipocrisia na política visa apoiar a verdade do princípio utilitarista e demonstrar a injustiça de outros princípios morais.

* Associate-Professor of moral and political philosophy, ESPOL, Catholic University of Lille.

Bentham asserts the articulation\(^1\) between truth and the principle of utility throughout his career. While the relation between truth and utility remains highly problematic and difficult to understand completely in his philosophy (Quinn, 2012; Schofield, 2015),\(^2\) lying, as the opposite of truth, questions how Bentham’s collective utilitarian thought deals with non-true statement in his effort to create impartial and just public institutions to realize the greatest happiness of the greatest number. In this debate, Bentham thought that discourses are always the mirror of individuals’ interests and, if each individual is the best judge of his own interest, it is highly probable that his discourse expresses his interests and his desire to promote them. As a collective utilitarian who devoted his entire career to reform institutions, legal codes and public policies, then, the issue of lies in public discourses remains of special importance. Indeed, with this topic, Bentham engages with a specific question in his thought: could the universal interest be promoted by public institutions while public officials looked out for their own interests? How could the adequacy between public official’s minds and their functions in the institutions be created and tailored on the basis of the principle of utility? As I will explain, Bentham presents a definitive answer to this issue that he strongly supports throughout his writings. He thinks that public officials’ lies are the symptoms of the “sinister interests” (Schofield, 2006, Ch. 5), i.e. private interests that will sacrifice the universal interest. Hence, in order to secure the greatest happiness of the greatest number, Bentham commands the complete exclusion of lying in the ethics of politics.

---

\(^1\) For instance: “D’abord à force de faire agir cette phrase, on s’est égaré du chemin de l’utilité ; seul guide qui mène à la vérité en faite de morale ; seul dictateur légitime de nos jugemens de la bonté et de la méchanceté des actions” (Bentham, 2016, p. 389).

\(^2\) In his article, Schofield (2015) argues that the relation between truth and utility are not unequivocal in the sense that the different alternatives that describe their relation might all appear in some domains of Bentham’s thought (whether utility subordinate truth, truth subordinates utility and, lastly, whether both are mutually supportive). In the case of religion, Schofield argues that utility and truth are mutually supportive.
In order to understand Bentham’s radical commitment to exclude lying in the ethics of politics, I will proceed in two steps. Firstly, I will explain how Bentham thinks that the mind of public officials and their interests should integrate the utilitarian code until it is perfectly internalized. Bentham’s ethics of politics proceeds in different stages from the selection of public officials to the development of specific aptitudes. Ultimately, all public officials should comply with the set of rules and aptitudes promoted by the utilitarian code. These aptitudes will lead public officials to fulfill a duty of sincerity that demonstrates their commitment to the greatest happiness of the greatest number. Secondly, the role of the liberty of the press remains central for securing the interest of the greatest number. Bentham thinks that newspapers and public opinion, even based on bad faith and lies, are necessary expressions of conflicting interests and remains the best control of public institutions. Bentham tolerates lies made against public officials and rejects all limitations of the liberty of the press and liberty of speech in general. Newspapers and the public have the power to control the rulers and this power should not be bounded.

1.1 Securing the Perfect Internalization of the Utilitarian Code.

Bentham’s ethics of politics is mainly developed under the name of “public deontology.”

By Deontology, taken in its largest sense, is meant that branch of art and science which has for its object the doing on each occasion what is right and proper to be done. In its ordinary acceptation however, the use of it was confined to that part of the field of thought and action which is considered as forming the subject matter of Government and Morality. For distinction sake, in so far as it takes for its subject matter that part of the field to which Government applies itself, Public Deontology is the name that may be employed: in so far as the application of it is considered as confined to those parts of the field of thought and action which Government has left free, Private Deontology is the name that may be employed (Bentham, 1983a, p. 249).

“Public deontology” constitutes the largest part of the field of ethics in Bentham’s thought. With this concept, Bentham aims to explain how just and impartial public institutions require that each and every public official’s interest supports the general utilitarian maxim, the greatest happiness of the greatest

---

3 The problem of compliance to the set of rules that shape public institutions is largely characterized as authoritarian in Bentham’s public philosophy.
number. Just and impartial public institutions rely ultimately on the ability of public officials to act, think and make decisions that are guided by the principle of utility. Bentham’s “public deontology” is mostly described and theorized in Bentham’s writings (Bentham, 1989a; 1989b; 1993; 1995; 2002) on constitutional law, institutional design, arrangements, publicity and transparency.

Bentham’s ethics of politics relies on a system of control and motivation devoted to create the artificial junction (Harrison, 1983, Ch. 5) between the interest and the duty of the functionary. The utilitarian ethics of politics is built on the idea that a system of motivation and control based on reward would be counterproductive since it would not find an objective measure between the reward awarded and the benefit obtained by the official and the good government. Instead of the promise of a gain and more generally of a positive utility, a system of control and motivation made of rules, arrangements, publicity and based on clear definitions of sanctions would influence the conduct of the public official in a more effective way so that he will be encouraged to act adequately with the utilitarian code. I would like to add that this argument is largely consistent with Bentham’s anthropology which considers that the value of pain is always greater than that of pleasure from the motivational point of view and thus offers, from the epistemological point of view, a more reliable prospective objectification of states of things that respond to the principle of utility. As such, a utilitarian ethics of politics based on a system of control of motivation is then the best approach for achieving the greatest happiness of the greatest number.

1.2 Selection of Public Functionaries.

That said, a utilitarian ethics of politics inspired by this general idea aims firstly to find the appropriate candidates for public office and needs then to proceed to the selection of functionaries in conformity with this system of motivation and control. The selection of public functionaries is guided by a maxim that Bentham quotes in many texts: “Official Aptitude Maximized; Expense Minimized” (Bentham, 1993). The application of this maxim is described in its complete form in Bentham’s masterpiece Constitutional Code. In Chapter 6, Bentham details how the conditions of remuneration of members of the
National Assembly are articulated in acts of presence where an “entry procedure” and an “attendance code” are fixed. The control of the presence of the Deputies leads to calculate the number of days of work corresponding and help to calculate their salary. The system of remuneration is also defined by the principle of “minimization” (Bentham, 1989a, pp. 297-310) which implies that the selection of the best candidate requires that they compete for the lower remuneration. These elements demonstrate how the system of motivation and control is a procedural, authoritarian and competitive method that describes the different stages through which a candidate has to undergo until he is hired by the State for a particular office. But the selection of candidates is not the only subject. The utilitarian code aims more generally towards the transformation of the interests of the public official in such a way that his interests will no longer be partial and personal but will suit perfectly with the duty that the principle of utility commands. Bentham’s ethics of politics is centrally based on the idea of the artificial identification of one public official’s duty towards his interests. This artificial identification of the duty and the interest is narrowly related to the concept of “aptitudes”. This concept explains how one public official internalizes the utilitarian code designed according to his office. After the selection and the long procedure of evaluation of the candidates, the new public official is trained and educated to incorporate the utilitarian code composed of moral and legal prescriptions. With this code, the “universal interest” is secured and, the public official also knows which sanctions will apply if his actions follow his own private interests and despise or misunderstand his duty.

### 1.3 The Doctrine of Aptitudes.

In order to secure the impartiality of public institutions, Bentham develops the idea that good functionaries have to internalize different “aptitudes”. The doctrine of aptitudes explains how the incorporation of the utilitarian code is realized. Bentham defines the “aptitude” as the “quality of constitutional arrangements” (Schofield, 1996) and not as an individual ethos. It is the capacity and disposition created by these arrangements for an individual to internalize the utilitarian code. Bentham distinguishes three different “aptitudes”.

Appropriate aptitude is aptitude with relation to the end. In the case of a functionary of Government, appropriate aptitude is appropriate official aptitude. Of appropriate aptitude with relation to the end of government and laws, three branches or elements
require to be distinguished, namely, 1. Appropriate moral aptitude; 2. Appropriate intellectual aptitude; 3. Appropriate active aptitude (Bentham, 1989b, p. 4).

These aptitudes articulate different relations between the individual mind and the content of the utilitarian code. The “moral aptitude” refers to the ability to act and think according to the principle of utility. This capacity is fundamental for public officials. The adjective “moral” qualifies the ability of the officials to think like a utilitarian.

... moral aptitude is a negative quality: it is constituted by the absence, in so far as possible, of a certain propensity universal in human nature. This propensity in the breast of each individual is the propensity to sacrifice all other interests to that which at each moment appears to him to be his own preponderant interest: to obtain happiness for himself to whatsoever amount by so doing, he abstracts or withholds it from all other individuals (Bentham, 1989b, p. 13).

The concept of “moral aptitude” explains how one individual considers subjugating his personal interests in order to fulfill the requirement of his office in the public institution. The artificial identification of his interest with his duty proceeds from his ability to sacrifice his personal interests.

... desire is kept down and weakened: the propensity exists but in the character of a lambent flame by which no combustion is ever produced. This is then what is possible to be done, done by a form of government such as the desire of rendering the exercise of it conducive to the greatest happiness of the greatest number is capable of discovering and recommending for adoption (Bentham, 1989b, p. 13).

The sacrifice of interests also accompanies the conversion of desires and interests so that their transformation leads to the artificial identification with the interest of the greatest number. The “moral aptitude” shows also that one individual does not master perfectly his own interests because they are dynamic, contingent and evolving if necessary. The plasticity of the interests is one dimension on which “moral aptitude” relies. However, the sacrifice of some of his personal interests is not only based on the force of his will but depends on the ability of the functionary to understand and think about the reasons of his sacrifice.

Bentham introduces then the “intellectual aptitude” which is grounded in the faculty of understanding in order to emphasize that the rationale of the sacrifice must be understood. The functionary should respect the sacrifice and not be blindly obedient to the utilitarian code. This “intellectual aptitude”
requires an “Examinee” (Bentham, 1989b, p. 77) who will assess the proper understanding of the functionary. The examination of the “intellectual aptitude” of the functionary relies on his ability to understand and master the appropriate knowledge of his office and the rights and duties attached to it. The knowledge of his duties and his ability to judge correctly are fundamental features of this aptitude. Lastly, the “active aptitude” is defined as the ability of the functionary to be involved in and attached to the realization of the greatest happiness of the greatest number. The public official should not only respect and act in conformity to his duties, and be informed of the normative dimension required by his office in order to act as a utilitarian, because it is also important that the public servant must want to be a utilitarian. This last point gives a full description of what the perfect internalization of the utilitarian code for a public official comprises.

This analysis shows then that Bentham’s ethics of politics aims to create an artificial junction between the duty and the interests of the functionary that will at the same time assure the impartiality and justice of public institutions and secure the universal interest from all sacrifices. Bentham thinks that the direct effects of these aptitudes will be concretely demonstrated by the “probity” and “sincerity” of public officials. Bentham focuses on this aspect in two texts in Political Tactics (1788) and The Book of Fallacies (1824). By describing the particular concrete situations, Bentham explains how deputes and public officials shall always speak the truth in the name of the “universal interest” and their own duty and interest to promote the greatest happiness of the greatest number. As such, lying is first and foremost always the mark of a corrupted mind, of the dissimulation of one’s own interests.

According to the notions commonly entertained of moral duty under the head of probity, and in particular under the head of that branch of probity which consists in sincerity, whatsoever be the nature and extent of the business in question, private or public, it is not right for a man to argue against his own opinion: when his opinion is so and so, to profess it to be the reverse, and in so doing to bend the whole force of his mind to the purpose of causing others to embrace this opinion thus reverse to his real one (Bentham, 2015, p. 486).

Lying is described as one individual masking his authentic and subjective interests. In the broad sense, lying is bad because one is dissimulating his own

---

4 “Any person who, in the character or situation of a Candidate for the Office, is regarded as being willing that his aptitude with reference to the Office shall be taken into consideration and enquired into, call him an Examinee” (Bentham, 1989b, p. 77).
interests. Applied to public officials, it means that the artificial identification between his interest and his duty is broken and that he does not fulfill his moral duty of sincerity. Masking one’s own authentic interests with a lie is masking the non-identification between one public official’s interests and his duty which is normatively shaped to respond to the utilitarian end. The dissimulation demonstrates that the aptitudes that the public official should have internalized in order to secure the greatest happiness of the greatest number can fail to guarantee his impartiality. In this case, Bentham considers that the source of a lie made by a public official relies most often in the lack of “intellectual aptitude”:

For the performance of it, on the part of the individual corrupted no extraordinary degree of intellectual aptitude – nor so much as the most ordinary human intelligence – is necessary: the degree of intellectual aptitude universally possessed by all domestic animals of an inferior degree is quite sufficient (Bentham, 1989b, p. 21).

The public official who lies is showing less intelligence than when he tells the truth. This idea shows that the selection of candidates and the internalization of the utilitarian code is demanding and that the intellectual aptitude must respond to the rationality that the principle of utility is asking for. Bentham’s utilitarian ethics of politics is rational and his liberal optimism interrogates ultimately the degree of rationality that a public official must possess in order to be recognized as fulfilling his duty and doing good work and, correlativey, the degree of authoritarianism the utilitarian code has to have to be able to eliminate lying in politics. If for Bentham, there is no possibility of justifying the necessity of lying in politics, the consequence of this view is that the commitment of the public official to say the truth is then necessary. This idea contradicts the specific political relations and dependencies that exist in politics in general.

Indeed, even during election campaigns, periods of high and intense political debates, the man in charge of the destiny of the interests of the greatest number shall not lie. Beyond all political dependence, beyond all commitment to his own political allies and parties, the words of one public official are the reflection of his interests and, as such, express how his will and his understanding are articulated, i.e. how his state of mind is prepared to respond to the “universal interest”. The principle of the greatest happiness of the greatest number requires a perfect and total incorporation of the utilitarian code, a sacrifice of all personal interests. However, the incorporation of the utilitarian code is not enough and
the system of control and motivation could be violated. Multiple errors and failures from public officials are always possible and can have many sources. The internal system of control and motivation needs an external support grounded in the role of the critic and free censure articulated by the public and the press. Bentham thinks that the “souls of justice” (Postema, 2014, Ch. 3) lies in publicity and transparency. The more controlled a public official is, the more he wants to respect his duties and act accordance with the “universal interest”. The duty of sincerity and truth is then encouraged most notably by the sanction of the public and the discourses of the press.

1.4 The Censorial and Critical Function of the Liberty of the Press.

Bentham’s writings on the liberty of the press belongs to a particular historical context. These writings are mostly inspired by Spanish politics in the 1820s and were initiated as a reaction to the prosecution and imprisonment of ‘Mora’ (an editor), a victim of the policing and censure system exercised by the Cortès. Bentham wrote many texts addressed to the people of Spain. In these letters, he defends the liberty of the press as being not only a fundamental and constitutional right but also a power which has a critical and informative function because it helps secure the maintenance of a good government. Bentham asserts that one can measure the ability of a good government – and of its ability not to lie – by looking at its attempt to limit the liberty of the press.

... to place, on any more advantageous footing, the official reputation of a public functionary, is, to destroy or proportionably to weaken, that liberty, which, under the name of the liberty of the press, operates as a check upon the conduct of the ruling few; and, in that character, constitutes a controlling power, indispensably necessary to the maintenance of good government (Bentham, 2012, pp. 12-13).

The liberty of the press is a symptom of a liberal government where the securities against misrules and the fundamental constitutional rights are secured and do not suffer any limits from the censure of the government. But if a free press is a proof of good government, it is also a condition for freely controlling and criticizing the government and the work of public officials. Yet, the liberty of the press is not a sufficient condition of the true-statement discourses developed by newspapers. Indeed, if in some way, the press is a ‘guardian’ of the good government, it is not committed to good-faith, evidence-based and enlightened criticism but can tell lies, make calumnies and others non-true statements. As
such, the critical role of the press leads to a paradox. In order to explain how Bentham solves it, it is necessary to look more closely at the text. Bentham creates several distinctions about the different modalities of lies that can appear in newspapers:

> If the imputation is, to a certain degree, particular, - imputing an individual act legally punishable, or at least disreputable, - it constitutes the sort of act expressible by the term defamation: if, to a certain degree vague and general, vituperation. But these sorts of acts are, both of them, commonly treated on the footing of offences: and this too, even where the person who is the subject of the imputation is a private individual, not bearing a part in the exercise of any of those powers (Bentham, 2012, pp. 11-12).

Bentham extends the classical understanding of lies as a false statement with the introduction of a distinction. The category of “defamation” is here a specific category of discourse which is characterized by the precise harm imputed to one individual and is opposed to “vituperation” where the negative effects of the discourse is not precisely defined. Bentham qualifies these acts as “offences”\(^5\). The acts of “defamation” can be defined as a legal offense of the type of “private extra-regarding offences” following the definition of acts that “are detrimental, in the first instance, to assignable persons other than the offender” (Bentham, 1970, p. 188). In this case, a society that expresses the will to define “defamation” as an offence can qualify legally this act as a “private extra-regarding offence” according to the principle of utility. “Defamation” is a border-line case for Bentham because its legal qualification as a legal offence is conditioned by other circumstances that include legal evidence, testimonies and the possibility for the accused to defend himself. In his writings on the liberty of press for Spaniards, Bentham seems to be much more concerned with the fact that the liberty of the press is not a medium for discharging wrong propaganda, but instead, that it should be used in its positive function of controlling the impartiality of public officials.

In the case of the public functionary, for vituperation, how gross so ever, there should be no punishment at all: for defamation, no punishment unless the imputation be false and groundless: nor even then, unless the false assertion, or insinuation, be the result of wilful mendacity, accompanied with the consciousness of its falsity, or else with culpable rashness: namely, with that which is exemplified by the giving credence and currency to an injurious notion, adopted without any, or on palpably insufficient grounds: no separate judicatory: no separate form of procedure, styled penal or criminal, while, in the other case, it is styled civil; and, in the case of

\(^5\) “Any act may be an offence, which they whom the community are in the habit of obeying shall be pleased to make one: that is, any act which they shall be pleased to prohibit or to punish. But, upon the principle of utility, such acts alone ought to be made offences, as the good of the community requires should be made so” (Bentham, 1970, p. 188).
defamation, in disproof of rashness of assertion, as well as of wilful falsehood, the
defendant should be at liberty to make proof of the truth of the imputation; and, for
that purpose, to extract evidence from the person who is the subject of it, as he might
from any other person at large.

For these notions, speaking in general terms, my reason is – that to place, on any
more advantageous footing, the official reputation of a public functionary, is, to
destroy or proportionably to weaken, that liberty, which, under the name of the
liberty of the press, operates as a check upon the conduct of the ruling few; and, in
that character, constitutes a controuling power, indispensably necessary to the
maintenance of good government (Bentham, 2012, p. 12-13).

In this text, Bentham asserts that the false statements and the deception
generated by these forms of lies *weaken* the liberty of the press, in the sense that
they diminish its ability to control and judge freely the power exercised by the
rulers. Indeed, the power to criticize and censure public officials is partly the
responsibility of newspapers and is crucial for good government. Bentham’s
defense of the absolute liberty of the press lists the danger and mischievous
political effects that could result from it. This list will never justify the necessity
of limiting the freedom of the press. In contrast, limiting the freedom of the press
is always the mark of a government that wants to promote the interests of the
rulers and sacrifice those of the people.

Bentham also presents a second justification in favor of the liberty of the
press that has its root in the analysis of the interests of newspapers and the role
of editors and the journalists. Indeed, there should be no boundaries to the liberty
of the press because a discourse always expresses an interest and whether the
editor of a newspaper is animated by the truth or by the intent to seduce new
readers in the public, there cannot be enough wrongs created that could justify
the limitation of the liberty of the press. Bentham thinks that public opinion is
largely shaped by the interests of the editors who participate in the fabrication of
the public opinion. As such, Bentham permits lies because he cares a lot more
about the collective effects of lies created by rulers than created by an editor. An
editor’s view is not separated from the interests of the individuals who comprise
the public opinion and, as such, an editor’s lies do not lead to the perversion or
the total corruption of the society. Bentham never takes the idea seriously that
permitting lies could end exacerbating the generalization problem6 (what would
happen if everyone did the same thing?). There is a gap between forbidding lying

---

6 On the generalization test in utilitarian thought, see: Lyons (1965, Ch. 1).
for public officials because of the risk of sacrificing the interests of the greatest number and the deceit and lies existing in a society. For Bentham, a society where there is a very significant number of lies, calumnies, false statements, unfaithful discourses produced by newspapers and the public is the mark of, what he calls, a society of “antipathy” (Bentham, 1970, pp. 58, 60-61) and this society is not ruled under the principle of utility.

Antipathy qualifies the relation between conflicting and opposed interests and these lead to reciprocal violence and conflicts. A divided society such as colonies represent a society where antipathy nourish the conflicts of interests between individual and public officials. In colonies, the degree of mistrust, disobedience and antipathetic sensibility is so important that the corruption affects the State, the rule of law and poisons all public institutions. Bentham’s criticism of colonies is grounded in the observation of the corruption of public officials, governors who exemplify how the abuse of power is always used to promote their own sinister interests against the interests of the people. The corruption of the State and its official in colonies is then a proof of how antipathy forms the basis of the collapse of a political society that runs the risk of falling back into the “natural society”.

In this section, I have explained why Bentham takes the moral and political implications of lying in public institutions very seriously. The utilitarian ethics of politics forbids lying in public institutions. The artificial identification of the duty and interest of the functionary with the utilitarian maxim is necessary to secure impartial and just institutions.

2. **Hypocrisy and Hypocrites: Utility against Asceticism.**

Bentham’s concern with lying is directly related to the utilitarian principle on which public institutions are based. As a collective utilitarian, his unease regarding lying is associated to the broader subject of the impartiality of public

---

7 “They hate you … You hate them. For what would you wish to govern a people who hate you? Will they hate you the less for governing them? Are a people the happier for being governed by those they hate? … For what can you wish to govern a people whom you hate? Is it for the pleasure of making them miserable” (Bentham, 2002, p. 294).

8 “The idea of a state of natural SOCIETY is, as we have said, a negative one. When a number of persons are supposed to be in the habit of conversing with each other, at the same time that they are not in any such habit as mentioned above, they are said to be in a state of natural SOCIETY” (Bentham, 1977, pp. 428-429).
institutions. Hypocrisy seems to be inscribed in the continuity of Bentham’s care to provide securities so that the universal interest will not be sacrificed by the rulers. In his study on *Political Hypocrisy*, David Runciman (2010) devotes a chapter on the philosopher from London. He explains that “the hypocrisy Bentham hates is the hypocrisy of empty words masking real power” or, in other words, “Bentham’s overriding conviction is that words that lack meaning are only insisted on by people with something to hide” (2010, pp. 123-124). While Runciman is absolutely right in explaining that Bentham’s theory of fictitious entity (Schofield, 2006, Ch. 1) is a central element of his attention to discourses in general, yet, the source of Bentham’s opposition and discussion on hypocrisy is first and foremost a debate on the nature of morality, its foundation and the truth of the principle of utility. In this second part, I will explain how much hypocrisy matters for Bentham’s defense of the principle of utility.

Bentham’s utilitarian thought is associated with the utilitarian motto *the greatest happiness of the greatest number*. The principle of utility is the foundation of Bentham’s entire utilitarian thought and aims to promote the happiness of the overall society:

> By the principle of utility is meant that principle which approves or disapproves of every action whatsoever, according to the tendency which it appears to have to augment or diminish the happiness of the party whose interest is in question: or, what is the same thing in other words, to promote or to oppose that happiness. I say of every action whatsoever; and therefore not only of every action of a private individual, but of every measure of government (Bentham, 1970, pp. 11-12).

Morality requires the promotion of pleasure for one individual and for the happiness of the greatest number. Following the research for rationality in morality, the principle of utility, Bentham argues, is the only rational principle that recognizes the “governance of two sovereign masters, pain and pleasure” (Bentham, 1970, p. 11). Grounded in a hedonist psychology, Bentham’s moral and legal thought targets the promotion of pleasure and the avoidance of pain for making the society and the world a better place. Bentham’s utilitarian ethics articulates the promotion of happiness with morality in every aspect of human’s life.

Yet, Bentham’s writings show that he developed a great deal of interests in the opposite theories (Bentham, 1970, Ch. 2; 2013; 2014) to the utilitarian ethics namely, the “principle of asceticism” and the “principle of sympathy and
antipathy”. The term “hypocrisy” is mainly used by Bentham to qualify the advocates of the non-utilitarian principles and to value negatively the nature of their error. But why does Bentham reserve the terms “hypocrisy” and “hypocrites” for his adversaries and mostly for the proponents of the principle of asceticism? How does Bentham understand “hypocrisy” and why does he root his understanding of “hypocrisy” in the principle of asceticism? The answers to these questions will, correlatively, enlighten the relation between utility and truth in Bentham’s thought.

2.1 Utility and the Principle of Asceticism.

Bentham advocates the importance of the principle of asceticism since Introduction to the Principles of Morals and Legislation. Contrary to utilitarian ethics, ascetic ethics “make[s] war upon pleasure”. Ascetic ethics commands not to promote pleasure and happiness; it is composed of “two branches” a negative (which requires the sacrifice of all pleasure) and a positive (which is dedicated to purity and nobility of the sentiment).

By asceticism I understand any system or article of doctrine in and by which endeavours are used to engage men to forego pleasure in any shape for any other cause than the procurement of still greater pleasure in the same or some other shape, or the avoidance of pain to an amount more than equivalent: or to subject themselves [to] pain for any other cause than the avoidance of still greater pain, or the procurement of pleasure to an amount more than equivalent (Bentham, 2013, pp. 10-11).

One individual is guided by the principle of asceticism when his moral conduct shows, generally speaking, the rejection of pleasure and happiness broadly understood as associated to a pretense of acting as a virtuous man. Bentham underlines the direct consequences of this ethical standard, namely its opposition to the principle of utility and its capacity to be “destructive and obstructive of happiness”.

Per asceticism, forego pleasure, not to avoid pain, but on its own account: so as to embracing pains. This is utility opposed directly (Bentham, 1983b, p. 30)

The motto of the ascetic is to oppose interests and pleasures to virtue, duty and morality. The sacrifice of all pleasures and interests is the first condition of morality. The ascetic denies the centrality of pain and pleasures and is radically opposed to any hedonist moral psychology. While Bentham does not ignore these oppositions, his analysis goes further and aims to describe precisely what sort of
morality and ethical principles is involved in the ascetic conduct. Firstly, says Bentham, the ascetic is not against all pleasures but prefers one type of pleasures which is connected to his own character and taste. The virtuous ascetic is looking for the “pleasures of a good name” or “love of reputation”. Bentham offers a description of these pleasures:

... the pleasures that accompany the persuasion of a man’s being in the acquisition or the possession of the good-will of the world about him; that is, of such members of society as he is likely to have concerns with; and a means of it, either their love or their esteem, or both: and as a fruit of it, of his being in the way to have the benefit of their spontaneous and gratuitous services. These may likewise be called the pleasures of good repute, the pleasures of honour, or the pleasures of the moral sanction (Bentham, 1970, p. 44).

Bentham describes the pleasures which form the basis of the motives of an ascetic conduct. An ascetic individual values the “pleasures of good repute” or “love of reputation” more than any other pleasures. This pleasure is rooted in the way others judge and estimate his conduct to be more demanding and prestigious than other regular conduct. As Bentham says, “The desire of becoming to men in general an object of admiration” (Bentham, 2014, p. 16) shows the social cause that are the basis of this pleasure. An individual is animated by the social prestige associated to one conduct that appears to be exceptional and highly estimate. This analysis unveils the foundation of the virtue by showing that the ascetic is not virtuous for the sake of virtue. Beyond this analysis, Bentham is also interested in the different discourses that try to justify and explain this moral attitude. Bentham identifies the philosopher and the religious man as two traditional types of men who develop ascetic theories. These two types of men are qualified as “moral hypocrites”:

... the sort of man in whom the propensity to asceticism has for its efficient cause the desire or love of reputation may be termed the philosophical ascetic: he in whose breast it has for its efficient cause the desire of ingratiating himself with the almighty being, the religious ascetic (Bentham, 2014, p. 16).

The philosopher – such as Socrates – is looking for the social distinction of others citizens and admiration from the young. The religious man – such as Paul – is interested in gaining God’s favor. With these two portraits, the “love of reputation” finds different application which, in both cases, result in the deceit of the virtuous conduct. The ascetic is not the authentic virtuous man that he pretends to be. He is very much like a person who creates a virtuous facade to
attract the favors of others whether in order to create a dependency, a domination or a form of admiration.

Hypocrisy, as a moral category, qualifies then how some individuals pretend to be virtuous, assert the absolute sacrifice of all pleasures while looking for one specific pleasure that Bentham calls the “love of reputation”. But Bentham’s interests lies in the “hypocrisy of the hypocrites”, i.e. the authors of the double discourse developed by some individuals, groups and authorities who, in the society, promote and seduce the public with this discourse. Bentham’s interest in “hypocrisy” is mostly concerned with the ultimate basis of these discourses, i.e. the principle of asceticism as the direct opposite of the principle of utility.

Most of Bentham’s analysis on hypocrisy demonstrates the wrongness of the principle of asceticism by describing its consequences (whether moral, political, social) and strengthens the truth of the principle of utility as the best moral principles and a powerful critical principle that reveals the errors of its opposing theories. Bentham detects a hypocritical approach in the discourse developed by most respectable figures such as religious authorities, aesthetic critics and individual authorities.

A man’s conduct may be said to be governed by the principle of asceticism in so far as, on any other consideration than that of a more than equivalent good considered as about to accrue either to himself or some other sensitive being or beings from the operation of any other than a supernatural cause, he subjects himself to pain in any shape, considered as pain, or avoids receiving pleasure in any shape considered as pleasure (Bentham, 2014, p. 16).

Bentham’s definition of the principle of asceticism explains in which sense the principle of asceticism is the strongest opposite of the principle of utility. In short, the terms of the opposition are known and can be listed in a series of distinctions: one claims to “make war against pleasure” while the other claims that pleasure is the only basis possible for morality, one is commanding the purest sentiment and the discipline of the will, the other despises all hierarchy of pleasures in morality, one is promoting the sumnum bonum, sentiments and model of virtue, the other is grounding morality in objective measure and calculation etc. This series of opposition has an end that Bentham situates in the description of the bad consequences that follow from the principle of asceticism in politics. Bentham believes that politics can end the dispute and demonstrate once and for all the truth and superiority of the principle of utility.
2.2 Asceticism and International Politics.

In her study on hypocrisy, Ruth W. Grant concludes by saying that “liberalism can be criticized, not for being hypocritical, but for refusing to acknowledge the necessity of hypocrisy” (1997, p. 176). She pursues with the following diagnosis:

... liberal theory does not take sufficient account of the distinctive character of political relations, of political passions, and of moral discourse and so underestimates the place of hypocrisy in politics. This is a real and often unrecognized weakness of liberal theory (Grant, 1997, p. 177).

While this assertion remains true overall in the case of Jeremy Bentham, Bentham’s view on international politics disproves in part this judgment. Indeed, in this last section, I would like to explain how the international consequences of the tenets of the principle of asceticism is, for Bentham, a key aspect of its moral hypocritical character. The principle of asceticism leads to the most disastrous consequences in world politics and this is not only an argument against the principle of asceticism; for Bentham, it shows the truth of the principle of utility. While Bentham’s argument remains unsatisfactory and problematic, the truth of the ethical principle has to be asserted from the tribunal of sensory experience. Bentham considers this to be the last criteria from which a moral theory cannot escape and should be judged.

Bentham identifies several historical and present experiences in international politics where the principle of asceticism appears to have caused devastating effects. Bentham thinks in particular of the causes and reasons for the “holy wars” and others massive human experiences of pains and misery.

It is true, that from the same source from whence, among the religionists, other doctrines and practices, from which misery in abundance was produced in one man by the instrumentality of another: witness the holy wars, and the persecutions for religion. But the passion for producing misery in these cases proceeded upon some special ground: the exercise of it was confided to persons of particular descriptions: they were tormented, not as men, but as heretics and infides. To have inflicted the same miseries on their fellow-believers and fellow-sectaries, would have been as blameable in the eyes even of these religionists, as in those of a partisan of the principle of utility (Bentham, 1970, p. 20).

Bentham identifies the causes of war of religion in the principle of asceticism. The same motive and passion, the “love of reputation” is at the source of the war of religion and the war of conquest. The man in power or the sovereign
is driven by the search for the promotion of the interest of his country; however, the interest of a country promoted by its sovereign is no less his own interest. The sovereign’s interest is distorted or confused with the country’s interest. The “love of reputation” leads step by step to the sovereign’s passionate quest for his own prestige. The primary causes of the war would thus be the close relationship of the passions with the interests of the sovereign. In this description, there is no difference whether it is the Crown of the Pope or the Crown of the King who is the subject to this passion. The “love of reputation” takes a particular form in international politics which has nourished Bentham’s analysis, i.e. the very idea of “glory”.

A king for the sake of gaining the admiration annexed to the name of conqueror (we will suppose power and resentment out of the question) engages his kingdom in a bloody war. His motive, by the multitude (whose sympathy for millions is easily overborne by the pleasure which their imagination finds in gaping at any novelty they observe in the conduct of a single person) is deemed an admirable one. Men of feeling and reflection, who disapprove of the dominion exercised by this motive on this occasion, without always perceiving that it is the same motive which in other instances meets perceiving that it is the same motive which in other instances meets with their approbation, deem it an abominable one; and because the multitude, who are the manufacturers of langage, have not given them a simple name to call it by, they will call it by some such compound name as the love of false glory or false ambition (Bentham, 1970, p. 107).

In this text, Bentham describes how “glory” is the passion, the love of reputation that plays a decisive role for the sovereign or the King in international politics. The passions of the sovereign, and the love of reputation precisely, causes war more than ever. The quest for “glory” leads the sovereign to make the most arbitrary and irrational decisions in order to assuage his own desire to be admired. The pleasure of the sovereign leads to the sacrifice of the happiness of the greatest number with the disastrous consequences of war.

However, Bentham differentiates between “true glory” and “false glory”. “True glory” is when the interests of the greatest number govern States and when the passionate competition and the competition of the selfish interests of sovereigns is replaced by the wisdom of the principle of utility and respect for international law for international justice. “Glory” is not reduced to the negative passion. In the same way, benevolence can qualify the action of the utilitarian and

---

9 From Bentham’s correspondence, I can say that this development on glory and the cause of wars is related to Bentham’s great memories and experience of Fénélon’s Telemachus: “Un roi entièrement tourné à la guerre voudrait toujours la faire. Pour étendre sa domination et sa gloire propre il ruinerait ses peuples” (Fénélon, 1965/1717, p. 107; see also de Champs, 2015, p. 26).
a sentiment, “true glory” is a positive sentiment which can be properly recognized when the action of the sovereign conforms to the principle of utility and participates in the promotion of the interest of the greatest number.

In the same way that the ethics of politics forbids lying for public officials, Bentham’s utilitarian international politics aims to control the sovereign and the government in such a way that “glory” could qualify the just actions and decisions that respect the rights and duties of each country on the international scene based on the principle of utility. The only principle which, at the international level, cannot be deceitful is the principle of utility. The principle of utility secures the promotion of happiness in the world through a utilitarian code that tries to set up a cosmopolitan theory (Bourcier, 2017) based on a legal system which will guarantee the rights and duties of each country in its relation to pleasure and pain.

Conclusion

Bentham’s utilitarian philosophy recognizes the necessity to fight against hypocrisy and lying though both do not have the same value and raise different questions in Bentham’s thought. Bentham’s ethics of politics aims to assure that public officials will not sacrifice the greatest happiness of the greatest number but that the rulers will, thanks to the artificial junction between their interest and their duty, act in accordance with the principle of utility. Public officials’ lies are then forbidden because they express the separation and opposition between one individual’s interests and the duty and function of his office in a public institutions. By exploring Bentham’s ethics of politics, the authoritarian dimension of the utilitarian code appears to be clear on two aspects. Firstly, the utilitarian ethics of politics is based on the premise that it is impossible to realize the utilitarian ends without utilitarian public officials. The good of public institutions depend on the conception of the good that public officials hold. Secondly, moral pluralism in public institutions is radically rejected from the understanding of public deontology: all public officials have to be utilitarian.

Contrary to lying, hypocrisy raises a fundamental question for Bentham which is the truth of the principle of utility. Bentham recognizes that hypocrisy qualifies a conduct that pretends to be virtuous and is grounded in a strong will and the sacrifice of pleasures while, on the contrary, it relies on the motive of “love
of reputation”. More fundamentally, the hypocrites are the theorists (whether in philosophy or in religion) who argue for an idea of morality that demands sacrifice of all pleasures, the promotion of social distinction and pure sentiment and a subjective commitment to morality that have no evidence, no objectivity and no rationality. Hypocrites are paradigmatically non-utilitarian theorists and believers in the principle of asceticism. In order to demonstrate the truth of the principle of utility, Bentham develops a negative argumentation that aims to reveal all the wrongs that the principle of asceticism creates while, at the same time, the truth of the principle of utility.

References


The aim of my paper is to discuss the question of whether in the political climate lying is to be seen as a skill, something that an effective politician must do, or whether it is an absolute “no go” realm. Are lying and deception necessary “skills” for achieving success in politics? Is truthfulness in politics a contradiction in itself? Is the political business as such not dirty by nature? When we think about the remarkable number of lies which have been concocted and distributed by politicians in the recent past, the impression that politics and lies indeed go hand-in-hand emerges, and that the ideal image of politics as a rational instrument for the formulation of generally binding objectives is fraudulent. Should a distinction be made between lies that pursue harmful goals and lies that aim to achieve a good aim? Should politicians be morally justified to lie in order to realise well-meant political objectives?

Keywords: lying, deception, success in politics, truthfulness, the problem of dirty hands, moral justification.

Abstract. The aim of my paper is to discuss the question of whether in the political climate lying is to be seen as a skill, something that an effective politician must do, or whether it is an absolute “no go” realm. Are lying and deception necessary “skills” for achieving success in politics? Is truthfulness in politics a contradiction in itself? Is the political business as such not dirty by nature? When we think about the remarkable number of lies which have been concocted and distributed by politicians in the recent past, the impression that politics and lies indeed go hand-in-hand emerges, and that the ideal image of politics as a rational instrument for the formulation of generally binding objectives is fraudulent. Should a distinction be made between lies that pursue harmful goals and lies that aim to achieve a good aim? Should politicians be morally justified to lie in order to realise well-meant political objectives?

Keywords: lying, deception, success in politics, truthfulness, the problem of dirty hands, moral justification.

Sumário. O objectivo do meu artigo é discutir a questão de saber se no mundo político a mentira é vista como uma competência, como algo que um político eficaz deve fazer, ou antes se é absolutamente interdita. Será que a mentira e o engano são competências necessárias para que haja sucesso em política? Será que a veracidade em política é uma contradição? Será a actividade política enquanto tal suja por natureza ou não? Quando pensamos no número considerável de mentiras geradas e disseminadas por políticos no passado recente, surge a impressão de que a política e a mentira são, de facto, uma dupla inseparável e que a imagem ideal da política como um instrumento racional para a formulação de objectivos gerais vinculativos é fraudulenta. Devemos fazer uma distinção entre mentiras que visam fins nefastos e mentiras que visam fins bons? Devem os políticos ter uma justificação moral para mentir de forma a realizar objectivos políticos bem-intencionados?

* Department of Philosophy, University of Vienna.

1 For the translation of my English abstract into Portuguese I would like to sincerely thank Professor Alexandra Maria Lafaia Machado Abranches. I would also like to thank Professor Pedro Miguel Páscoa Santos Martins for his excellent support during the publication process.
0. Introduction

The paper provides a profound overview of the positions of several philosophers, theologians and politicians about the question of lying in political and moral terms. In the first section, I will analyse the question of how the true nature of politics can be defined. In the second section, I will discuss the requirement of politicians to get their hands dirty. In the third section, I will examine the claim that lying is not generally reprehensible in a moral sense. In doing so, I will try to defend the thesis that lying is, in certain exceptional situations, not only a morally good thing but also a skill that is morally required. The fourth section will focus on a discussion of absolutely prohibiting lying. The fifth section will show the wide range of skills the successful liar has to master. In this section I will discuss positions that claim lying is a skill. In the sixth and last section of the paper, I will defend my position that lying is a morally neutral skill that can be used to realise something good as well as to create horrible evil as such. I will demonstrate that the moral direction of lying is absolutely dependent on the responsibility of the agent.

Jean-François Revel claims that the very first of all powers that rule the world is the lie (cf. Revel, 1990, p. 11). This automatically begs the question of what a lie is. A lie is generally defined as a false statement made with deliberate intent to deceive or as an intentional untruth or falsehood. However, what is frequently disregarded is the fact that the lie is only one single facet of the contrary of truthfulness, but it is the most obvious and most relevant kind of untruth. This is supposedly also the reason why within the field of moral philosophy the lie has since time immemorial played an almost exclusive role when the opposite of truth should be analysed. Nonetheless, the range of untruth should not completely fade into obscurity. A few other examples of mendacity are dissimulation, guile, flattery, breach of promise, obfuscation, diversionary tactic, perjury, forgery, secrecy, deception, whitewashing, boastfulness, hypocrisy,
allegation, manipulation, trickery, fakery, subterfuge, stratagem, collusion or withholding the truth. Michel de Montaigne realises early on in his *Essais* that the flipside of the truth has one hundred thousand forms of appearance and that it has at one’s command an indefinite latitude (cf. Montaigne, 1998, p. 23).

The aim of my paper is to discuss the question of whether in politics lying is to be seen as a required skill – that is, something that an effective politician must do for achieving success – or whether it is something that should be absolutely prohibited as claimed, for instance, by Augustine of Hippo or by Immanuel Kant. A further goal of this paper is to clarify whether truthfulness in politics is or is not a contradiction in itself.

1. The true nature of politics

In *The Prince*, Niccolò Machiavelli outlines quite a dirty image of the politician since he claims that it is neither possible nor necessary for the prince to have all the virtues, but that it is, in practice, even absolutely destructive to have them and to use them all at the same time. According to Machiavelli, for the prince it is sufficient to pretend to have the virtues: “Therefore it is unnecessary for a prince to have all the good qualities I have enumerated, but it is very necessary to appear to have them” (Machiavelli, 1993, p. 139). This principle formulated by Machiavelli claims that seeming is extremely important. It can be seen as the very antithesis of what Cicero asserts in *De officiis*. Namely, that one should genuinely be something instead of merely seeming to be (cf. Cicero, 1992, p. 181).

Hannah Arendt also claims in her essay *Wahrheit und Politik* that there never was a doubt that truth and politics are on rather bad terms with each other, and that no one has ever counted truthfulness to be among the political virtues. According to Arendt, it would seem that lying has always been regarded as a necessary tool not only of the politician’s or the demagogue’s but also of the
statesman’s trade (cf. Arendt, 2013, p. 44). Friedrich Schlegel also argues that the political business as such is by its very nature dirty or something that spoils the character of anybody who comes into contact with it when he says that where “politics or economy is involved, there is no morality” [Translated from the German] (Schlegel, 1800, p. 22). And, to cite a more recent example, Erich Straßner states in his essay _Dementis, Lügen, Ehrenwörter. Zur Rhetorik politischer Skandale_ that the more a politician masters the entire disposable apparatus of manipulative tactics and strategies, the more his reputation increases because solely the virtuosity with its application guarantees the political success (Straßner, 1992, p. 2). Even Plato in his work _Republic III_ says rulers have the exclusive right to lie in order to protect and guarantee the ideal state, whereas subjects are to be punished for lying (389 b6 – 389 d4) (Plato, 1997a, p. 1026). This passage in Plato exhibits a quite dangerous image of political power. However, the greatest potential for causing a stir is the passage in Plato’s _Republic V_ where it reads that for the purpose of the optimal production of children the rulers may deceive and engage in control and selection pertaining to the reproduction among the subjects:

> the best men must have sex with the best women as frequently as possible, while the opposite is true of the most inferior men and women, and, second, that if our herd is to be of the highest possible quality, the former’s offspring must be reared but not the latter’s. And this must all be brought about without being noticed by anyone except the rulers, so that our herd of guardians remains as free from dissension as possible (459 d4 – 459 e2) (Plato, 1997a, p. 1087).

These lines are, without a doubt, a prime example of the dangers that can emanate from political power. At the same time, they raise the question of whether it is the power or the foray into the political business that spoils the politician’s character or whether the reverse is true: a spoiled character spoils politics. On this note, the Austrian politician, Julius Raab, has claimed that it is not politics that spoils the character but a spoiled character spoils politics. Gustav Radbruch, on the other hand, sees politics as a chance to prove one’s character: “Politics spoils character. No, politics tests character. People who work in the political arena where so much is considered allowed which would be impermissible in private life, and who know no other liberal moral other than the one in private life, have a character that has proven itself in politics” [Translated

---

4 For the German original see: Schlegel (1800, p. 22).
from the German] (Radbruch, 2002, p. 216). However, Jacob Burckhardt has quite a negative perspective on power in his work *Reflections on History* and considers power as something evil: “Now power is of its nature evil, whoever wields it. It is not stability but a lust, and *ipsa facto* insatiable, therefore unhappy in itself and doomed to make others unhappy” (Burckhardt, 1950, p. 86, see also p. 38 and 120).

Since power is to be seen as the essential factor in politics, it might be concluded, according to Burckhardt's remarks, that politics is by its nature dirty and evil. In terms of the role of power in politics or, to be more precise, the inextricable amalgamation between power and politics, Straßner asserts in his essay that the acting of our politicians is primarily geared to the acquisition, the enhancement and the maintenance of power (Straßner, 1992, p. 1).

However, as far as power is concerned, the Wounded Knee Massacre (cf. Brown, 1981) is an example of an abuse of power. When we think about the remarkable number of lies which have been concocted and distributed by politicians in the past, the impression emerges that politics and lies are, indeed, an inseparable couple and that the ideal image of politics as a rational instrument for formulating and implementing generally binding objectives is fraudulent or nothing more than a pure illusion. Some examples for the endless range of political lies and scandals, include the Watergate Affair (cf. Bernstein, & Woodward, 1974; see also Woodward, 2005; Woodward, & Bernstein, 1976; Zuelzer, 1975, Ch. 3), the CIA Coup in Chile, the Tyrolean Rebellion in 1809, the Iraq Crisis, the Suez Crisis, the Spiegel Affair (cf. Doerry, & Janssen, 2013), the Chernobyl disaster where important information was concealed, something which we can also see regarding the Pentagon Papers, and, last but not least, the Irangate Scandal. These cases reveal that power and lies are inherently dirty and accompany the pursuit of harmful goals.

Contrary to this thesis, which maintains the dirtiness of power, we may offer a statement from Romano Guardini who considers power to be something inherently neutral whose moral value is inevitably linked with the agent or the agent's responsibility: “In itself, power is neither good nor evil; its quality is

5 For further information see: Green (1996).
determined by him who wields it. In fact, of itself it is only potentially constructive or destructive, since it is essentially governed by freedom” (Guardini, 1961, p. 6). In the end, according to Guardini, it will always depend on the agent and his moral norms as to whether power is something good or not. On closer consideration, in my view this thesis can ultimately be applied one-to-one to the lie. The lie is itself a neutral competence. Originally, it is neither something good nor something evil. It is the agent who gives the lie its moral value or its moral orientation.

2. The problem of dirty hands

On the requirement of politicians to get their hands dirty, Michael Walzer claims in *Political Action: The Problem of Dirty Hands* that no “… one succeeds in politics without getting his hands dirty” (Walzer, 1973, p. 164). In this context, the term “dirty hands” has its origin in Sartre’s play *Dirty Hands* in which Hoederer, for whom purity is nothing more than “… a phantasm of monks and fakirs …” and who interprets politics as a dirty business, declares the following: “I have dirty hands. Right up to the elbows. I have plunged them in filth and blood. But what do you hope? Do you think you can rule innocently?” [Translated from the German] (Sartre, 2005, p. 114). Walzer’s answer to Hoederer’s or, that is to say, Sartre’s question is the following:

My own answer is no, I don’t think I could govern innocently; nor do most of us believe that those who govern us are innocent … even the best of them. But this does not mean that it isn’t possible to do the right thing while governing. It means that a particular act of government … may be exactly the right thing to do in utilitarian terms and yet leave the man who does it guilty of a moral wrong (Walzer, 1973, p. 161).

For Walzer, there are situations when politicians are compelled to lie to defend public interest. Walzer supposes that certain actions can be morally wrong and good at the same time. In other words, the ethics of dirty hands supposes that one and the same action can be morally false and right at the same time. For example, a person who kills somebody to save another person’s life is acting morally wrong and right at the same time. Walzer puts forth the following example of a dilemma which a politician can find himself in:

... he [the new leader, KB] is asked to authorize the torture of a captured rebel leader who knows or probably knows the location of a number of bombs hidden in apartment buildings around the city, set to go off within the next twenty- four hours. He orders the man tortured, convinced that he must do so for the sake of the people
who might otherwise die in the explosions – even though he believes that torture is wrong, indeed abominable, not just sometimes, but always (Walzer, 1973, p. 167).

In Walzer’s example, the action taken by the politician is morally wrong, but the consequences of the action are morally good. The politician decides to do a bad thing in order to do good.

In discussing the problem of dirty hands, I would like to mention Hans-Georg Soeffner who sketches the theory of a kind of ritually supported substitution agreement that is characterised by the practice that the electors give their ambassadors a kind of collective permission to commit moral violations inasmuch as such a violation is needed for the general interest. In this way, the representatives are able to wash their hands of moral guilt (cf. Soeffner, 1998, p. 224). According to Soeffner’s theory, the problem of dirty hands ceases to be a problem. Something similar would happen if Niklas Luhmann’s position was applied to the problem of dirty hands, since Luhmann releases or uncouples the political system from any form of control on the basis of moral criteria. Luhmann’s idea is that the political system regulates for itself the manner and form which is morally relevant. According to Luhmann, political ethics should reflect, in particular, the self-regulating character of political systems (cf. Luhmann, 1993, p. 40). By applying Luhmann’s notion of self-regulation of political systems, the problem of dirty hands would resolve itself. Niccolò Machiavelli has also addressed the problem of dirty hands, albeit indirectly, and provides a general justification of the political lie. In the chapter 15 of The Prince, he suggests that the man who enters political life is required “... to know how to do wrong, and to make use of it or not according to necessity” (Machiavelli, 1993, p. 117). This line of argument can also be found in chapter 18 of The Prince, where Machiavelli states that the prince has to be a master of hypocrisy and dissimulation, a “great pretender and dissembler” (Machiavelli, 1993, p. 138) and that there are situations in which the prince is justified in breaking his word if the compliance of his promises is disadvantageous for his political targets and interests (Machiavelli, 1993, p. 137). The problem posed by Machiavelli is that the emerging politician will, at least sometimes, find himself in a dilemma where the skill to act immorally to reach certain goals will be required. There are circumstances that require one to commit immoral things in order to reach a good
goal. This principle is expressed even clearer in chapter 9 of the first book of *The Discourses*, where Machiavelli points out:

> It is a sound maxim that reprehensible actions may be justified by their effects, and that when the effect is good, as it was in the case of Romulus, it always justifies the action. For it is the man who uses violence to spoil things, not the man who uses it to mend them, that is blameworthy (Machiavelli, 1976, p. 132).

This principle can be applied to a moral dilemma where a person has to choose between two courses of action both of which would be wrong for him to undertake. Such a dilemma happens whenever someone is forced to choose between sticking to certain important moral principles and sidestepping some catastrophe. Dilemmas of this nature show that, at least sometimes, crossing moral principles is unavoidable and, indeed, that it is possibly even more moral to cross them than to adhere to them. In other words, there are goals that are reached by crossing moral borders that could be of a higher moral value than the moral value of the strict compliance with the moral principles.

### 3. Lying is not generally reprehensible in a moral sense: the white lie and other good lies

I defend the thesis that lying is, in certain exceptional situations, not only a morally good thing but also a skill that is morally required. I am convinced that there are circumstances when not telling the truth is not only a good thing but also a moral obligation, in particular if the act of telling the truth would have the inevitable consequence that one life or even several lives would be lost. There are several philosophers and surprisingly even theologians who have formulated numerous exemptive rules which concede a partial permissiveness to lying, if, for instance, the lie is used in a situation of extreme hardship as a kind of saving medicine (Müller, 1962, p. 325), the lie is utilised in a situation of conflict of obligations (Müller, 1962, pp. 274-275), the lie is used for reasons of charity (Müller, 1962, p. 275), the lie is uttered with the purpose of self-defence (Müller, 1962, pp. 275-277), the lie in the sense of a *falsoquium compensativum* (a good purpose justifies an immoral method) (Müller, 1962, pp. 272-273) or if it is a

6 Müller quotes the following authors as examples for such a position: Clemens of Alexandria, Origenes, Johannes Cassianus and Germanos of Constantinople.

7 Examples for this position are: Giovanni Vincenzo Bolgeni (Müller, 1962, p. 327), Hans Christian Oersted or Helmut Thielicke (Müller, 1962, p. 334).
white lie that saves life and limb. Matthias Laros in his book *Seid klug wie die Schlangen und einfältig wie die Tauben* discusses the exemplary case of the friend who is arrested and incriminated by the Gestapo for listening to foreign stations. Laros goes on to discuss the justifiability of the white lie to protect the accused from years of prison (Laros, 1951, pp. 37 ff.). On the topic of white lies, Eberhard Schockenhoff mentions that in early modern times in the course of the confessionalisation of the political regimes everywhere in Europe religious dissidents were forced either to choose martyrdom or to escape into a feigned assimilation and to acquiesce the demanded confessional practice of faith. The feigned participation in public ceremonies of the state’s religion, forced change of religion and mock-baptism were for English Catholics, French Protestants and Spanish Jews the only way of saving their goods and chattels and often even the only way to protect their life (cf. Schockenhoff, 2000, p. 89).

Schockenhoff’s example of circumstances where the state and churches infringed on the right to the freedom of religion with their methods of oppression, is a situation where lying is more than just legitimate. The justifiability of the white lie is also discussed by Sissela Bok. She considers nearly every kind of lie as unjustifiable but regards the white lie that saves a person’s life as justifiable. However, she notes that it is only justified if the other person’s life can, without a doubt, be protected by the lie (cf. Bok, 1980, pp. 65–66, 136–139).

In my opinion, a white lie that protects one’s own life or that of another person should always be legitimate. The immorality of lying is without exception abolished by the act of preventing someone’s death. To my mind, the perfect example for the legitimacy of white lies are those told to save the lives of persecuted ethnic or religious groups. One specific historical example involves the lies told during the Second World War to protect Jews from persecution. This shows once again that the protection of a life should invariably carry a higher moral weight than any imperative moral obligation to tell the truth.

4. The absolute prohibition of lying

---

8 Augustine of Hippo and Immanuel Kant principally refuse all forms of white lie, even the one that protects the life of an innocent.

9 For further details about the discussion of the justifiability of the white lie see also: Schockenhoff (2000, pp. 106–108).
Analysing the role of the lie in politics and society requires a discussion of the position of the absolute prohibition of lying. In doing so, there is no way around Aurelius Augustine who defines and approaches the lie methodically in *De Mendacio* (cf. Augustinus, 1953) by distinguishing eight different levels of the lie. Basing his thesis mainly on the Holy Bible, Augustine considers every form of the lie as a sin and as something that is morally reprehensible, even if the lie injures nobody or if it is the only way to defend people. According to Augustine, in matters of faith lying is not permitted under any circumstances. In the chapter XVIII of the book XIII of his *Confessions*, Augustine also leaves no doubt about the important role truth should play in the life of a faithful Christian (Saint Augustine, 1984, p. 116). I consider it noteworthy that Augustine refers in his *Confessions*, in chapter XXV of the book XII, implicitly to John 8:44 (Saint Augustine, 1984, p. 107), where it is written that the devil is the father of the lie:

You are from your father, the devil, and you want to work the lusts of your father. He was a man-killer from the beginning and hasn't stood in the truth, because no truth is in him. When he speaks the lie, he speaks from his own, because he is a liar and its father (*The Gospel According to John*, p. 89).

In the Augustinian direction but arguing in a different way, namely on the basis of pure rationality, the philosopher and theologian, Johann David Michaelis, pleads in his work *Moral* in favour of an absolute prohibition of lying that he defends with the argument that every form of wilful untruth destroys the social confidence, any credibility and, as a direct result, any ability to convince other fellows of the truth. The result of Michaelis’ intellectual efforts can be seen in the knowledge that the lie makes social cohabitation of people impossible. To demonstrate the direct consequences of deliberate lies, which Michaelis sees in the loss of any convincibility, he cites the example that, given the general permissibility to lie, a chased person cannot be saved anymore from the danger of being killed by a persecutor, who has the intention of killing, by telling a lie since the murderer will probably believe that he has received incorrect information about the abode of his victim and, for that reason, any attempt to deceive him is condemned to failure from the beginning:

---

10 In this context, it should be mentioned Johann Georg Walch who formulates a polemic against the rigorous Augustinian prohibition of lying by creating an example in which is demonstrated that an intentional lie can be legitimate, indeed required (cf. Walch, 1775, p. 2318).
... what is to be done if someone asks me in a rage where the man whom he wants to kill has run? Answer: To tell the truth. The truth will be considered a lie by the murderer. If the murderer is at least not stupid, he will choose the other path and not that one I have showed him. Therefore, with a lie I endanger the life of the chased man just as by telling the truth [Translated from the German] (Michaelis, 1792, pp. 160–161).

Johann Adam Bergk also claims that a general permissiveness of lie would displace any truth since everybody would tell the untruth as soon as anything can be gained, such as when breaking out of a dangerous situation: “Where should the truth be searched in such a case? Is the truth not changed into a fairy tale if everybody who considers oneself to be in distress is permitted to say the untruth?” [Translated from the German] (Bergk, 1798, p. 143). In those days, the positions of Bergk or Michaelis provoked predominantly reactions of disconcertment and refusal since the enlightened zeitgeist supports the position of the lie being permissible inasmuch as a disadvantage can be prevented or an advantage can be achieved by lying. So Michaelis’ absolute prohibition of lying is an attempt to attack the moral theories of the Enlightenment. Michaelis tries to defeat the position of the permission of the lie of the Enlightenment with its own argumentative instruments since he argues not theologically but rationally, exactly like the philosophers of the Enlightenment argue. Immanuel Kant, who read Michaelis’ Moral and whose positions are very similar to those of Michaelis, expresses in The Metaphysics of Morals an absolute ethical and juridical prohibition of lying in an equally rigorous manner: “The greatest violation of a human being’s duty to himself regarded merely as a moral being (the humanity in his own person) is the contrary of truthfulness, lying, (aliud lingua promptum, aliud pectore inclusum gerere)” (Kant, 1996, p. 182). Kant condemns categorically every kind of lies because in his estimation lying and reason are two things that are by all means incompatible. The reason, in turn, is considered by Kant as a condition for the status as a person. Therefore, the liar steals his status as a person, he destroys his human dignity, as Kant concludes in The Metaphysics of Morals:

By a lie a human being throws away and, as it were, annihilates his dignity as a human being. A human being who does not himself believe what he tells another (even if the other is a merely ideal person) has even less worth than if he were a mere thing (Kant, 1996, p. 182).

According to Kant, lying is for a man
... directly contrary to his character as a moral being (in terms of its very form), that is, to inner freedom, the innate dignity of a human being, which is tantamount to saying that they make it one’s basic principle to have no basic principle and hence no character ... (Kant, 1996, p. 175).

In 1796, Benjamin Constant addresses stiff criticism to the rigorous conception of an absolute prohibition of lying of a “German philosopher”, by publishing a pamphlet entitled Des réactions politiques (cf. Constant, 1988, pp. 95–157) in which he refers above all to the example of the murderer who wants to know whether his victim is at home. In this pamphlet, Constant formulates a rejection of the absolute prohibition of lying in a situation outlined in the prominent example of a murderer. Constant argues that moral principles cannot impetuously be adapted to practice, especially to the political and juridical practice but they need conciliatory principles.

In Constant’s estimation, an absolute prohibition of lying that is independent of the validity of the situation can be interpreted as a kind of ruin of any faculty of social cohabitation of people. According to Constant, society is politically inviable if the same is not permitted to lie at least in certain situations. Karl Friedrich Cramer11, the translator and editor of the German version of Constant’s pamphlet, remarked in a footnote that Constant told him that the “German philosopher” he is referring to is Kant. In truth, the position of an absolute prohibition of lying as described in the example of the murderer was defended by Michaelis before Kant, namely in 1792, and Kant’s work, The Metaphysics of Morals, in which Kant develops his casuistry of the lie, contains word-for-word neither the aforementioned example nor the complete argument that Constant uses as a reference point for his pamphlet. On top of this, Kant’s ethical prohibition of lying as formulated in The Metaphysics of Morals cannot be Constant’s reference point because The Metaphysics of Morals was published after Constant’s pamphlet came out, namely in 1797, and also probably after the publication of Cramer’s German translation of Constant’s pamphlet, in autumn 1797. Apart from this fact, a spicy detail to note is that Kant himself confesses in On a supposed right to lie because of philanthropic concerns (his reply to Constant’s pamphlet published in 1797) that he cannot at all recall where precisely it was that he discussed the example of the murderer in his writings, but

11 For further details about the oral annotations of Constant towards Cramer see: Geismann, & Oberer (1986, p. 11).
he admits shortly afterwards with suspicious emphasis that the example was discussed indeed by himself “in some passage” of his writings just like Constant had recited it. Kant assumes, in the aforementioned reply, the complete responsibility and the onus of proof of a provoking example or position that he mentioned not at all up until then. In *On a supposed right to lie because of philanthropic concerns* Kant picks up on the example, raised by Constant, of the potential murderer who wants to be informed whether his intended victim is at home and uses it for justifying his position of an absolute ethical and juridical prohibition of the lie that remains valid even when an assailant with the declared intention to kill and with the weapon in his hands asks for the abode of the innocent whom he is bent on assassinating. Kant takes an explanatory approach to an example that he himself has never mentioned before. A rather obvious reason for Kant’s eagerness to assume responsibility for an example that is not his own could be, in my estimation, that Constant’s example enables Kant to strengthen and enlarge his position in terms of rigorousness for an essential element, i.e. for the principle that there are no exceptional cases in which a man is exempted from his obligation to tell the truth. In *The Metaphysics of Morals*, Kant defends the position that the obligation to truthfulness has to apply irrespectively of all consequences: “That I ought not to lie, no matter how great the benefits to myself and my friend might be. Lying is mean and makes a human being unworthy of happiness” (Kant, 1996, p. 224).

Kant criticises, unlike everybody else, the lie as a communication form that is under no circumstances allowed for a moral subject, distinguishing between an external lie that is directed to another person and an internal lie that is directed against the subject itself. The external lie is considered by Kant as a usage of language against the intention of linguisticality:

But communication of one's thoughts to someone through words that yet (intentionally) contain the contrary of what the speaker thinks on the subject is an end that is directly opposed to the natural purposiveness of the speaker's capacity to communicate his thoughts, and is thus a renunciation by the speaker of his personality, and such a speaker is a mere deceptive appearance of a human being, not a human being himself (Kant, 1996, p. 182).

12 Original title: Über ein vermeintes Recht aus Menschenliebe zu lügen.
Kant holds the position that it is absolutely not allowed that thought things and said things differ. On closer consideration, in this argument a parallel can be drawn to Augustine who claims in *De Mendacio* that not every untruth is a lie: a person who believes in the truth of his statement even if it is not true is not lying: “For which purpose we must see what a lie is. For not everyone who says a false thing lies, if he believes or opines that to be true which he says” [Translated from the German] (Augustinus, 1953, p. 2.). In Augustine’s view not the untruth is the requirement for the lie but the relation that a subject has to the things he is saying. If the subject is convinced that the said things are true, he is, according to this conception, not lying. Augustine claims that the opposite of the lie is not the truth but the reliable statement on which reliability depends, on the one hand, on the accordance between the expressed things, and, on the other hand, on the inner conviction of the speaker that the expressed things are true. This conviction between what is meant and what is said could be described as truthfulness in the conception of Augustine. Augustine’s approach could be seen as a way to untie the lie from the problem of the truth. In this interpretation of Augustine, it is the difference between what is said and what is known which defines the lie. Arthur Schopenhauer, in turn, considers all forms of lying that are manipulating or negating another person’s will, or which bring another person to accept a truth as *his* own when it is not *his* truth (cf. Schopenhauer, 1969, p. 337) as reprehensible and unjust and, therefore, as something that is to be rejected. In this context, it must be remarked that in Schopenhauer’s conception all forms of moral wrong are generally reduced to the following constellation: “The other cases of wrong can all be reduced to the fact that I, as the wrongdoer, compel the other individual to serve my will instead of his own, or to act according to my will instead of to his” (Schopenhauer, 1969, p. 337). Schopenhauer considers the origin of all forms of unjust actions the “motivation”, under which he understands any attempt of subjugating or manipulating another individual’s will. For this very reason, the refusal to tell the truth is for Schopenhauer no form of injustice because it doesn’t manipulate another person’s will. In contrast to this, in Schopenhauer’s view, telling something false is always wrong because it is a kind of force to accept a truth as my truth that is not my truth. In that sense, lying is a kind of behavioural control. It is a form of “motivation” and, therefore, morally reprehensible: “... every imposition of a lie is a wrong. The person who refuses to show the right path
to the wanderer who has lost his way, does not do him any wrong; but whoever
directs him on to a false path certainly does” (Schopenhauer, 1969, p. 338).

In Schopenhauer’s view, the protection of the own will takes on such an
important role that all actions which serve to protect the will are allowed,
inclusively killing another person if this person threatens my own will and this
menace cannot be averted in another way:

If an individual goes so far in the affirmation of his own will that he encroaches on
the sphere of the will-affirmation essential to my person as such, and denies this,
then my warding off of that encroachment is only the denial of that denial, and to
this extent is nothing more on my part than the affirmation of the will appearing
essentially and originally in my body, and implicitly expressed by the mere
phenomenon of this body; consequently it is not wrong and is therefore right. This
means, then, that I have a right to deny that other person's denial with what force
is necessary to suppress it; and it is easy to see that this may extend even to the killing
of the other person whose encroachment as pressing external violence can be warded
off with a counteraction somewhat stronger than this, without any wrong,
consequently with right (Schopenhauer, 1969, pp. 339-340).

Montaigne, in turn, argues quite distinctly compared to Schopenhauer, and
considers lying as a remarkably cursed vice because in his estimation exclusively
by telling the truth we are humans and capable of community. By lying, according
to Montaigne, we are destroying the premise for the personhood and for the
politician and human rights activist, Václav Havel, provides a completely
different approach in formulating a plea for being truthful. In his book Versuch,
in der Wahrheit zu leben he condemns the mendaciousness of the post-
totalitarian communist system and vehemently stands up for the way of
truthfulness, regardless of the risk of unjustified imprisonment, work-related
discrimination and social marginalisation. What really counts in Havel’s view is
the daily fight for the possibility of a life in truthfulness, dignity and freedom.
Havel considers the purity of this fight as the best weapon against the post-
totalitarian structures (Havel, 2000, p. 82). In the 1990 New Year speech, Havel
deems the worst thing to be that he and his fellow citizens are living in a spoiled
moral milieu. Havel speaks about a moral illness because of the habit of saying
something which goes against what is thought (Havel, 1991, p. 9). In this context,
Augustine and Immanuel Kant can once more be recalled in their demand that
thought things and said things must never differ. During his New Year speech,
Havel brought back to his fellow citizens their accompliceship that has its origin,
according to him, in getting accustomed to the totalitarian system, in tolerating
it as an irrevocable fact because this was how the totalitarian system was actually kept alive. Such a tolerating attitude can be interpreted as the breeding ground that enabled a totalitarian system to survive for so long. Havel considers himself and all his fellow citizens, even though to a different extent, as responsible for the lengthy existence of the totalitarian system. According to Havel, nobody is solely its victim but everybody is at the same time its co-creator (Havel, 1991, p. 10). In pursuing his systematic campaign for a life in trueness, Havel urges his fellow citizens in his pre-election speech to vote exclusively for those candidates who are used to telling the truth instead of those who construct a new “truth”, namely “their truth” on a weekly basis (Havel, 1991, p. 83). Havel mentions George Orwell or, to be more precise, the atrocious Orwellian world where an absolute manipulation and oppression of truthfulness predominates (Havel, 2000, p. 81). However, Havel’s critique of the mendacity of the totalitarian system reminds me of a specific passage in Orwell’s novel *Nineteen Eighty-Four* in which even history is something invented by the “Ministry of Truth”. I would like to quote the aforementioned passage because it shows the metamorphosis of the lie into an incontestable truth: “It might very well be that literally every word in the history books, even the things that one accepted without question, was pure fantasy… Everything faded into mist. The past was erased, the erasure was forgotten, the lie became truth” (Orwell, 1996, p. 78).

Another appeal that is strongly related to political and social practice against nearly every form of mendacity comes from Sissela Bok who claims in her book *Lying: Moral Choice in Public and Private Life* that confidence and integrity are valuable resources that are easy to lose but hard to win back. According to Bok, liars should think about the consequences of lying which she considers as harmful in particular for society, namely for the social confidence on which our social coexistence is based. In other words, Bok regards practically every form of lie, even its well-intentioned variants, as detrimental for the society and for the individual (Bok, 1980, pp. 36 ff.). Here, the demerit of the lie is situated in its derivation of the truth. According to Bok, truth is a positive value and every form of falsified transmission diminishes this positive value and, in addition to it, it gradually undercuts the basic function of language that consists in transporting the truth. In my view, this constitutes an unambiguous borrowing from Immanuel Kant who holds a similar position concerning the true purpose of
language. Bok explains the harmful effect of lying for society or rather the uncontrollable nature of lies through the slippery slope argument according to which every lie is inevitably and unstoppably comparable with a chain reaction of more lies with the result being that a general climate of reciprocal deception emerges (Bok, 1980, pp. 132 ff.). Michel de Montaigne had discerned the uncontrollable nature of the mendacity which is close to what we nowadays call a slippery slope. Montaigne considers mendacity as the vice whose origin and development should be disputed relentlessly because otherwise mendacity continues to increase with the liars or the lies brought into orbit. Once the lie is allowed, in Montaigne’s estimation, it soon becomes apparent how absolutely impossible it is to undo the lie (cf. Montaigne, 1998, p. 23).

Bok evaluates the lie in politics in an extraordinary morally stringent way since she claims that wherever there is a special risk of lies becoming routine towards the general public, as is undeniably the case in politics, special protective measures should be insisted on against any form of lying (Bok, 1980, p. 219).

A further approach along these lines is taken by Eberhard Schockenhoff in his book *Zur Lüge verdammt?: Politik, Medien, Medizin, Justiz, Wissenschaft und die Ethik der Wahrheit*. Against the “compulsion to the lie” as claimed, for instance, by evolutionary biology or behavioural science, Schockenhoff cherishes the claim to absoluteness of the ethical norm of the truth. According to Schockenhoff, the thesis of the omnipresence of the lie, or any documented violation of an ethical norm, should not on any account be an argument for its relativisation or even its abolishment. Schockenhoff argues that it is impossible to equate the thesis of the ubiquitous lie in the animal world with the phenomenon which is described as a lie in humans in a moral sense (Schockenhoff, 2000, pp. 32 ff.). A further example for a position that asserts the exceedingly important role of truthfulness for every man in general with simultaneous application of this issue to politics, namely in concrete terms to kings and princes, can be found in Aegidius Romanus’ *De Regimine Principum*. Truthfulness is, in the view of Aegidius Romanus, the avoidance of understatements and overstatements in regard to their own person whereby he views it as more important to avoid the latter since, in his estimation, all humans tend by nature to overstate things for reasons of self-protection and to overvalue
themselves. Aegidius Romanus demands that the merciful and gracious king must not be either boastful or its opposite, but that he has to be truthful and sincere. The most important virtue for a good king nevertheless is, in Aegidius Romanus’ judgement, that he never promises something he is not able to realise afterwards (Romanus, 1968, pp. 80–82).

5. Lying as a skill

I defend the thesis that lying demands several extremely sophisticated skills of the liar who needs a gapless memory, a vivid faculty of imagination, empathy and, beyond that, the right balance in the use of these skills. The successful liar is a brilliant performer because, by plying his art under the guise of truthfulness, he needs to have his utterance accepted as the truth in the end. While lying, the liar has to manage the contradiction between his conviction and fiction. Balancing this is quite a demanding skill that not everybody has. The sovereign liar has to control the risk that he may come to believe in the end in his own lies since in that case the freedom, gained initially with the lie, would be lost. In such a case, the original witting lie changes gradually and unnoticeably into unwitting self-deception. Michel de Montaigne claims that in his view it is said with good reason that whoever cannot trust his memory completely should be vigilant against lying (Montaigne, 1998, p. 23). In Pierre Corneille’s comedy *The Liar*, there is another allusion to the urgent need for an excellent memory when lying, namely in the 5th scene of the 4th act when Klito remarks towards Dorant: “It takes a good memory to keep up a lie” [Translated from the German] (Corneille, 1954, p. 49). The notion that lying is a skill is also defended by Friedrich Nietzsche in the 54th aphorism of *Human, All Too Human I* with the title *The lie*. Being surprised about his observation that people lie so rarely, Nietzsche’s question in the aphorism is the following one: “Why do almost all people tell the truth in ordinary everyday life?” (Nietzsche, 1996, p. 4). To this Nietzsche gives the following answer:

Certainly not because a god has forbidden them to lie. The reason is, firstly because it is easier; for lying demands invention, dissimulation and a good memory (which is why Swift says that he who tells a lie seldom realizes what a heavy burden he has assumed; for, in order to maintain a lie, he has to invent twenty more) (Nietzsche, 1996, p. 40).

These lines show Nietzsche’s explanation of why people lie so infrequently: the reason is not a moral one but the fact that the liar has to take a quite oppressive onus that claims sophisticated competences from the liar if he wants
to lie successfully. According to this perspective, telling the truth is the more convenient, the more indolent way. Nietzsche’s conviction that telling the truth is a manifestation of an intellectual deficit since lying presupposes much more knowledge and skills than sticking to the truth can also be found in the Unpublished Writings. The virtue of truthfulness is interpreted by Nietzsche as a striking intellectual inability as the following fragment in the Unpublished Writings unambiguously shows: “To look at the things coldly in order that they lie there nakedly and without fluff or colour – this is called ‘love of truthfulness’ and is in reality nothing else than the powerlessness to lie” [Translated from the German, fragment 3[241]] (Nietzsche, 1999, p. 82). In addition to his withering criticism of the “love of truthfulness”, Nietzsche repeatedly clarifies his point of view that he does not consider it scandalous that people lie. In Nietzsche’s estimation, the real éclat is the fact that people often and to a great degree do not have the skill to lie, or, that is to say, that they are lousy liars. Nietzsche is disturbed by the fact that there are so many ungraceful lies that are so easily unmasked as a lie as he writes in the 183rd Aphorism of Beyond Good and Evil: “I’m not upset because you lied to me, I’m upset because I don’t believe you anymore” (Nietzsche, 2002, p. 74). In this context, it must be specified that in Nietzsche confidence would not be unsettled, not in the least, if the liar had lied competently without being brought to light. With this perspective of lying as a competence Nietzsche hints, in my view, unequivocally at Plato’s dialog Lesser Hippias that is perhaps the earliest philosophical text about the phenomenon of the lie. In my opinion, in the Unpublished Writings there is a fragment which alludes directly to Plato’s dialog Lesser Hippias: “To lie wittingly and intentionally is more worthy than telling the truth involuntarily – in this respect Plato is right” [Translated from the German, fragment 26[152]] (Nietzsche, 1999, p. 189).

In the dialog Lesser Hippias Plato formulates a defence of the liar, evaluating the liar as someone who can decide independently between truth and falsehood. Starting from the question, which man of Homer’s poems is the better man, Achilles, who is true and simple, or Odysseus, who is wily and false, Socrates demonstrates that the ability to lie requires not merely saying something wrong, but – and this is the essential point of being capable of lying – to know that an argument or something else is false. The ability to lie presupposes indispensably
the knowledge of the truth and a huge number of intellectual qualities. An ignorant man, who is unable to speak falsely except by accident, is, as shown by Socrates, unable to be a liar. Only the competent man, the good man is able to lie because he knows the truth and, in addition to it, exclusively such type of man has the skill to lie voluntarily. Arguing with Hippias about the question which kind of liar is the best, the man who deliberately contrives a lie, or the man who lies unwittingly, Socrates achieves the alarming moral conclusion that the man who is willingly bad or who lies knowingly is superior to one who lies inadvertently. By implication, Odysseus is assessed to be better than Achilles. Socrates demonstrates that only the good man can voluntarily do wrong, and that it is better to do wrong voluntarily and intentionally, willingly and wittingly, than involuntarily and unintentionally, since voluntary falseness is based on intellectual competences. In a moral sense, the conclusion of Socrates' induction that the more competent soul must be the better soul is without any doubt problematic: “So the one who voluntarily misses the mark and does what is shameful and unjust, Hippias – that is, if there is such a person – would be no other than the good man” (376 b7- 376 b9) (Plato, 1997b, p. 936). In this regard, it is important to note that in Plato the “if” is obviously important. In Plato, the postulation that only good men can err willingly does not imply that there are good men who err willingly. But if it is false that good men voluntarily err, then the argument implies that no man voluntarily errs. This, however, would lead to the conclusion worthy of discussion that all wrong doing is involuntary and that the evil is something determined. This conclusion can be found in the Laws IX:

Athenian: ... all wicked men are, in all respects, unwillingly wicked. This being so, my next argument necessarily follows.

Clinias: What argument?

Athenian: That the unjust man is doubtless wicked; but that the wicked man is in that state only against his will ... (860 d0 – 860 d4)(Plato, 1997c, p. 1518).

Another passage in Plato which talks about the issue of the involuntary misconduct is in the Laws V: “... the first thing to realize here is that every unjust man is unjust against his will. No man on earth would ever deliberately embrace any of the supreme evils, least of all in the most precious parts of himself – and as we said, the truth is that the most precious part of every man is his soul” (731 c2 – 731 c5) (Plato, 1997c, p. 1414). In the Timaeus it is written: “But it is not right
to reproach people for them, for no one is willfully evil” (86 d6 – 86 eo) (Plato, 1997, p. 1286). Plato's conclusion of the involuntary evil was picked up on by several authors within the debate about determinism and indeterminism (cf. Rosenberger, 2006, p. 26), especially in the discussion about the appropriate application of the criminal law (cf. Seebaß, 2007, pp. 68, 146). Plato pointed out the criminal significance not only of the criterion of the lack of will, but also of the knowledge criterion (cf. Seebaß, 2007, p. 158).

6. Conclusion

Should the lie be considered an imperative essential skill in the current political climate? I think lying is something that shouldn’t be morally rejected without exception. Politicians should be allowed to lie in order to achieve good political objectives, such as to save lives. An effective and responsible politician who intends to do the best for the people has the duty to reflect carefully on how to act to protect citizens from hazardous situations. In some cases, lying can be the only possible way to avoid a devastating disaster. I consider the rigid or blind application of moral principles immensely dangerous and also an easier and more convenient way to go in comparison with adopting an unremitting and critical attitude that scrutinises every moral rule. In my view, lying is too often hastily condemned as being morally wrong by society. There are certain situations in which telling the truth is, in my opinion, the more immoral way. This is true for personal life as well as for political life. With regard to the lie in politics, the interesting aspect is that although society in general knows that all politicians lie, politicians themselves do not speak freely about this fact. They must not confess to the fact that lying is something that is in certain circumstances the only right way. The people don’t want to hear that their favourite politician is a liar although they know it in reality. Paradoxically, they want the act of lying to be concealed and, on top of that, they do not wish to recognise that the act of concealing is kept from them. Ultimately, people are practicing a kind of pseudo self-deception. If a politician violates morality in order to avoid a big disaster, the truth is that everybody expects that from him. The act of crossing the moral boundary is, without any doubt, a move that can be interpreted as something for which the

13 In this context, I want to refer back to Friedrich Nietzsche's 183rd Aphorism of *Beyond Good and Evil* that I have quoted in a previous passage of this paper.
politician can be accused, but the result of the transgression should basically excuse him, if the lie results in saving the lives of many people. Nevertheless, in most cases reality is different. Even if the result arouses enthusiasm within the public sphere, the politician has to disguise just the same the act of crossing the moral borders. In this context, I would like to refer to Niklas Luhmann who in his reflections about scandals, cynically comments that if a politician is caught engaging in misconduct, he will be sacrificed in order that everything else can continue its usual run unchanged (Luhmann, 1993, p. 39). However, the truth of the matter is that there are paradoxical situations in which the violation of morality is indeed more moral than following the moral rules. In other words, being moral is more immoral than to act immorally in the sense of violating the moral rules on one occasion to achieve a certain result. So, I think that the person who enters political life has to acquire the ability to discern situations where it is more moral to act immorally. Machiavelli’s prominent demand on the prince to learn “… how to do wrong, and to make use of it or not according to necessity …” (Machiavelli, 1993, p. 117) could be interpreted as implying that there are certain situations where it is absolutely important for the prince to cross the moral borders to prevent big disasters.

To conclude, I defend the thesis that acting immorally can, in certain situations, be more moral if violating moral standards can save a life. In my view, Michaelis or Kant were not right on the point that the obligation to truthfulness applies regardless of all consequences. With reference to this, Dieter Bonhoeffer describes in Was heisst die Wahrheit sagen? Immanuel Kant as a cynic who exhibits only a dead idol of the truth, claiming that the imperative to tell the truth pertains without exception everywhere, anytime and in relation to anyone in the same way. According to Bonhoeffer, Kant destroys the existence of the vivid truth between the humans, pursuing a truth-fanatism that cannot consider any form of human weakness (Bonhoeffer, 1948, p. 285).

Given its origin, lying is a morally neutral skill that can be used to realise something good as well as to create horrible evil. The moral direction of lying is absolutely dependent on the responsibility of the agent. A lie utilised for the purpose of saving a life should be morally justified in any case.
References


ARGUMENTATIVE HYPOCRISY AND CONSTITUENT DEBATES: THE ITALIAN CASE

HIPOCRISIA ARGUMENTATIVA E DEBATES CONSTITUINTESE: O CASO ITALIANO

Giovanni Damele*
giovanni.damele@fcs.h.unl.pt

Abstract. Jon Elster suggested that even speakers who are not moved “by a concern for the common good”, but whose concerns are “purely self-interested”, may be still forced or induced “to substitute the language of impartial argument for the language of self-interest”. This substitution would be the fruit of the civilizing force of hypocrisy. This argumentative hypocrisy is a key concept for understanding a process of negotiation through persuasive strategies typical in constitutional debates. Particularly, Elster believes that “the most important requirement” of a bargaining theory should be “that we are able to specify what will happen during a temporary breakdown of cooperation”. The constituents can get out of an impasse caused by a non-cooperative situation resorting to argumentative hypocrisy. The paper will analyse some examples taken from the debate which led to the final version of the Italian Constitution.

Keywords: Argumentation theory, Political argumentation, Rhetoric, Constitutional law, Elster.

Sumário. Jon Elster tem sugerido que até os falantes que não são movidos “por uma preocupação com o bem comum”, e cujos interesses são puramente egoístas, podem ser obrigados ou induzidos a substituir a linguagem do egoísmo pela linguagem do argumento imparcial. Esta substituição seria o fruto da força civilizadora da hipocrisia. Esta hipocrisia argumentativa representa um conceito-chave para interpretar um processo de negociação por meio de estratégias persuasivas, típico dos debates constituintes. Em particular, Elster defende que o requisito mais relevante de uma teoria da negociação deveria ser o facto de ter a capacidade de explicar o que acontece no caso de uma suspensão temporária da cooperação. A hipocrisia argumentativa pode permitir aos constituintes ultrapassar um impasse. Este artigo irá analisar alguns exemplos retirados dos debates que conduziram à versão definitiva da Constituição Italiana.

Palavras-chave: Teoria da argumentação, Argumentação política, Retórica, Direito constitucional, Elster.

* Professor Auxiliar Convidado, Faculdade de Ciências Sociais e Humanas, Universidade Nova de Lisboa.
0. Introduction

In his *Arguing and Bargaining*, Jon Elster applies his argument from the *civilizing force of hypocrisy* to the analysis of two historical examples of constitution-making processes: the Philadelphia Convention of 1787 and the French revolutionary *Assemblée Constituante* (1789-1791) (Elster, 2000). As a source of impartiality, even if through a kind of deliberative *heterogenesis of ends*, the *civilizing force of hypocrisy* works as a powerful mechanism of mediation. In sum, even if used hypocritically, impartial arguments can lead to agreements based on general interest.

In this paper, I will briefly describe Elster’s approach, underlining mainly its argumentative dimension. I will focus the analysis of the second paragraph on the double parallelism between, on the one hand, self-interested and impartial argumentation, and, on the other, “threat-based” and “warning-based” bargaining. In the third and the fourth paragraph I will present some examples, taken from the *travaux preparatoires* of the Italian Constituent Assembly (*Assemblea Costituente*, 1946/48), applying Elster’s model to them. I will also address the main issue of the choice between publicity and secrecy of debates. In the fifth, I will try to show correspondences and differences between Elster’s model and a debate in a real context, in particular from the point of view of the rhetorical dimension.

1. Elster on the “civilizing force of hypocrisy”

According to Jon Elster, the process of constitution-making can be described resorting to two types of – in his words – “speech acts”: arguing and bargaining. The first type, based on “rational argumentation” is subject to “criteria of validity”. The second, based on “threats’ and promises”, is subject to “criteria of credibility”. Both types together form “the main vehicles by which the parties seek to reach agreement” (Elster, 2000, p. 372).

Considering the “type” of arguments used in constitutional debates, Elster makes a broad reference to two general categories: consequentialist or deontological arguments. In the first case, the framers “appeal to overall efficiency”, while in the second “to individual rights”. This second type, that of
“right-based arguments”, seems to be more “impartial”, because “the rights are assigned to everybody” (Elster, 2000, p. 378).

Elster notices also that “consequentialist” arguments can imply a certain amount of impartiality. However, the use of deontological arguments seems to be much more effective: “framers can go to great lengths to make it appear that a measure whose real justification is obviously utilitarian can also be defended in terms of rights” (Elster, 2000, p. 391).

Taking into consideration the debates developed in two constituent processes – the Federal Convention in Philadelphia and the Assemblée Constituante in Paris from 1789 to 1791, Elster insists mainly on the role of rational argument in constituent decision-making, claiming that even the actors with “purely self-interested” concerns may be “forced or induced to substitute the language of impartial argument for the language of self-interest” – a kind of “substitution” that Elster attributes to what he calls “the civilizing force of hypocrisy” (Elster, 2000, p. 349).

This consideration leads Elster to conclude that the two original Habermasian commitments to truth and impartiality can coexist with a strategic use of impartiality and sincerity. Even if the framers are not genuinely committed to these values, “they may find it in their interest to appear to be so committed”. In this case, “they engage (...) in strategic uses of purportedly non-strategic argument”. In other words, self-interested actors may try to “ground their claims in principle” when “their self-interest tells them to appeal to an impartial equivalent of self-interest”. This strategic function of “citing a general reason” has, according to Elster, an obviously persuasive goal, which is to increase the ability of the speaker to persuade others, and particularly, to persuade “the neutrals to agree” with him (Elster, 2000, pp. 405-409).

Finally, a kind of “substitution”, analogous to that established in an argumentative context by the “civilizing force of hypocrisy”, can also be found in a bargaining context. In the same way in which they may substitute an impartial argument for a direct statement of their interest, strategic actors “may also find it useful to substitute truth claims for credibility claims”. In this case, “instead of making a threat whose efficacy depends on its perceived credibility, they may utter a warning that serves the same purpose and avoids the difficulties
associated with threats”. Indeed, according to Elster “warnings” are “factual equivalents” of a threat. Threats, “are statements about what the speaker will do”, while warnings are statements “about what will (or may) happen, independently of any actions taken by the speaker.” Thus, this substitution seems to be motivated by a strategic preference for (more persuasive) claims based on truth instead of (possibly less persuasive) claims based on credibility (Elster, 2000, p. 415). The “substitution” seems to be based on the rhetorical dimension of verisimilitude. A reference to (supposedly true) factual consequence is used in order to increase the persuasiveness of a mere appeal to (subjective) credibility.

Interestingly enough, this passage, which constitutes the argumentative side of the “civilizing force of hypocrisy” and which takes us from a sincere commitment to truth to a strategic sincerity, partly coincides with the Aristotelian definition of rhetorical discourse. In rhetoric, the possibility of the discourse being true is taken into consideration only as a means of persuasion: independently of the truthfulness of the premise or of the sincere commitment of the speaker to truth. What matters is the appearance of truth (Viano, 1955, p. 284): the persuasion of the audience about the truth. In other words, the civilizing force of hypocrisy is, in its essence, based on a rhetorical move.

If, as according to Elster, rational discussion is supposed to be based on “the power of the better argument”, then constitutional bargaining “rests on resources that can be used to make threats (and promises) credible” (Elster, 2000, p. 392). Referring to these resources, from the viewpoint of the distinction between “the role of rational argument” and “threat-based bargaining in collective decision making”, Elster explains that the paper focuses on “bargaining on the basis of extra-parliamentary resources, such as manpower and money” (Elster, 2000, p. 348). In other contexts he also makes reference to “extra-political” or “extra-constitutional” resources.

2. The Italian Constituent Assembly

---

1 Whether these substitutions refer to a strategic use of argumentation or to a deceptive use of warnings is a question that – although relevant – is beyond the purposes of this paper. Both are examples, according to Elster, of “impartial equivalent of self-interest”. As far as the “impartial equivalent” is concerned, the parallelism is relevant from my viewpoint. Although there can be clear differences between the two strategies.
According to Elster, every constitution reflects not only the relative strength of the contending political forces within the Assembly, but also the political climate.

The Italian Constituent Assembly (Assemblea Costituente) was elected in 1946 (using proportional representation) with the first free and full elections in Italian history. The result of the election set up a balance between the main anti-fascist parties, the centrist Christian Democrats (DC) on one side (which won 35 percent of the vote) and the two leftist main parties on the other (the socialists of the PSI, with 21 percent of the vote, and the communists of the PCI, with 19 percent). A third political position, the liberals, was also very influential even if numerically less significant (less than 10 percent, adding together all the liberal parties).

As for the political climate, the work of the Assembly was threatened by external and internal circumstances. Internally, there were the obvious tensions inherent in the anti-fascist parties belonging to different political traditions: the Christian democrats, the Social-Communists, and the Liberals. Externally, there was the risk of a political crisis leading to civil war and, eventually, to foreign intervention. However, the main external influences affecting the work of the Assembly were the relations between political parties in view of the political elections of 1948. In particular, in May 1947 the Prime Minister (Presidente del Consiglio) Alcide De Gasperi, general secretary of the Christian Democratic Party, drove the communists out of the government with the aim of forming a new government, without the extreme left and obtaining the full support of the government of the United States. Italy could have followed the path of Greece in 1946-1947, with a bloody civil war and foreign intervention (Foot, 2003). However, De Gasperi and the general secretary of the Communist Party, Palmiro Togliatti, continued to collaborate even after May 1947, to draw up the Constitution.

The work of the Italian Constituent Assembly lasted 18 months, from 1946 to 1947 (with over 170 sittings).

As Elster noticed, a fundamental aspect of the institution of a constituent assembly concerns its internal procedural rules. The following issues can arise: a) it must decide whether to proceed in closed sessions or open the debates to the
public; b) it may decide to create one or several subcommittees to prepare a draft of the constitution or to discuss special issues (Elster, 2000, p. 367).

Regarding this problem, a main concern expressed by the communist trade union leader Giuseppe Di Vittorio was that the secrecy of the sessions of the constitutional commission would have finally excluded “the people” from the drafting of the new Constitution (Pombeni, 2016, p. 112).² As a matter of fact, the Italian solution was a mixed one: the sessions of the Assembly were public, but those of the commissions were closed. Indeed, in order to organize the work of the Assembly, the task of drafting constitutional provisions was delegated to a “Constitutional Commission” of 75 deputies (also known as “the Commission of the 75”), divided into three sub-commissions, each one chaired by a deputy of one of the three main parties³. The first sub-commission, tasked with drafting constitutional articles on the “Rights and Obligations of the Citizens”, was chaired by the Christian Democrat deputy Umberto Tupini. The second and the third, on “Constitutional Organization of the State” and “Economic and Social Relationships”, were chaired by the communist Umberto Terracini and the socialist Gustavo Ghidini, respectively. Finally, a Committee of 18 deputies (the “Committee of the 18”) was given the task of writing an overall draft of the constitution, in accordance with the work of the three sub-commissions.

Working in subcommittees made it possible to have preliminary votes that did not commit the delegates to premature decisions. From this viewpoint, Elster quotes Jean-Joseph Mounier, leader of the moderates in the Assemblée Constituante: according to him, the committees favour “cool reason and experience,” by detaching the members from everything that could stimulate their vanity and fear of disapproval. Particularly, in the case of closed sessions (such as in the Federal Convention) there is little risk of being prematurely locked into one opinion (Elster, 2000, p. 411). According to Elster, at the Federal

² According to Di Vittorio, the “most negative aspect” of the work method used in the Constituent Assembly were precisely the “small committees” which tended to exclude the “large public” from the debates and, above all, from the decision making process. It is necessary to note that the secretary of the Italian Communist Party took part in the debate in the commission and in the relative sub-commissions, being one of protagonists of the constituent process - one of the few leaders who was directly involved in the debates and who was among the responsible for the organization of the Assembly. From this point of view, he acted in a pragmatic way, often driving his parliamentary group away from (potentially controversial) positions of principle or radical stances and towards both pragmatic positions and compromise solutions (Pombeni, 2016, p. 112).

³ “The next aspect of the constitution of the Constituent Assembly concerns its internal procedural rules. The following issues can arise ... It may decide to create one or several sub-committees to prepare a draft of the constitution or to discuss special issues” (Elster, 2000, p. 367).
Convention the fear was that the pride and vanity of the delegates, as well as pressure from their constituencies, might have prevented them from backing down from an opinion once they had expressed it (Elster, 2000, p. 386). Another interesting effect of secrecy – Elster notes – is that of pushing the debates away from argument and towards bargaining (Elster, 2000, p. 386). Conversely, in public sessions (such as in the French Assemblée Constituante) there is stronger pressure against the expression of self-interest. A public setting encourages the use of pre-commitment through principle, with the larger audience serving as a resonance board for the claim and making it more difficult to back down (Elster, 2000, p. 410). According to Elster,

this difference between the two assemblies is reflected in the debates. Many of the debates at the Federal Convention were of high quality: remarkably free from cant and remarkably grounded in rational argument. By contrast, the discussions in the Assemblée Constituante were heavily tainted by rhetoric, demagoguery and overbidding. At the same time, the Convention was also a place where many hard bargains were driven, notably the deal between the slave holding and the commercial states (Elster, 2000, p. 411).

According to Elster, a positive consequence of secrecy also allows the possibility for a participant in the debate to change his or her mind. This can help reach a more cooperative and reasonable context, in which every participant is not strongly committed to coherence due to the external pressure of public opinion. As Elster noted, “the pursuit of understanding” can also be “constrained by a commitment to truthfulness or sincerity” and, in a habermasian fashion, “the outward form of truthfulness is consistency”. However, while “a participant who is seen as choosing normative arguments à la carte ... will often be viewed as insincere”, people “often modify their views by exposure to an argument.” Elster notes that “framers in both assemblies believed that a major virtue of rational deliberation was that of allowing this to happen” (Elster, 2000, p. 377). We may also say that a (not necessarily genuine) change of mind can be instrumental in reaching an agreement on fundamental issues. During the debates of the Commission of the 75 a general problem concerning the framework of the Constitution was presented with the so called Bozzi Motion [Ordine del giorno Bozzi], subscribed by, among the others, Giovanni Lombardi (PSI), Mario Cevolotto (Liberal), Amintore Fanfani (DC), Tomaso Perassi (Republican), Piero

---

4 “Nevertheless, genuine changes of mind can often be distinguished from opportunism. Explicit disavowal of one’s earlier views, and attempts to remedy earlier decisions, would be one criterion. Claiming to be persuaded by normative arguments that are counter to one’s self-interest would be another. Although neither criterion is infallible, both can be helpful” (Elster, p. 2000, p. 377).
Calamandrei (Liberal-socialist) and Francesco M. Dominedò (DC). This motion committed the framers to write a Constitution as far as possible simple and clear, so that it would be comprehensible “for the entire people”, and limited to “essential norms” and “concrete dispositions”, and to abstain from stating “general programs” and political “desires”. Key leaders both of the Christian democrats (Giuseppe Dossetti and Costantino Mortati) and of the Communist Party (Togliatti) criticized the motion as “equivocal”, defending the programmatic character of the Constitution (Togliatti) and the freedom of the framers (Mortati). At this point, Fanfani (one of the main subscribers) withdrew from his initial position and partially agreed with Togliatti on the programmatic nature of a Constitution: a program which was not – in his words – a mere “hope”, but a clear “will”. Interestingly enough, this change of mind led to a major compromise on the motion, based on an ambiguous reformulation. The Motion was ultimately approved, but a vote declaration of Tupini opened the door to a more flexible interpretation of the text (Pombari, 2016, pp. 206-208). Tupini’s vote declaration can also be considered a kind of hypocritical commitment and an “incompletely theorized agreement” (Sunstein, 2007), something that originates from conflicting opinions, but which nevertheless concludes with participants agreeing on a single outcome, but for different reasons (Mansbridge et al., 2010). We will come back to this issue infra. What is relevant to show here is that Fanfani’s (at least) partial change of mind would have been probably if not impossible, at least very difficult in a public plenary assembly.

Generally speaking, it is possible to say that, in the Italian case, the joint process of the (closed) committees and of the (public) assembly favoured the drafting of a Constitution largely based on a set of political compromises. The main example of such a compromise is Article 7, which established a privileged status for the Church in the new Republic, including (with the approval of the communists) the Concordat between the Church and the fascist State, included in the “Lateran Pacts” (Patti Lateranensi) of 1929. Other compromises ended with the independence of the Constitutional Court – a defeat for the Communists – or the Christian Democratic retreat from a corporativist and regionalist model, which can be viewed as a victory of the liberals. One of the main Christian

---

5 The Motion had been presented by the liberal MP Aldo Bozzi on October 25, 1946.
Democratic defeats came when the final constitutional text no longer proclaimed the indissolubility of marriage (Einaudi, 1948, pp. 662-663). Thus, it may be interesting to take into consideration some examples of such “compromises” and compare them with Elster’s framework.

3. The debate in the sub-commissions

Analysing the debate in the commissions, we find many characteristics included in Elster’s model. First of all, we find a bargain conducted resorting to internal and external resources. Among the internal resources, we may find the use of logrolling or vote-trading, which is the practice of exchanging favours by reciprocal voting for each other’s proposed legislation.

During the debate concerning the inclusion in the Constitution of the indissolubility of matrimony (strongly endorsed by the Christian democrats and equally strongly opposed by the left wing parties and some of the liberals) a first compromise was ultimately found, including in the provisional draft a definition of family as a “natural society”, and dropping any reference to the “indissolubility” of marriage. Also in this case, the agreement was reached by Togliatti (PCI) and Aldo Moro (DC) on the basis of a (strategically) ambiguous reformulation of the text. The expression “natural society”, indeed, allows for multiple interpretations, and the agreement was clearly on the level of the wording, and not on that of the (different) possible interpretations.

When the text was submitted in the Commission of 75 for approval, Togliatti, following the previous agreement, gave instructions to the communist deputies about how to vote. According to this logrolling strategy, the Communists allocated their votes in a way that ultimately permitted the approval of the expression “natural society”.

The debate on the issue of the definition of marriage was thus made difficult by the strong resistance of the Catholic MPs (Giorgio La Pira and Dossetti, particularly) who sensed that this issue was absolutely crucial. During the debate, the direct threat of withdrawing from the commission was substituted by the warning that it could cause a secession in the commission. The Socialist Lelio Basso, in particular, warned the Christian Democrat MPs against causing a secession with what he interpreted as stubborn behaviour. A kind of substitution
of a threat with a warning, based, in this case, on “extra-parliamentary resources”\(^6\).

However, these strategies where ultimately ineffective, and the final draft included both the definition of family as “natural society” and the indissolubility of marriage. The issue was finally solved by resorting to a strategic use of the rules of the Assembly. In the General Assembly, the last passage before the final approval of the article was the vote of the amendment presented by the socialist MP Umberto Grilli, who proposed again to eliminate the term “indissolubility”. Before voting, twenty deputies called for a secret ballot. This was quite an unusual demand, based however on the rules of the pre-fascist Chamber of Deputies, which the Constituent Assembly decided to adopt. Thanks to this strategic use of a pre-existing institutional framework, which reduced the influence of the parties (and also of public opinion) over the deputies, the indissolubility of marriage was not included in the Constitution.

Considering the argument used, every party tried to use “impartial” arguments, even if of a different kind. We know that every party had strongly “egoistic” reasons, mainly related to the imminent electoral campaign. However, they substituted the language of impartial argument for the language of self-interest.

The debate on article 29 and the possibility of introducing the indissolubility of marriage also provides a good example from this point of view. Even if the Christian democrats MPs were convinced that this position was not simply related to Catholic doctrine, but was ultimately based on the very “nature” of marriage, they also felt that it should have been justified on the basis of “universalizable” arguments. For this reason, they made it clear that they were referring to marriage not as a sacrament (which was already recognized by the church as indissoluble), but as a legal institution. Thus, they presented philosophical, consequentialist (based on “the unity of family”), and even “scientific” reasons for defending its indissolubility. La Pira used a kind of “tu quoque” argument, suggesting that the communist’s opposition to the

\(^6\) Making reference to “extra-constitutional bargaining”, Elster notes that “one party may act on the assumption that the other will be unwilling to be seen as responsible for breaking off negotiations” and he gives a similar example: “during the debates over the Spanish constitution in 1978, the Union of the Democratic Center was accused of breaking a painstakingly negotiated set of compromises, leading to the withdrawal of the Socialist member on the subcommittee” (Elster, 2000, p. 394, and n. 223).
Argumentative hypocrisy and constituent debates: the Italian case

Indissolubility of marriage was ultimately contradictory, considering that divorce was practically banned in the Soviet Union (Damele, 2016).

This type of argument was somewhat common in the Assembly at the time. It was also used in its version of a kind of “impartial” use of the argument from authority, based on the appeal to someone who is mainly considered as a political authority by the adversary. As Mario Einaudi noticed, “the Communists called in the authority of George Washington and relied on Benjamin Franklin to weaken the argument for an upper chamber, while the Christian Democrats quoted at length the authority of Stalin to support the thesis that the two chambers had to be of equal power” (Einaudi, 1948, p. 662). This kind of use of the argument from authority can also be considered hypocritical, since it makes reference to somebody who, under normal circumstances, would not be considered a political or moral reference point by the speaker. In this case, the hypocritical nature of the argumentation is quite explicit. On the other hand, the appeal to someone who is regarded as a political authority by the adversary, increases the appearance of impartiality, for persuasive purposes. From Elster’s viewpoint, the speaker is not necessarily impartial: he exploits impartiality for self-interested purposes. Nevertheless, Elster thinks that even this self-interested use of impartiality can have a “civilizing” force, due to the fact that, in a sense, it speaks (even if for egoistic purposes) the language of universalizable arguments.

Other examples show a strategic use of a hypocritical discourse, with the purpose of persuading different audiences. It can be interesting, from this viewpoint, to consider the different strategies used by Christian democrats MPs in order to justify their work in the Constituent Assembly towards their electorate and the Catholic hierarchy, and compare them with the persuasive techniques used by the same politicians within the commissions, in order to reach agreement.

There was strong pressure, coming from the Vatican and the Catholic hierarchies, particularly on issues such as family, education and the role of the Catholic Church. In many cases, Catholic deputies tried to defend Catholic positions resorting to impartial arguments. Then, when this strategy was effective, they presented it as a victory for a Catholic point of view. This strategy was especially common during debates on social issues, when Christian Democrat
deputies invited Communist and Socialist deputies to leave aside terminological questions, and consider the essence of the problem, in order to reach a solution. However, when they had to defend their decision on articles published in the official press of the Catholic Church or of the Christian Democratic party, they presented that same solutions as purely based on the social doctrine of the Catholic church.

Discussing the approval of the whole Constitution, in the Christian Democrat newspaper “Il Popolo”, La Pira (one of the leaders of the “left wing” of the Christian democrat party) stated that “it is possible to say, without errors, that the organic conception, typical of the Christian sociology … it is on the very base of this Constitution.” A kind of argument used with the aim of persuading the Catholic hierarchy about the acceptability (from their own viewpoint) of the Constitution. However, during the meetings of the Commission of the 75, La Pira, together with other Christian Democrat leading figures such as Dossetti and Moro, resorted to a different strategy, based on a universalization of the argument, with the aim of showing that, under the surface of different ideological and religious viewpoints, there was a deep common ground. Reference was made, for example, to a general evolution of the contemporary thinking, which resulted in a wide convergence between different ideological positions, on fundamental issues, such as, for example, labor law. In other words, on the one hand, the strategy was that of persuading the electorate and Catholic hierarchies (suspicious of the socialist and communist influence in the Constituent Assembly) that they were succeeding in drafting a constitution which was “Christian” in its fundamental inspiration. On the other hand, it was that of persuading political adversaries that, under different expressions, they all shared the same fundamental ideas. The objective of this double strategy was, fundamentally, a legitimation of the compromise, now represented as a strategic move necessary to obtain political success on issues that were instrumental from the point of view of Catholic social teaching (Pombeni, 2016, p. 117). Such a strategy can be considered, from Elster’s point of view, a good example of the civilizing force of hypocrisy: the necessity to reach an agreement force to “universalize” their own point of view (that is, to give an “impartial” version of a

---

7 Session of December 23, 1947
“partial” or “egoistic” point of view), only to come back to the original “partial” argument when the strategy is that of persuading the electorate or also, in the case of Christian Democrat MPs, the Catholic hierarchy. Clearly, in this second case, a kind of strategy like that used by the Commission, based on the idea of a deep common ground between different ideological points of view, would have been less effective, due to the position of the Catholic church (quite suspicious regarding “compromises” with the social-communists) or the polarization of the electorate.

4. Conclusions

The choice made by the constituent fathers to organize the Constituent Assembly’s work into two sessions – closed drafting commissions and a public plenary assembly – favored compromises, allowing, according to Elster’s scheme, shifts from arguing to bargaining and (in relevant cases) a change of mind or withdrawal. Many of the strategies described by Elster as ways to (tacitly or expressly) reach agreement through negotiation can be found, such as logrolling or the strategic use of procedural rules in general. From a discursive point of view, examples of substitution both of impartial arguments for the language of self-interest and warnings for direct threats can be found. Particularly relevant, from a rhetorical point of view, is the strategy used by the Christian Democrat MPs of using “partial” arguments while addressing their electorate or the Catholic hierarchies and “impartial” arguments in order to persuade their political adversaries – a clear example of adaptation of the persuasive strategy to the (particular) audience.

Many of these strategies can represent, from Elster’s point of view, examples of the “civilizing force of hypocrisy”, in as much as they imply a kind of universalization (mainly in the case of the substitution of the “impartial” for “partial” arguments) and the substitution of a subjective threat with an objective warning. Both strategies are, in the examples taken into consideration, instrumental to negotiation.

It has become common to refer to the Italian Constitution as an example of “high-profile agreement”, different from the low profile compromises of the “politique politicienne”, which resulted in a process of “higher law making” (in Bruce Ackermann’s words). This is certainly true, due in part to the general
context and to the sense of responsibility of the “constituent fathers” (or at least of their leaders). First of all, the constituent fathers were well-aware of being in a kind of “veil of ignorance” situation. A crucial problem was that of the future electoral weight of each political party, something that suggested a solution based on a strong system of checks and balances, with the purpose of preventing the exploitation of a strong position, transforming an electoral advantage into a powerful domination of the whole system. Secondly, in a (highly) pluralistic political landscape, compromises are not only inevitable, but the desired result of the political (in this case constituent) process. Arguing, bargaining through threats and promises, are fundamental strategies instrumental to reaching such an agreement. However, it is also important to look, realistically, at the outcomes of such a strategy. We made a reference, supra, to Sunstein’s concept of “incompletely theorized agreement. This peculiar method of statute-making, aimed at reducing the potential for conflict, is based, according to Sunstein, on an agreement on abstractions, to which an agreement on the particular meaning of those abstractions does not correspond.

Thus, ambiguity is used strategically to foster agreement on abstractions without limiting specific interpretations (Eisenberg, 1984). This technique, which according to Eisenberg corresponds to a “more rhetorical view of communicator as strategist”, is particularly used by collective agents like legislatures. In this context, people “confront multiple situational requirements, develop multiple and often conflicting goals, and respond with communicative strategies which do not always minimize ambiguity, but may nonetheless be effective” (Eisenberg, 1984, pp. 227–238)

As Hans Kelsen has noted, this exploitation of ambiguity can be part of a legislative technique which tends “to leave a large latitude to the authority charged with its application or interpretation” by “intentionally formulating the rules in a vague, indefinite, equivocal way” (Kelsen, 1939, p.11, 17). This choice of vague or indefinite terms seems to make the decision “easier when the interests are strongly opposed”. However, “the real concordance of wills” can be “only an illusion”, which is “very quickly dissipated when it becomes necessary to apply the norm formulated in indefinite terms”. Then, “it is realized that the ... parties
have aimed at very different ends under the cover of the words with which the norm to be applied is expressed”.

References


WHY THE DELIBERATIVE IDEAL JUSTIFIES PUBLICITY – EVEN IF PUBLICITY MAY UNDERMINE DELIBERATION

POR QUE O IDEAL DELIBERATIVO JUSTIFICA A PUBLICIDADE – AINDA QUE A PUBLICIDADE POSSA COMPROMETER A DELIBERAÇÃO

John Pitseys*
john.pitseys@uclouvain.be

Abstract. This paper critical assesses those arguments in democratic theory which defend the role of secrecy in politics in epistemic terms, that is, by emphasizing some of the negative effects that publicity in deliberation can have (namely: obscuring the informational process, favoring plebiscitary rhetorics, and replacing open discussion with conformist behaviors and/or bargaining). Based upon the analysis of the 2010-2011 Belgian Sixth State Reform, the paper argues that, even if publicity can produce negative effects on deliberation, it cannot be proven that these effects are more severe than those produced by a closed-doors deliberation. Furthermore, it argues that an epistemic justification of closed-door deliberation could not be accepted by a reasonable citizen: the justification of publicity does not rely on its epistemic positive effects, but on the fact that the assessment and definition of these cannot be left to the negotiating parties.

Keywords: political publicity, political secrecy, public deliberation.

Sumário. Este artigo avalia, criticamente, os argumentos que, na teoria política, defendem, por razões epistémicas, o papel do secretismo na política. Desta maneira, enfatizamos alguns dos efeitos negativos que a publicidade pode ter na deliberação (designadamente: obscurecendo o processo informacional, favorecendo retóricas plebiscitárias e substituindo a discussão aberta por comportamentos conformistas e/ ou barganha). Com base na análise da Sexta Reforma do Estado Belga de 2010-2011, este artigo defende que embora a publicidade possa produzir efeitos negativos na deliberação, não se pode provar que estes efeitos são mais severos do que aqueles produzidos por uma deliberação à porta fechada. Além disso, argumentamos que uma justificação epistêmica da deliberação à porta fechada não poderia ser aceite por um cidadão razoável: a justificação da publicidade não depende dos seus efeitos epistêmicos positivos mas sim do facto de a avaliação e definição destes não poder ser deixada ao critério das partes em negociação.

* CRISP (Brussels) / Catholic University of Louvain, Belgium.
Democratic theory links political legitimacy to two possible sorts of criteria. First, the political decision must reflect on an equal basis the interests expressed by the social space. Second, the political decision must reflect the common good, which relies upon a rational decision-making process. These two conceptions are subject to contention but converge on the point that publicity is a political virtue that must necessarily be pursued and which is descriptively and normatively opposed to secrecy and lying.

A closer look at the notion of political publicity reveals however a number of difficulties. Publicity can hinder the conditions of rational collective deliberation. Reversely, meeting the conditions of a rational collective deliberation may lead to restricting the exercise of political transparency. Deliberation is supposed to favor an inclusive dialogue between the members of the citizenry. However, there can be some strong divergences on the extent to which the decision-process must be open to citizens, and – consequently – on the nature of the justifying base of a deliberative system.

Political publicity seems thus both a central democratic principle and a mined field for philosophical dilemmas: in this context, to what extent can a closed-door deliberation be legitimate in the framework of a political decision? Must the decision making process be open if it can be proven that it harms the epistemic quality of the deliberation, and thus the average level of fairness of the content of the decision?

I do not deny that publicity can harm the overall quality of political deliberation. I think nevertheless that there is no reason to restrain publicity in the name of public reason. I am not saying that there are no valid normative grounds to limit publicity, and the conclusion of the paper suggests political equality might be the best yardstick on which to base an option for public or closed-door deliberation. But I think that the deliberative argument against publicity is weak in this regard.

In order to structure my point, I will first remind the reasons why publicity can harm the deliberative quality of the political discussion. Here I do not aim
However to show that the practical conditions for public discussion do not match the deliberative ideal - which would be a trivial assertion. Nor will I recall that deliberation and democracy fall under practically and analytically distinct categories. Rather, this paper aims to show that the best way to approach the conditions of a rational deliberation might require restraining political publicity and, reversely, that the arguments against publicity might actually fit with a non-ideal understanding of the deliberative ideal.

The second part of the paper suggests that the possible negative epistemic effects of publicity do not suffice to justify secrecy. From a consequentialist view, I develop the idea that the epistemic arguments against secrecy can all be replicated in the context of a closed-door deliberation. And from an internal normative aspect, I try to demonstrate that the deliberative argument against publicity is not consistent with the normative basis it ought to stand for. Should closed-door deliberation promote the quality of the discussion and respect citizen’s political rights, it contravenes the egalitarian understanding of impartiality deliberative democracy is expected to be grounded on. In this framework, the conclusion of the paper suggests that if secrecy must be rejected when it disregards impartiality, it could be justified if and when it respects this principle.

Finally, the reader will notice that this paper finds a major inspiration from my practical experience as a political adviser during the Belgian sixth State Reform negotiation process that took place between June 2010 and December 2011 – and broke the record for the longest formation of a government. This period was a testing experience. It gave, I think, a rare source of reflection for understanding the complex relationship between representative and deliberative legitimacy, and between ideal and non-ideal theories of legitimacy.

1. Deliberation and Publicity

1.1. Three conceptions of publicity

Three conceptions of publicity may be distinguished in political theory. First, publicity can be understood as the principle of public information (A). Citizens are informed of the practices, arguments and decisions that take place within the political system: publicity designates the process by which political
facts are presented and made available to the public (Gutmann, Thompson, 1999, chap. III). Such a process relates to the access to political information, that is to say the content of the norms that apply to the collectivity, the arguments and positions of the parties, the justification process of the decision (Rawls, 1993: 66-67). It designates also the ownership modalities of the information.

A second conception understands publicity as a collective background for the justification of public discussion (B). This background encompasses the public information onto itself but also the substrate of its collective treatment, that is to say a common space of political understanding and justification (Luban, 1996, pp. 171-172; Manin, 1996, pp. 215-216). In this framework, the conception B both describes and justifies the conditions for the constitution of the public debate. It includes the legal procedures of public disclosure as well as the general access to the content of the political discussion. But most of all it encompasses various institutions that are to cover political information, organize the collective discussion and civilize the citizen’s behaviours: we think here about the informational networks of the civil society, the action of the mass media, the conversations within informal public sphere, etc. The conception B covers both the access to the information and to the public space, and what makes possible the discussion about them.

These two conceptions respond to a practical ideal, as the social and political institutions must meet effectively their requirement of justification. This ideal serves two different possible functions. First, it represents an ideal of public discussion, insofar as the principle of publicity is supposed to enhance the volume, the quality and the responsiveness of the arguments put on the table of the political discussion. Second, it serves an ideal of democratic participation. On the one hand the constitution of a public sphere of information is supposed to connect the exercise of the government with popular sovereignty: the political publicity constitutes both a means of control of the represented on the representatives, and a way for the representative to have access to the issues, perspectives, perceived interests and political pulse of the represented. On the other hand, the principle of publicity contributes to the creation of a common general public space.
Finally, publicity can also be understood as a hypothetical way to test the rightness of the decision (C) (Bohman, 1996, p. 25). The conceptions A and B see publicity as a benchmark to organize the public space and the actual conditions of the political debate. The conception C approaches the meaning of the transcendental principle of publicity as defined by Kant in his *Treaty on Perpetual peace*: “All actions relating to the right of other human beings are wrong if their maxim is incompatible with publicity” (Kant, 1795, p. 130). According to Kant, a policy “which cannot be publicly acknowledged without thereby inevitably arousing the resistance of everyone to my plans, can only have stirred up this necessary and general (hence *a priori* foreseeable) opposition against me because it is itself unjust and thus constitutes a threat to everyone” (Ibidem: 126). Being hypothetical, such a test doesn’t entail any real public but only the putative unveiling of the argument (Ibid.: 131). In the words of Axel Gosseries, “the test is thus hypothetical, not only because it does not necessarily require *actual* publicity (as we shall see), but also because any outcome of actual publicity would inevitably remain at best a rough approximation of what the test would lead us to with an ideal public” (Gosseries, 2005). Following this acceptation, the principle of publicity does not compel the political actor to communicate his reasons to the other parties. Nonetheless he has to universalize these reasons assessing whether they can be accepted by all. A reason which is valid for me cannot be accepted unless it is valid for the others as well: an action which does not pass the test of the universal audience cannot be considered and, consequently, has to be rejected.

In this context, there is to be a converging relationship between the three mentioned definitions of publicity: “The best way to make sure that officials formulate policies that could withstand publicity is by increasing the likelihood that policies will withstand publicity” (Luban, 1996, p. 157). More information is more contextual justification. And more contextual justification paves the way for a higher level of generality.

1.2. The Three Classical Critiques to publicity and The Deliberative Ideal
Political publicity faces three classical critiques. The first one is related to the argument for the Reason of State, that is to say the title by which the State can use discretionary powers in order to preserve the existence of the community – and protect it from its enemies within and outside. If all political information is available to the public, it means it is also available for these enemies: the information strategically related to the subsistence and vital interests of the community must thus stay concealed. The argument of the Reason of State is not intrinsically linked to a state of emergency but to the identification of two distinct spheres: the sphere of war, according to which a political collectivity has for primary purpose the pursuit of its own continuity; and the political sphere, according to which a political collectivity has for fundamental ideal the pursuit of the common good of its constituency. These two spheres are analytically distinct but their issues, arguments and involved actors are entangled so that it is impossible to prevent the circulation of information from reaching enemy hands without blocking its access to the citizenry.

The second one is the elitist argument, according to which the principle of publicity can be useless, and even counterproductive in the framework of a pluralist and liberal democracy. As a consequence of the social division of labor, the main part of the citizenry is not competent to judge the way political representatives handle public affairs. Moreover, citizens are for the most part not willing to participate in public life, whether it be because they are uninterested or because they lack the time to dedicate themselves to the ruling of public affairs. In this context, a democratic regime has to ensure that the general tenets of the formal political discussion are accessible to the public and that the laws, their application and their adjudication be properly publicized, but the political space cannot be held responsible for the lack of information and/or political expertise of the general public space.

The third one is related to the dilemma of the plain political information, which can be summarized through the following sentence: the political information can be clear or complete but not both, and certainly not neutral. Either we assume that the principle of publicity requires translating the political information into a clear and accessible content. In that case, such mediated public information stems from a cognitive selection process. In what conditions
is this process handled? Who is the actor taking charge of the selection of the information? The ideal of plain, clear information unavoidably raises the questions of the possible biases accompanying the selection process, and the power asymmetries which at the same time cause and arise from it. The other option is to assume that providing the citizen with complete political information is an attainable goal. Various studies and experiments tend to show that complete information can paradoxically harm the level of collective understanding. First, the profusion of public information makes it more difficult to find the accurate input. The apparent “neutrality” of the information reinforces paradoxically its opacity for the citizen.¹ The total and overabundant disclosure of information can even be used to dissimulate information that the authority would prefer to be kept discrete. Second, complete information might stay undecipherable for most of the members of the community, and especially those who suffer from social and educational inequalities. The technicality of the laws and administrative regulations feeds the already existing gap between the most and least resourceful and/or specialized actors. Third, pure information doesn’t entail a clear and transparent reception of its content. The communication is a process by which not only the sender but also the receiver of the message contributes to color the content and the meaning of the message. Various experiments in social psychology show for instance that, when confronted with uncommented information, people tend to give more attention to the unexpected, emotional, negative (Rottenstreich, Hsee, 2001; Hsee, Hastie, 2006), loss-related (Tversky, Kahneman, 1991), more accessible or striking (Wilson, 2002) information – despite the objective relevance of this information.

These three arguments can be taken as purely factual: but they also have a normative dimension, since they assume the factual features of the public give a solid normative ground for limiting publicity. Abundantly discussed in democratic theory, the deliberative ideal seems to offer a convincing answer to these three classical critiques of political publicity.

The deliberative ideal derives the legitimacy of a political decision from the deliberation between free and equals parties leading to a rationally justified

¹ “Consider for example, public knowledge and public opinion during the period prior to the invasion of Iraq in 2003. Long after sufficient information existed to disprove the contention, large segments of the American believed (...) in a proved link between Saddam Hussein and al Qaeda” (Fenster, 2006: 929)
consensus (Cohen, 1986, 1989, 1996; Elster, 1986; Habermas, 1992). It relies thus upon two conditions: a condition of collective participation and a condition of argumentation. *The condition of participation* means that the discussion is open to all members of the political community. It has to allow them to express themselves in an autonomous way. The deliberative model proposes a fair, inclusive and discursive procedure “to which all the citizens may participate and by which they are brought to a fair cooperation” (Blondiaux, Sintomer, 2002, p. 18). The condition of argumentation stipulates that the general interest does not result from a transcendent principle or the sum of the interests/preferences of the citizens, but from the formative process of a collective agreement dictated by the rule of the best argument.

In this context, deliberation can be first conceived through its *capacity to attain the conditions of a fair procedure* (I) (Christiano, 1996). A good deliberation gives everybody an equal/fair say. It supposes that the citizens be granted equal political rights and equal opportunities to participate in the decision-making process. To this end, the epistemic quality of the discussion is seen as a tool to clarify the different interests involved, the stakes of the discussion, the rules and narratives of the discussion. Deliberation can be defined as well through *its capacity to contribute to the epistemic quality of the discussion* (II) (see Estlund, 2009). The definition of a rational decision is neither given nor defined by the deliberative process. The deliberative process provides the discursive tools for approaching the conditions of a rational decision. The best arrangement for approaching the independent criterion of a reasonable decision is thus the democratic deliberation. For the classical version of the epistemic model, political deliberation promotes a balanced dialogue, a reasoned exchange of reasons, or a mutual amendments process (II.1). And for the extended version of the epistemic model, deliberation helps to clarify the parties' preconceptions, and to encourage a collective recognition of the different collective identities involved. Finally, political deliberation can be defined as the *recognition criteria of a rational decision-making process* (III) (see Habermas, 1981 and 1992; Cohen, 1989 and 1996). A true deliberation defines what a rational decision-making process is: the ideal conditions of a true deliberation do not only provide a way to assess the quality of the decision but also its correctness. The political deliberation must approach the conditions of the ideal deliberation, which
determines the conditions for a fair and/or rational result. Habermas’, Cohen’s or Benhabib’s accounts of public deliberation not only give an account of the public reason but a standard of rationality.

In this framework, the deliberative conception of political legitimacy strictly frames the epistocratic or decisionist arguments underlying the principle of the Reason of the State, as public deliberation is supposed to contribute to the collective definition and control of the sphere of secrecy. Considering the conceptions I and II of the deliberation, the Reason of State is no longer – and cannot be – defined by the mere fait du prince. Public deliberation either justifies why the decision must be made public, or the conditions under which it can be concealed from the public’s eyes. In this framework, even the decision by which a sphere of secrecy is allowed in some determined cases should ideally be subject to a collective deliberation, before or after creating such a sphere: the question remains open whether the reasons for disclosing or concealing information require having access to its content. The deliberative ideal is not incompatible with the existence of strategic secrecy, but constitutes its hypothetical justification, the practical form by which it is justified, and the channel through which a part of the secret information ends up being disclosed in the public space.

Besides, the deliberative ideal rejects the idea that democracy can be assimilated to the peaceful competitive process by which the political elites could speak and act for the people – in the respect of individual liberties and the rule of law - and assumes that the dilemma of plain political information can be resolved by submitting their content to (the selection and diffusion process of) deliberation. As envisaged here above, political deliberation conceives the common good through the meeting of two conditions: a condition of collective participation and a condition of reasoned argumentation. In this context, the deliberative ideal aims to demonstrate that the incompetence and disinterest of the citizen are not a fatal consequence of the democratic regime. And it assumes that the dilemma of plain political information can be resolved by submitting their content to (the selection and diffusion process of) deliberation.

In this framework, the principle of publicity constitutes both a tool and an objective to achieve with regard to formulating the conditions of argumentation
and participation. As explained by Mill, publicity both promotes the circulation of true ideas and verifies the relevance of the ones which could be wrong:

the peculiar evil of silencing the expression of an opinion is, that it is a robbing of human race; posterity as well as the existing generation; those who dissent from the opinion, still more than those who hold it. If the opinion is right, they are deprived of the opportunity of exchanging error for truth; if wrong they lose, what is almost as great a benefit, the clearer perception and livelier impression of truth, produced by its collision with error (Mill, 1859, p. 20).

Thus, publicity provides the citizen with various tools to have a grasp on the political debate, build up the epistemic capacities of the parties, and foster the collective recognition of the interests (see I). It contributes to the citizen’s political maturity (II). It fuels the argumentation process, forces the parties to universalize their reasons, and - finally – show that these reasons are compatible with the general interests (II and III). Assuming that the practical conceptions A and B of political publicity are supportive of the hypothetical conception C, it might even constitute one of the necessary conditions of the narrow conception of deliberation (III). Reversely, a successful deliberation is assumed to increase publicity. The collective deliberation promotes a large circulation of the facts, arguments, views and interests (A). It produces arguments and information (A and B), and contributes to the formation of the public space by creating a common discussion substrate (A and B). Finally, the conception II of the deliberation is meant to approach the conditions of hypothetical publicity (C), while the conception III is to give a practical definition of it.

Second part: Why Transparency Is Not Reasonable for Political Deliberation

Whatever status is given to the principle of participation, deliberative theorists assume that the creation of deliberative spaces or system contributes to its intensity and quality. Different arguments lead one however to be doubtful about the double converging relationship we described here above, and to envisage the hypothesis according to which a closed-door deliberation might be a better fit, than public debate, for the conditions of a reasoned democratic debate.
First, a secret deliberation within the political space is not incompatible with the free participation of all in the public debate. Deliberative theories could have been criticized for their underlying elitism and the ambiguity they bear on the status of participation in the deliberative systems: is participation only an instrumental factor of deliberation? What if it can be proven that the participation of all undermines the conditions for a rational collective deliberation? In reality, the publicity of the political space to the general public space is not necessary for a vivid general public space. Even if the tenets of the negotiation remain concealed from the public, the actors of the civil society can still debate about the political programs of the actors involved, their possible strategies or the public policies susceptible of being implemented. Besides, it would be wrong to argue that close-door deliberation would create a technical gap between the expertise of the negotiating/deliberating parties and the members of the public space as a whole, as the members of the public space can of course dispose of expertise on their own. The research centers of the political parties and the organized civil society, the investigative work of the mass media, the specific expertise of the academic environment do give the citizen most of the technical information he needs to take position in the public debate.

Second, and getting back to Luban’s sentence, publicity is not necessarily withstood by publicity: closed-door deliberation is not only compatible with the pursuit of the hypothetical conception of publicity (see C), but also the extended practical conception of publicity understood as a background for political justification (B). Better than the public disclosure of the decision-making process, closed-door deliberation is able to support an epistemic deliberative input (II) within the public space and a (more) rational deliberation (II and III) amongst the negotiating parties. Reversely, publicity might harm the quality of the deliberation.

Deliberation relies on a procedural requirement of “sincerity”. This requirement of sincerity implies, on the one hand, that the arguments put on the table are ones considered to be true (Habermas, 1983). Whether they are effectively believed or not, they must be defended as such since they are publicly revealed. Should these arguments be put forward as true, that would imply that they are in fact believed to be true, as they are representative of a substantial
position and can be taken as such by any other reasonable party. On the other hand, the requirement of sincerity does not entail that the public space be provided with a total transparency with regard to personal beliefs, positions and strategies. However, it does require that the arguments set out by one party be true, further calls for other parties to behave the same way, and acceptance that the discussion is ruled by the rule of the best argument. Deliberation does not only encompass a collective recognition of the different available arguments, but also a dynamic of mutual adjustment which seeks to make the general interest prevail over the aggregation of particular interests. Sincerity is here understood as a general principle of deliberative goodwill. Public deliberation does not require sincerity in one’s positions only, but also in the way other parties’ arguments are treated.

Political publicity supports the first meaning of sincerity at the expense of the second one. First, with transparency the actors could become more reluctant to change their minds and less encouraged to take other points of view into account (Elster, 1998, p. 117; Elster, 1994, p. 249). Publicity forces actors to stay coherent with the reasons they originally gave, and allows them to reduce the cognitive gap potentially existing between the public and their personal beliefs. Nonetheless, the cost of maintaining coherence is high for the deliberative process, as any position change is taken as a lack of (self) coherence. The negotiating parties are more reluctant to stand out from their original stance (Fung, 2003, pp. 348-349). ² The public concession of an argument is seen as a defeat for the negotiator for it is perceived as a setback for his political clientele and supporters. And the publicly seen success/influence with regard to an argument is perceived as a victory by the clientele. In this context, being the first to score remains the most accessible way to defend later compromises but also a privileged way to develop pure “plebiscitary” rhetorics (Chambers, 2004).

Second, the political discussion is not a linear process progressively leading toward converging interests, arguments or positions. It is an explorative process, alongside which the parties experiment various negotiative stances, heterodox proposals, hybrid compromises or unpopular second/third-best

---

² As demonstrated by MacCoun et Kerr, open voting done by raised hands locks in the positions of the various actors more that secret voting. (Kerr, Mc Coun, 2004).
outputs. It requires of course taking distance with one’s original positions, enduring some argumentative risks, and pushing creativity to its limits (Bok, 1982; Gutmann, Thompson, 1996). Unfortunately, these elements of dialogue tend only to emerge when the actors are guaranteed they won’t suffer any prejudices for speaking straightforwardly. As defended in the Seale Case about secrecy surrounding the White House’s internal advisory process; presidential advisers will tend to avoid any serious consideration of novel or controversial approaches to presidential problems if they are held publicly accountable for all approaches that were advanced, considered but ultimately rejected. In this context, the Frankness and Candor argument assumes that publicity tends to inhibit innovative positions and favor actor’s aligning to conformist or authoritarian arguments. Taking the example of the Belgian sixth State Reform in which I had the honor to participate as a political advisor – that left Belgium without any federal government for more than five hundred days between June 2010 and January 2012 – I have described elsewhere that the negotiators could not have reached any common ground without adopting and experimenting with borderline interpretations of the Constitution, informal expert consultancy or crude exchanges on each parties’ negotiation interests (Pitseys, 2011). The transformative process underlying the discussion process could not occur without the shared assumption that each submitted proposal goes along with an (in)formal “read and delete” principle.

It can be argued of course that publicity still operates as a filter for the collective decision. Deliberation does not guarantee a correct solution but it permits to neutralize the worst solutions (Evans, Wurster, 1997) and “the more fallacious and superstitious forms of reasoning” (Christiano, 1997, p. 248; see Estlund, 2003, p. 82).

---

3 Thus, closed door deliberation allows to (in French) “explorer toutes les avenues, à l’abri des émois et des pressions de l’opinion, et puissent trouver … une solution telle que les concessions respectives soient moins onéreuses et leurs gains plus élevés” (Pestiau, 2001, p. 39)


5 Analyzing the internal governance of various central banks directing boards, Meade and Stasavage showed for instance that the members of these boards express more easily their divergences with the bank director when the meeting is held on behind closed-doors (Meade Stasavage, 2008)
Prolonging this idea, Elster’s “civilizing force of hypocrisy” positions itself as a powerful objection to the deliberative defense of closed-door deliberation (Elster, 1998; see also Christiano, 1997, p. 248). According to Elster, publicity compels the citizen to show that his reasons are – even for purely instrumental motives – compatible with the general interest and not only his private or particular interests:

there are certain arguments that simply cannot be stated publicly. In a political debate it is pragmatically impossible to argue that a given solution should be chosen just because it is good for oneself. By the very act of engaging in public debate – by arguing rather than bargaining – one has ruled out the possibility of invoking such reasons (Elster, 1997)

Elster not only shows that the private interests are better defended when they are formulated through a public interest argument, but that public pressure prevents the representatives from setting on the agenda interest-driven arguments (even though these interests are thought to correspond to the common good) to the benefit of general interests arguments. The public political discussion prevents the representatives from putting forward pure opportunistic arguments or partial arguments, but also those normative arguments that do not find common acceptance. Reversely, the publicity of the discussion gives the parties various political and socio-psychological incentives to propose public interest arguments. Furthermore, it would benefit the actor’s objectives to truly believe in his line of argumentation – and so reduce his internal cognitive dissonances (Elster, 1995, p. 251; Goodin, 1992, pp. 137-138).

The argument of the “civilizing force of hypocrisy” is not without raising serious questions yet. First, what does “civil” mean? The “civilizing force” argument supposes that the nature of the public interest is uncontroversial, or at least that a social consensus can be found on some fundamental values. But the case for/against political publicity precisely occurs when the political actors have to find an agreement between different perceptions of the common good. A closed-door deliberation is envisaged when the actors have to deal with controversial issues, which spark substantial normative divergences. We could of course counter-object that the original disagreement between the parties is not really caused by divergences about the nature of the public interest, but by disagreements on the consequences it bears or on its practical translation in the political reality. Each party can agree on the fact that “Belgium is worth being
saved as a unified state” or that “The Belgian sixth State Reform must preserve the fundamental pillars of the social security system”, but diverge on what a “unified State” or “fair social security system” is. But in that case, the counter-objection actually gives another reason to dismiss the argument of the “civilizing force”. Even if publicity orients the actors towards public interest arguments and creates a collective notion of common decency, the real reason why the actors negotiate is not linked to the fact they do not share the same public interests concerns, but that they do not understand, interpret and use them in the same way.

Second, what does “civilizing” mean? Simone Chambers shows in “Behind Closed-Doors Deliberation” that the case for/against publicity not only opposes private and public reason arguments, but stages as well a third register of political rationality: the plebiscitary reason. As put forward by Chambers, “speakers still appeal to what they think are common or public values but with a twist, under the ‘glare’ of publicity these arguments may become shallow, poorly reasoned, or appeal to the worst that we have in common” (Chambers, 2004, p. 394). Manipulation, flattery, pandering or agenda-setting strategies can perfectly be implemented through arguments calling for the common good or the public interest. Finally, even within a civilized public space the idea of a “publicly justifiable” argument does not have any necessary relationship with the idea of deliberation, and more broadly a transformative discussion process. The representatives are still publicly accountable to their electorate and will tend to determine their position according to it (MacCoun, 2006, p. 115). Even a sincere and shared call for common decency can be incompatible with the procedural rule of the best argument.6 As a matter of fact, Elster himself acknowledges that public interests arguments are compatible with undebated emotions, conformism, or pure majoritarian rhetoric. There is no reason to think that the rational treatment of the public interests is less controversial than the substantial definition of its nature.

Third, who is the “civilizer”? The argument of the “civilizing force” not only supposes that the general interest can be collectively defined by the general public

---

6 If the views of the public remain unknown, Lerner and Tetlock show however that the representatives will tend to be more careful in their auto-justification process (See Lerner, Tetlock, « Accounting for the effects of accountability », Psychological Bulletin, vol. 125, n°2, pp. 255-275, in MacCoun, 2006: 115).
space as a whole, but that such a space practically exists. Belgium – but also ideologically highly-divided liberal countries like the United States (Jamieson, Cappella, 2008; Krebs) – illustrate how various autonomous public spheres can coexist within the same national space, and interact like opaque and independent constituencies (Hahn, 2002). There are both Flemish and French-speaking distinct public opinions, each of them constituted by their own medias, pundits, television channels, informal political networks and civil society. In this framework, the mere notion of public debate is warped from the very start, as – taking here some random examples – the very meaning of a “fair compromise” between the two linguistic communities, the historical interpretation of the previous State Reforms or the understanding of the concept of “federalism” do not even find a common political space for these divergences to be discussed. The different fractions of the society do not even confront their disagreements. Each negotiating party can thus defend his own conception of public interest to his own constituency without taking any risk of being contested by the others’, further knowing that their (the others’) arguments will remain opaque to his own electoral/sociological/linguistic clientele.

Another counter-objection is that the deliberative defense of closed-door deliberation relies on an overly restrictive conception of deliberation. First, deliberative theories have for long extended the notion of deliberation beyond its classical definition as a rationally motivated consensus. Deliberation cannot be reduced to the pursuit of a reasoned argument, or the assessment of the epistemic quality of the discussion within the political space. Understood as fair proceduralism, deliberation aims to free the citizens from the distortions resulting from unequal political power (see I). Understood as a way for approaching an independent criteria of collective rationality, it can be seen (see II.2) also as a tool for raising citizen’s sense of civility, favoring the collective recognition of actors (Williams, 1995; Young, 1990), or identifying and clarifying the various actors’ positions (Mansbridge, 2006). Second, the epistemic disadvantages of publicity are to be counterbalanced by the advantages it might have for the social deliberative system as a whole, at the level of all the differentiated but interdependent elements which, within society, contribute to tackle political issues through “arguing, demonstrating, expressing and persuading” (Mansbridge et al., 2012, p. 46). The deliberative system designates
the different levels of power, the institutions and actors composing the deliberative substrate.7

Unfortunately such a counter-objection is both ambiguous and reversible. If deliberation is ultimately defined through its capacity to define or approach (even minimally) a rational decision, we still have to show that the civilizing or equalizing effect of publicity – as desirable as it may be – has a comparatively more positive (or less negative) effect on the epistemic quality of the decision than a closed-door deliberation. First, the notions of procedural fairness or common civility do not evoke intrinsically any public reason: their relationship is instrumental at best (Estlund, 1997, 2002). Second, the notion of deliberative system:

does not require that every component have a function or that every component be interdependent with every other such that a change in one will automatically bring about a change in all others. If a component does contribute to a function, it is not necessary that the function be fulfilled optimally in one location, since in a deliberative system the same function may be distributed across various subsystems (Ibidem, p. 5).

In other terms, it’s not because the political discussion takes place out of the public eye that the different institutions and actors of the deliberative system are not able to debate about – or control - the content and stakes of the decision to come. As envisaged above, a closed-door deliberation doesn’t prevent the actors of the civil society from nourishing the public debate and allowing the deliberation to fulfill its different possible functions. Chambers wrote that “the democratic function of publicity - are policies in the public interest? - can be achieved in secret via a thought experiment but the Socratic function of publicity - are policies based on the best possible reasons? - often cannot” (Chambers, 2004, p. 407).” The sovereign can seek advice from the public - and even ask to the public whether secrecy is desirable or not - without disclosing the content of the political discussion.

If we can admit that perfect communication would guarantee both a perfect deliberation and the conditions of the hypothetical test of publicity, we must concede that such a transparency is out of reach. First, epistemically weak

---

7 For the authors mentioned here above, the deliberative system includes « roughly, speaking, four main arenas : the binding decisions of the state (both in the law itself and its implementation ; activities directly related to preparing for those binding decisions ; informal talk related to those binding decisions ; and arenas of formal or informal talk related to decisions on issues of common concern that are not intended for binding decisions by the state” (Ibidem p.9).
conceptions (I and the extended version of II) of deliberation could oppose the conceptions B and C of publicity. Second, the two practical (A and B) conceptions of publicity could oppose the two epistemically strong (the classical version of II as well as III) conceptions of deliberation. Third, epistemically weak and strong conceptions of deliberation could oppose each other respectively. And fourth, weak and strong conceptions of publicity could oppose each other. Consequently, defending an epistemically demanding conception of deliberation can lead to preferring the hypothetical conception of publicity (C) to the practical conceptions of publicity (A and B), and to disentangle the existence of a collective background for the justification of public discussion (see B) from a systemic requirement of public information (A). Even in a democratic system, there are thus some strong epistemic reasons to prefer a closed-door deliberation to a public deliberation as long as the autonomy of the civil society and the possibility of an ex post control of the decision are guaranteed within this system.

Third Part: When Do We Have To Accept Or Reject Political Publicity?

Notwithstanding the arguments mentioned above, we think that there are some valid reasons to prefer publicity rather than a closed-door deliberation. My first argument against closed-doors deliberation is both internal and descriptive: it contests that secrecy would further the deliberative process better than a public discussion. My second argument is both external and axiological, and contests the status and definition given of the deliberative ideal.

3.1. The Regressive Argument.

A variety of reasons lead to think that closed-door deliberation presents worse inconveniences than publicity; that it reproduces them at least; and that the balance between these two conceptions of political visibility cannot be established with certainty anyways.

Assuming that deliberation diminishes the public pressure on the negotiating actors relies on the idea that the level of public pressure is related to the presence of the public eye. We think this idea conflates the practical concept of publicity, which designates the public access to information, with the concept
Why the deliberative ideal justifies publicity – even if publicity may undermine deliberation

of accountability, which covers the idea that the political representatives are politically – and sometimes legally – liable. It is not because a discussion is not public that its parties are not legally or politically accountable. In this framework, the public pressure on the representatives could be even heavier than in the framework of a public debate, insofar as many relevant elements have been knowingly hidden from the public, and citizens can assume that these elements were hidden precisely because their disclosure could have influenced the parties’ behaviors. In this context, the parties are tempted to anticipate what will be the pressure of the public spotlight once they will have to disclose the final result of the negotiations. It is all but certain that late disclosure would diminish the level of public pressure on the negotiators.

We could object to this argument that closed-door deliberation still gives the negotiators more autonomy to debate in a dispassionate way, explore new arguments, put forward general interests rather than partisan arguments. In reality, the conception of secrecy defended by supporters of closed-door deliberation is no less naïve and abstract than the conception of publicity they pretend to highlight. The reason why publicity yields negative effects on the deliberation is linked to the fact that the epistemic ideal would be total transparency, that the mere fact people choose to deliberate proves that such a transparency is a mirage, and that the second epistemic/deliberative best paradoxically becomes the figure of the closed-doors deliberation. Nonetheless, the reason why closed-door deliberation is supposed to prevail also relies on the problematic assumption that total opacity is possible. In reality, such opacity never exists: among many others, the case of the Belgian sixth State Reform illustrates how the information constantly flows from the negotiation table to the public (twitter, negotiators’ entry speeches, leaks…) and from the public to the negotiation table (boards of the party, civil society…). The leaked information is important enough to grasp the public’s attention but not complete enough to ensure trust between the public and the political space.

---

8 Going to the point that some of the negotiation documents distributed to the negotiators were each given a slightly different presentation (spelling mistake, words missing…) in order to give the possibility to recognize the source of the leaks to the media.
Besides, all the epistemic arguments used against publicity can be reproduced in the negotiation space. For Elster a secret discussion favors negotiation more than deliberation insofar as it is easier to negotiate on strategic arguments rather than deliberating on general interest positions (Elster, 1995). It doesn’t mean that closed-door discussion does not give any space for public interests or reasonable arguments. But there is no reason to think that the “frankness and candor” of the closed-door deliberation serve a deliberative dynamic rather than unaccountable arguments. The progression toward a common agreement is all but the same as the progression toward a common good. The fact that the “civilizing force of hypocrisy” may be counterproductive does not mean that a closed-door deliberation would be comparatively more civilized. The negotiators may decide to concede a point by pure weariness, or link their consent to the short-term gain they would take from the proposed reform - a punctual financial advantage for the competent executive in the framework of a global reform of the social security system, an electoral advantage due to reframing the electoral districts. It is possible to contest the nature of the common good summoned through the public debate, but it is impossible to control to what extent this common good is even discussed during a closed-door deliberation.

We could object again that a closed-door deliberation is never totally secret. There is an informal dialogue between the organized civil society and the representatives in charge of the political discussion. Moreover, the negotiation field constitutes a social sphere in itself. Taking again the example of the Belgian sixth reform, the negotiating parties were part of a collective story, which pushed them into a relationship of mutual trust. The discussion is an ongoing process that accumulates partial negotiation drafts, intermediate arguments, and informal discussion clauses. The parties are often constrained to negotiate in a specific physical space – the Belgian Senate, the Cabinet of the Minister of Social Affairs... - eat together, are stuck in the same room until the late hours of the night. The negotiation sphere ends up forming a micro deliberative system in itself, where the collective narrative and the individual social capital of the negotiators play a determining role. It is even difficult to step out of this collective narrative that it supposedly agreed by all the parties. It is thus highly difficult for a negotiator to put forward an openly destructive position, a pure personal interest argument, or a totally eccentric proposal.
In reality, the existence and legitimization of such a micro deliberative system helps to understand why the “frankness and candor” argument and the “civilizing force of hypocrisy” are so reversible in the framework of closed-doors deliberation. Indeed, it is far from certain that closed-door deliberation allows the parties to explore new arguments or test audacious ideas out of the public spotlight. The pressure within the negotiation field is as heavy as it would be in front of the public, it engages the political and social capital of each participant. The bargaining power of the negotiators depends on their social resources amongst the other parties. The success of an argument also depends on the way it is presented, the way it fits the narrative of the negotiation, its common linguistic codes, informal devices... There are thus some arguments that just cannot be defended because they do not fit this common narrative. Closed-doors deliberation creates thus an “enjailment effect” that inhibits the negotiators’ creativity, and generates its own endogamous rhetorical biases. As for the Belgian case for instance, four of the eight political parties (the French-speaking and Dutch-speaking ecologist parties as well as the French-speaking and Dutch-speaking liberal parties) originally pled for the creation of a federal electoral district at the level of the whole constituency, allowing thus translinguistic electoral debate for ten to twenty-five percent of the elected federal representatives. This position had been abundantly discussed in the newspapers, and never explicitly rejected by the other four political parties around the table. Nonetheless, the idea turned out to be progressively excluded from the negotiations field – and even ridiculed – without even having been discussed by the partners: the negotiators endowed with the largest social capital constantly dismissed the point from the negotiation agenda; the single world “twenty-seven” (referring to the amount of new-elected nationalist Flemish representatives) pronounced by a negotiator was enough to rule out the point before any single discussion about it; and the expression “my party wants...” by the supporters of the electoral district served as a clear signal that their official support for the measure would not correspond to their effective strategy around the negotiation table. Frankness doesn’t mean sincerity and candor does not apply for audacity: the reason why the civilizing force of hypocrisy can lead to conformism or plebiscitary rhetoric are exactly the same why the micro society composing the
negotiation field is far from guaranteeing the conditions of a reasonable discussion.

The arguments raised for defending secrecy can be reversed so that they bring us to criticize it on the exact same base: besides, closed-door deliberation has to face specific objections regarding its epistemic perspectives. It does not mean that publicity is preferable from an epistemic point of view, but (a) we can think that closed-doors deliberation is not necessarily more rational than public deliberation, and (b) that the comparison between these two configurations can hardly be determined.

In this context, a good reason left to justify closed-doors deliberation is the necessity to find an agreement and give a conclusion to the deliberation. Even if the internal narrative imposed to the negotiation field comes to undermine the conditions of deliberation, it also contributes to aggregate the different available positions: *a contrario*, and all epistemic consequences aside, we do not contest that public pressure hinders mutual concessions and the progression to a collective agreement. Justified or not, it is interesting to note however that the consensus argument is no longer related to any epistemic or deliberative argument. A reasoned discussion does not necessarily lead to a common agreement, and a collective agreement does not necessarily lead of course to a reasoned discussion. The possible justification of closed-door deliberation would be thus independent from its reasonable/deliberative nature.

### 3.2. The Indeterminacy Argument

The defense of secrecy relies upon the assumption that the definition of ‘rational deliberation’ is known and shared by everybody. Claiming that closed-door discussion is best presupposes that we all know and agree on what a reasoned – and more fundamentally, what a legitimate – decision is. But if it is the case, why should we keep the deliberation secret? And even if a common shared conception of deliberation does not entail a collective practice, is this conception really shared? Who decides such a conception is justified to defend either publicity or closed-door deliberation?

Assuming we live in a pluralist society, we must expect that the arguments and procedures that the negotiators take as important might not be the
arguments that the observers take as important, whether it be at a normative or factual point of view. In this framework the arguments and procedures that the negotiators take as important might not be the arguments that they communicate to the observers as important. It means that the observers are likely to disagree, at least partly, with the positions and procedures discussed within the space of political decision. This is precisely the reason why the content of the discussion must remain concealed to the public. It means moreover, that the observers are not supposed to know when and why such a disagreement occurs.

In this context, it may be hasty to conclude that closed-door deliberation is compatible with an open debate. The asymmetry between the information available to the negotiation and the information available to the observers affects the terms of the discussion taking place within the general public space. Such an asymmetry could affect the terms of the public debate. The normative arguments discussed in the public space do not necessarily align with those that are debated in closed-door deliberation (the public space debating for instance the terms of a possible federal district even though the question is practically precluded from the negotiation field). The factual elements of the closed-door deliberation remain hidden from the public eye as well, leading the public opinion to debate on the basis of erroneous or irrelevant data. The possible obstacles and threats to the negotiation are subject to an analysis gap between the public opinion and the negotiation sphere: as experienced during the preliminary phase of the Belgian sixth State Reform, for instance, the atmosphere between the negotiators was never better than when it was described as ‘terrible’ by the mass media (and the other way round). Finally, those asymmetries of information nurture themselves insofar as the observers could not even know the relevant information to discuss and assess the decision-making process. The observer does not know what influences or controls the political discussion.

Consequently, the informational asymmetry between the public space and the negotiation sphere can of course pervert the terms of the secret deliberation. The way imperfectly informed arguments influence public debate is likely to orient on its turn the terms of the secret negotiation. In this context, the (real) pressure exerted on the negotiating parties can be based on (wrong) rumors and assumptions – as well as the corresponding threats and demands – emerging
from the public space: the negotiators even make up to leak partial or even incorrect information in order to influence the power balance of the internal negotiating agenda.

But above all, those asymmetries end up reinforcing the belief according to which only the negotiators are able to deal with the political issues that are submitted to them, and choose whether or not it would be rational/strategic/legitimate to disclose the political information. In this framework, the argument against publicity is a new formulation of liberal elitism. The information gap is justified by the fact that the negotiators supposedly know better than the public and it is neither possible nor necessary to bring the expertise level of the public and the political space up to par. The deliberative argument against publicity precludes thus the very possibility of a “deliberative system” (Mansbridge et al., 2012) at the level of the common public space. Besides, it preempts the definition of what is a rational decision-making process. In a closed-door deliberation, the negotiators not only presuppose that a secret deliberation is better than public deliberation; that the search for the highest possible epistemic input within a small set of deciders is preferable to the search for a higher epistemic input within the public space; that the conditions of deliberation in the framework of a closed-door deliberation are preferable to the conditions of deliberation within a public deliberation; and that the deliberative second best is thus a closed doors-deliberation whose result will be later controlled by the public opinion. They assume that the observers do not need to have the possibility to access, discuss, and assess the terms of this definition. They also take for granted that a discussion on the conditions of desirability – and by extension, the practical implementation – of the deliberation does not need any deliberation. Staging themselves as an epistemic open process of rational justification of the collective decision, the deliberative defense does not take into account the fact that there can be a reasonable disagreement on what a deliberative second best is, and whether it is desirable or not and even whether deliberation must be the ultimate criterion for assessing the legitimacy of a political decision. The observers cannot debate the conditions underlying their exclusion from the political decision.
The justification of publicity does not rely on its epistemic positive effects on deliberation, but on the fact that the assessment and definition of these positive effects cannot be left to the negotiating parties. Public deliberation is preferable to closed-door deliberation if and when it increases the available information on the different interests at stake, and because it allows staging a broader variety of common good perspectives (Christiano, 1997, p. 260).

In this context, the only reason left for justifying closed-door deliberation would consist in admitting that some people have more legitimacy than others to define the public interest or what a true deliberation should consist of. As underscored by Chambers or Thompson, such an assumption can be debated at the occasion of a collective democratic debate defining the legitimate conditions of secrecy, or at the occasion of an ex-post assessment of the negotiation process: secrecy would be thus legitimate if it receives a deliberated assent. Unfortunately, the secrecy of the negotiation field creates its own acoustic room. As envisaged supra, the observers do not have the possibility to assess whether these pre-defined conditions are effectively respected and can never be sure how and where to orient their spotlight in the ex-post assessment phase. The negotiators keep the upper hand on the interpretation of what a reasoned discussion is and the definition of the legitimate sphere of secrecy.

**Conclusion**

We may admit that political publicity can produce negative effects on the quality of the political discussion. We cannot be sure, however, that these negative effects are more certain and severe than those of a closed-doors deliberation. We have showed why closed-door deliberation could not be accepted by a reasonable citizen and why the principle of political publicity is justified despite its possible negative effects on the quality of the deliberation.

This paper leads us then to two concluding points, the first of which relates to the status of publicity. As envisaged in this paper, political publicity is both an obstacle and a necessary factor for deliberation. On the one hand, publicity stages inescapably diverging conceptions of political discussion. The mere coexistence of more or less self-interested, activist, agonistic conceptions of political
discussion makes a true deliberation impossible, as the principle itself is subject to a practical disagreement. If there is no congruence between the practical principle of political publicity and that of formal publicity (required by the rational motivation of the practical decision), it is not because this practical principle of publicity cannot ensure the conditions of communication thanks to which this requirement is conceived but because there is no consensus on these conditions. On the other hand, publicity allows for the recognition of deliberation as a “publicly significant” theme (Habermas, 1992) which can be submitted subsequently to a collective discussion, judgment or agreement.

The second point relates then to the way the case for publicity helps to reshape the question of political legitimacy. Secrecy does not allow treating the different possible conceptions of deliberation – and by extension, the different possible deliberative second best patterns – on an impartial ground. Reversely, publicity must not be defended because it ensures a higher deliberative quality but because it is contextually impossible to anticipate what constitutes a rational decision. The discussion on the conditions of the discussion must remain equally open to discussion, as doing otherwise would be contrary to the procedural conditions of political equality. Publicity must be supported because there is no reason to give the negotiating parties a discriminating prerogative to define, assess and control what should be a rational decision. Deliberative theories of democracy assume that the participation and argumentation conditions of the deliberation are to be mutually supportive. In the case they are not, this paper shows that deliberation is at best an instrumental way to ensure political equality. In this sense, closed-doors deliberation must be rejected because it undermines the conditions of political equality within the public space.

Finally, the ultimate criterion for justifying – or not justifying – publicity may be not its deliberative appeal but the way it fosters political equality. If closed-doors deliberation has to be rejected when it opposes a general principle of political equality, it could be equally supported if compatible with a fair deliberative system. Assuming it is impossible to offer an a priori practical hierarchy between the different available conceptions of the political discussion, secrecy might then be justified if it gives the actors the right and means to choose what they think is the most accurate discursive way to defend (build alliances,
Why the deliberative ideal justifies publicity – even if publicity may undermine deliberation

conceive policies, organize common communication) their political preferences – as long as these tools don’t harm other people’s equal political rights. This point deserves to be developed further.

References


RADICALISM AND COMPROMISE
INTRODUCTION

The Centre for Ethics, Politics and Society (CEPS) at the University of Minho held the third edition of the Braga Colloquium in the History of Moral and Political Philosophy on the 1st and 2nd February 2018. This annual conference has been conceived as an encounter among researchers fostering an amicable discussion among peers which would attract prestigious scholars and provide young researchers with an opportunity to present their work and promote the study of the tradition of political and moral philosophy and its legacy in shaping our institutions, culture and beliefs. But the goal of this series of conferences is also to focus on how this tradition can contribute to tackling the challenges our societies are facing today. Every year the conference will have a specific theme, which will be chosen by taking into consideration the current global political situation. So, in line with the spirit behind this new series of conferences, the third edition of the Braga Colloquium in the History of Moral and Political Philosophy explored the ideas of “radicalism” and “compromise”. The Centre for Ethics, Politics and Society (CEPS) at the University of Minho held the third edition of the Braga Colloquia in the History of Moral and Political Philosophy on the 1st and 2nd February 2018. This annual conference has been conceived as an encounter fostering an amicable discussion among peers which would both attract prestigious scholars and provide young researchers with an opportunity to present their work and promote the study of the tradition of political and moral philosophy and its legacy in shaping our institutions, culture

---

* Integrated Researcher at CEPS (Center for Ethics, Politics and Society), University of Minho.
and beliefs. But the goal of this series of conferences is also to focus on how this tradition can contribute to tackling the challenges our societies are facing today. Every year the conference has a specific theme, which will be chosen by taking into consideration the current global political situation. So, in line with the spirit behind this new series of conferences, the third edition of the Braga Colloquium in the History of Moral and Political Philosophy explored the ideas of “radicalism” and “compromise”. Politics has frequently been defined as the art of the possible or the art of compromise. More dramatically, it has been depicted as the realm of Faustian deals and tragic choices. Max Weber famously wrote that the political call demands endurance in the face of disappointment. It is the realm of frustration and sacrifices, of fragile equilibriums between *fiat justitia* and *raison d'état*.

Sometimes the existence of political structures of accountability relaxes the dependence on character, reputation, and honor among conflicting parties. Reasonable civic duties suffice to deflate social conflicts and to compensate offended actors. Lacking these institutions, integrity becomes non-negotiable for social trust.

On occasion, however, a social order of tolerance would not emerge without sacrificing the moral integrity of former heroes that we now consider dogmatic integrists. Conversely, this institutional order of tolerance also allows the political space for the reconstruction of identity claims for recognition that derive their radical strength from their intrinsic aversion to political settlement.

From a historical point of view, our political languages and attitudes towards compromise, negotiation, bargaining, and agreement have changed in a myriad of contexts and traditions. As so did our conceptions of what seemed once worth sacrificing or defending**.

The format of the conference includes a keynote speaker, a prestigious scholar in the field either of Ethics or Political Philosophy, or both, who is generally recognized for their contributions to the discussion of the conference’s topic. Professor Avishai Margalit, Professor Emeritus at the Hebrew University

---

**My special thanks to Professor David Alvarez, from the University of Vigo – and a very much missed former integrated member of CEPS - for drafting this year’s call for papers.
of Jerusalem, who, sadly, wasn’t able to travel due to health problems, would have been our guest this year. But we were fortunate enough to have as keynote speaker Prof. Simon Cabulea May, from Florida State University, who not only generously accepted our late invitation, but also participated actively in the many discussions arising in the various sessions of the conference. As organizer of this edition of the conference and as a member of CEPS, I wish to express here my deep gratitude to Professor Simon Cabulea May for his critical role in making this a successful event, attracting participants from all corners the world.

Another feature of this series of conferences is that the authors of those presentations deemed most relevant are invited to submit an article based on their presentations for publication in a special dossier in the Ethics, Politics and Society journal. This year, six articles which impressed by their quality, have been collected and are now herein presented.

In “A Politics without Compromise: The Young Hegelians and Politics” Vivien Garcia explains how compromise is rejected by the Young Hegelians since it is incompatible with their Hegelian stance on reality as a dialectical process. Julio. A. Andrade, in “Enacting Levinas’s infinite responsibility as an ethico-political compromise”, argues for a particular reading of Levinas on the relationship of ethics and politics as compromise which permits the possibility of there being continuity, rather than a gap, between the two, and discusses the implications of this interpretation for business ethics. Paolo Vanini, “A reactionary fascination: Emil Cioran and Joseph De Maistre” explores Cioran’s reading of De Maistre’s thought as an instrument with which to investigate the ideological debate on reaction and revolution that took place in the twentieth century. In "What Does Being An “Aristotelian” Really Mean" Sinan Kadir Çelik addresses the ambiguity with which terms like “(neo-)Aristotelian” or “Aristotelianism” are used in current literature on ethics and political philosophy, and proposes a definition or characterization of an “Aristotelian” way of inquiry in these fields. In "Compromise and/or Integrity? Some Normative Remarks.", Beatrice Magni offers a conception of compromise which distances itself from negotiation and preserves the demanding requirements of moral fairness. Finally, Yuval Eilon discusses the cost of compromise in terms of political virtue and proposes that being uncompromising is sometimes the best course of action.
Abstract. This article aims at exploring how most of the Young Hegelians came to reject all forms of compromise. It will first show how Young Hegelianism itself was born from a process of radicalisation. Then, it will expound some of the theoretical developments that this process produced and explain why and how all forms of compromise came to be rejected. For Young Hegelians, a compromise is an anti-dialectical position. It consists in the adoption of a median posture, which does not correspond with a real mediation. It is a way of deflating conflicts and, more precisely, to avoid the oppositions at work in history being unveiled in their purity.

Keywords: Young Hegelianism, compromise, dialectics, “juste milieu”.

0. Introduction

For a long time, the Young Hegelian movement has been considered as an epiphenomenon — or a transitional form — in the history of the Hegelian school...
and in the history of ideas in general. Fortunately, an increasing number of recent studies have revealed the great interest of Young Hegelian reflections and demonstrated they can greatly deepen our understanding of political modernity and, consequently, our present political condition. Jürgen Habermas even wrote in *The Philosophical Discourse of Modernity*: "Today the situation of consciousness still remains the one brought about by the Young Hegelians when they distanced themselves from Hegel" (Habermas, 1987, p. 53). Nonetheless, one of the key elements of Young Hegelian views seems somehow marginal in the history of modern political philosophy: the absolute rejection of all forms of compromise.

This article aims at exploring this aspect. It will first show how Young Hegelianism itself was born from a process of radicalisation, as a result of several breaks. Then, it will expound some of the theoretical developments that this process produced and explain why and how all forms of compromise came to be rejected.

1. Young Hegelianism as radicalisation process

1.1. Ruptures in the Hegelian School

As suggested by its name, the Young Hegelian movement was closely related to the Hegelian philosophy. It was actually born inside the Hegelian school as a consequence of its several splits. Its emergence can be described as the result of a theoretical and political radicalisation process.

The first decisive break was the division between a left-wing and a right-wing in the Hegelian school. It was first formulated in 1837, in David Strauss defence of his scandalous book *Life of Jesus* (Strauss, 1837). As put forward in its title, this writing was a historical study (in a Hegelian sense, of course) of the life of Jesus. The approach he adopted was radically different to that of dogma and theology. Strauss intended to read the Gospels in the light of the cultural context of their time and to replace this context in the framework of a universal and philosophical history. He aimed at hunting down the mythical additions of these texts. Even if the inspiration of such a perspective was obviously Hegelian, it became matter of debate inside the Hegelian school. Strauss completely accepted it, he even justified the controversy. He thought that some passages in the
Hegelian corpus were not totally clear, and that they permitted a variety of interpretations. At the heart of the matter lied the fact that Hegel explained that religion and philosophy had the same content but that they differed in their form (see e.g., Hegel, 1830, § 573, p. 582). Religion offered a representation of the truth, whereas philosophy developed a conceptual knowledge of it. Here was, to Strauss's mind, the ambiguity: to what extent is the content independent of the form? Strauss considered that the religious form implied a conservation, a crystallisation of the past, and that it could not follow the historical becoming. Hence the mythifications. As Douglas Moggach puts it, “Strauss depicts the Hegelian distinction between philosophy and religion, concept and representation, as an antinomy. The contents of religion are not identical to those of philosophy but reflect a lower, pre-rational awareness. Religion is thus displaced within the system of absolute spirit” (Moggach, 2007, p. 63).

Of course Strauss's views were not shared by the whole Hegelian school. In his defence of his *Life of Jesus*, Strauss provided an overview of these divergences of opinion. He distinguished three branches in the school, based on the different reactions his work received. The first one, the right wing, tried to prove that the Hegelian corpus did not contain one shred of irreligion. The second one, the left wing, accepted and used the possibility of a rationalistic, historical, and critical reading of the religious writings. The third one, the centre, was undecided on the issue (see Angaut, Buée, Clochec, & Renault, 2015, pp. 15–24).

Rather than being strictly descriptive, this categorisation deepened some tensions inside the Hegelian School. The *Annals of Halle*, the review edited by Arnold Ruge and Ernst Theodor Echtermeyer, offers an interesting example of this evolution. When created in 1838, its initial objective was to compete with the *Annals for Scientific Criticism*, the official review of the School — which had been founded by Hegel himself and Eduard Gans. The editors of the *Annals of Halle* considered the official review “too centred on the defence of the Hegelian system and on academic questions” (Angaut et al., 2015, p. 19). Without any reference to the straussian typology, they just aimed at creating a more spirited publication addressing the problem of its time. But the *Annals of Halle* gradually became one of the main voices of the left wing of the school. The convergence between their initial aspirations and the debate on the place of religion in the system led them
to a broader scope of concerns and questions. One of them was directly political: does Hegelian philosophy support conservative or progressive views? Inspired by the French Revolution, the left Hegelians went for the progressive answer. Nonetheless, their positions covered a spectrum of views ranging from constitutional monarchy to radical republicanism — at least initially (see, e.g., Clochec, 2014).

If the first split in the Hegelian School occurred from the inside, the second one came from the outside. It was provoked by the attacks on Hegelianism (in general) by some conservative and neo-pietist intellectuals. In the autumn of 1838, Ernst Wilhelm Hengstenberg published an article called “The Decline of the Hegelian School” — published in Friedrich Tholuck’s *Literarische Anzeiger* [Literary Gazette] in August, September, and October, 1838 —, where he argued that the Hegelians who opted for a rationalistic perspective on religion and a political defence of the French Revolution only revealed the true nature of the Hegelian system. From his Christian and monarchist point of view, the whole Hegelian school was a danger, therefore it had to be fought. During the same year, Heinrich Leo, another neo-pietist, published a book called *Die Hegelingen* (Leo, 1838) — which could be translated as “The Hegeliards” or “The Hegelsters”. Leo had until recently been influenced by Hegel, so if he was as vehement as Hengstenberg, it was also more precise. He thought that some ambiguities of the Hegelian corpus could be exploited in an atheistic and revolutionary way, as left Hegelians did. Furthermore, his words were more catching than Hengstenberg's ones and some of them acquired an unexpected impact. As a synonym for *Hegelingen*, Leo used “Young Hegelian”. The reference to the Young Germany, this “group” (see Taillandier, 1848) of German writers who advocated (beyond new literary forms) a set of liberal political positions inspired by the French Revolution, was obvious. The purpose of such an identification was clearly disparaging.

These attacks, of course, elicited responses from the left side of the Hegelian School. Bruno Bauer (who recently converted to left Hegelianism after having defended a right Hegelian position for a while) became particularly involved in this controversy (on the evolution of Bruno Bauer's thinking see Moggach, 2007). His entanglement started with a theological and philosophical debate with
Hengstenberg, his future colleague at the University of Bonn. At first, he mainly tried to prove that the pietists's assertions were slanderous because they were based on fallacies; he still regarded Hengtenberg as a worthy adversary, and his remarks were quite tempered. But Bauer gradually radicalised his comments. This process followed a double movement. The first one lied in the development of some criticisms against Strauss's ideas. As an example, he believed that Strauss's views on history and religion fell back in some form of substantialism that made no room for subjectivity (see Moggach, 2007, pp. 45–46). In doing so, he took some distance from a point of reference of the controversy but he did not come to agree with neo-pietists. The second movement that characterises Bauer's radicalisation is not a purely intellectual one; but the neo-pietists' attacks were not purely intellectual either. They fitted into the political context of the investiture of Frederick William IV of Prussia. If, when he came to power, the new sovereign rose hopes of a new political sequence breaking with the repressive politics of his father, he soon followed a very conservative and reactionary (as Young Hegelians came to qualify it) policy. One of its specific features was the struggle against the Hegelian ideas. Not surprisingly, both Hengstenberg and Leo exerted a powerful influence on the Prussian Camarilla, a circle of court advisers counselling the King behind the scenes. As a consequence, the main left Hegelian organs were censored and banned. As repression progressed, Bauer assaulted more directly the religious and political base of the regime. In doing so, he opened the way for the second movement of his radicalisation process: he came to claim that, in reality, the neo-pietists' statements were fully accurate and that “true” Hegelianism, as Hengstenberg presented it, was atheistic and revolutionary. The rationalist Straussian views on the Gospels were transformed into atheism, and the progressive views on politics became fully revolutionary. The title of the most famous satirical book written anonymously by Bruno Bauer (and in part by Karl Marx), *The Trumpet of the Last Judgement Against Hegel the Atheist and Antichrist*, speaks for itself (Bauer, 1841/1989).

The growing repression further strengthened Bauer's position. In 1842, his dismissal from the University of Bonn — upon a request authorised by the (anti-Hegelian) Secretary of State for Education Johann Albrecht Friedrich Eichhorn — carried a great symbolic weight. Philosophy had become a political affair!
1.2. Swerving from Hegel

Paradoxically the Young Hegelians had to face an obvious problem: the political substance of Hegel's writings was different from their own, and assuredly less radical. But to radicalise, in that sense, does not mean to change course; the radical is not the pure other, even if he is no more the same. But how can they justify that kind of swerving, and especially in a systematic theoretical framework? Views differed in that respect. Had Hegel accommodated his political considerations with the existing order? And in this case, did the theoretical principles of the system remain intact? Or were these principles unsatisfactory and should be criticised? Then what would remain of the system? Two main views opposed (for an in-depth presentation, see Clochec, 2018).

The first one was the “accommodation” hypothesis. Its origins can be traced back to a distinction made by some Hegelians, and Heinrich Heine notably, made (even before the first break in the Hegelian School) between Hegel's esoteric doctrine and its exoteric coating (on the Meinersian origins of this distinction, see, Clochec, 2018, p. 303). Heine claimed that if the regime could approve of the second one, the first one was intrinsically dissenting and liberal. The complexity of Hegel's language and style masked his true views on politics and religion (see, e.g., Heine, 1856, p. 292). Some Left Hegelians, and some Young Hegelians then (Arnold Ruge in particular), took up that pattern. But according to Ruge, the matter was not simply about reading Hegel's writings adequately; it was also about correcting some specific sections of the system: those that justified the existing state of affairs. By an obvious analogy with the German campaign of 1813, he came to portray his own philosophical outlook as a “liberation war” against three “foreign powers”: the closed philosophical system, the presupposed dogmatics, and the empirical reality (Ruge, 1841, p. 2). Hegel let these powers invade his thinking because in the indolent times when he lived, he could enjoy a privileged position (thanks to such a compromise). So he never suffered repression, and he never had to make a real political stand (see, e.g., Ruge, 1842, p. 761). Then his idea that philosophy must “look behind”. But in the opinion of Ruge, the owl of Minerva did not have to wait for the falling of the dusk to spread its wings. Philosophy, to Ruge's mind, had to be the constructive force of the future; its sclerotic and compromised elements had to be hunted and eradicated,
so that it would no more reject the “ought to be”. The philosophical criticism was to produce the historical movement, and the theoretical activity was to become practical. The philosopher, Ruge wrote, must be “sat at the loom of its time” (Ruge, 1841, p. 3). Hegelian philosophy had to be realised, namely redefined (distinguishing the principles of the implications) and completed.

The second perspective did not dismiss the need for rectifying the system, on the contrary! It claimed that the division between principles and implications was misleading. Bruno Bauer was the first to develop such a position but his friend, in those days, the (young) Karl Marx explained it more clearly and defended it more firmly (on Bauer's hesitations, see Clochec, 2018, pp. 305–306). As he put it in his doctoral dissertation: when Hegel's pupils “explain one or the other determination of his system by his desire for accommodation and the like, hence, in one word, explain it in terms of morality” (Marx, 1841/2010, p. 84). Marx used two arguments against the accommodation hypothesis. The first one consisted of recalling that, a short time ago, some of these pupils still shared the same views; as a consequence they “rendered themselves suspect of not having been serious before” (Marx, 1841/2010, p. 84). The second (and more interesting) argument lied in a criticism of the psychologising aspects of the accommodation hypothesis. It assumed that some content of Hegel's philosophy depended on personal and arbitrary decisions — and this is how we must, here, understand Marx's usage of the term “morality” — in a particular context. So those who defended that kind of interpretation deviated from a real philosophical understanding of the problem: “It is quite thinkable for a philosopher to fall into one or another apparent inconsistency through some sort of accommodation; he himself may be conscious of it. But what he is not conscious of, is the possibility that this apparent accommodation has its deepest roots in an inadequacy or in an inadequate formulation of his principle itself” (Marx, 1841/2010, p. 84). The principles were as problematic as the implications: the conservative positions of Hegel were rooted in the internal logic of his thinking, and the Hegelian system required an inner revision. So the goal was no longer to be consistent with Hegel's words or to reveal a hidden doctrine. Unfortunately, in his dissertation, Marx did not precisely designate the fallacious aspects of the system. The Bauerian influences of his text — the emphasis on self-consciousness in history as opposed
to an abstract and transcendent “absolute spirit” — help us to better understand his outline.

Thus, one of the key aspects of the Bauerian reconfiguration of the Hegelian framework was a renewed dialectic deprived of sublation and affirming the supremacy of the negative against the positive. For Bauer and for most of Young Hegelians, the term “positive” came to designate the fixed and sclerotic elements of the existing state and theories, the persistence of dead and rotted materials in history that get in the way of the development of self-consciousness and that must be purged by criticism — as it discovered the contradictions and the mythical aspects of the gospels (On Bauer's studies of the gospels and the difference between his approach and Strauss's one, see Moggach, 2007). At that time, and as mentioned by Marx in his doctoral dissertation, some philosophers (such as Schelling, whom the Young Hegelians perceived as the main theoretical enemy of Hegelianism) and jurists (such as those of the very conservative German Historical School) highlight the value of the “positive” could only strengthen such a claim. Nevertheless, the struggle against the positive could not affect the mere enemy. Hence the idea that the principles of the Hegelian system could not remain undamaged: “Criticism must therefore direct itself against itself, and dissolve the mysterious substantiality in which up till now it and its object have been contained. The development of substance itself drives it on to the universality and determinateness of the idea and to its real existence, to infinite self-consciousness” (As translated by Moggach, 2007, p. 75).

2. Theorising the refusal of compromise

2.1. Moderates and “juste milieu”

The relationships with the moderates had been a determining element throughout the evolution of the Young Hegelian movement. One of their first attitudes lay in a defence of the whole school and an attempt to radicalise it. The article of the (young) Friedrich Engels untitled “Diary of a Guest student” offers an example of this kind of position (Engels, 1842/2010). In this report of some lectures given by the “Old” Hegelians Philip Konrad Marheineke and Leopold von Heinning, Engels highlighted their criticism of Schelling in order to give an impression of unity in the face of the enemy. At the same time, he transposed the words of these professors in a Young Hegelian vocabulary in order to make their
remarks sound more radical. It should be noted that, in May 1842, when this article was published, Marheineke, who had been the master and the collaborator of Bruno Bauer, defended his former student and tried to prevent his dismissal. But soon, this support transformed into criticism. “Old” Hegelian did not agree, either with the Young Hegelian views or the idea of a common front, and they made this clear. The Young Hegelians rupture with the Hegelian School hastened as a consequence. During the summer of 1842, the radicalisation process becomes clearer and more conceptualised. It also exceeded the framework of the Hegelian School.

Edgar Bauer (Bruno Bauer's brother) introduced an important concept in that sense. He considered that all politics of conciliation stumbled into the pitfall of the *juste milieu* [middle way politics]. Gradually, this term became widely used in Young Hegelian literature. It was generally quoted in French; it had indeed first been employed in France to characterise the July Monarchy. It had been popularised by French Press caricaturists; they had made the expression a kind of motto of Louis-Philippe's regime, which both rejected absolute monarchy and revolutionary republics. In such a position, Edgar Bauer and Young Hegelians denounced a *lack of principle*, which undermined the antagonism necessary for any historical development. Bruno Bauer's negative dialectics was the evident background of such a claim. Obviously Edgar Bauer did not just target the French regime. For example, he attacked the German liberals (such as Karl von Rotteck) who mainly defended, in the French Revolution, the achievements of the National Constituent Assembly. To Bauer's mind, those who rejected the Revolution as a whole were more consistent; at least, they perceived the real principle of the revolution and they spurned it. Edgar Bauer also wrote an entire book entitled *Bruno Bauer and his enemies* to advocate for his brother's against both the Old Hegelians and the anti-Hegelian establishment, following the same argumentative pattern (Bauer, 1842). Two kinds of enemies were identified: the real and consistent enemies and the moderate false friends. He assumed that the best way to refute both of them was quite simple: revealing the antagonism in its purity and leading to its summit the principle of the negative.

But how can one know whether this principle is the good one? How to know if it really tends to the universal? Isn't, by definition, the negative a particular
opposed to another particular (the positive)? Here again, Edgar Bauer answered such questions in the light of his brother reading of Hegel’s philosophy of history. For Bruno Bauer, as a matter of fact, “the universal must not be treated as a transcendent or hypostatised entity or force [...] because this would be to impose a new form of heteronomy on subjective action. A universal perspective is opened rather when we reflect on the entire course of history as a process of alienation and struggles for freedom. [...] The essence of liberation is the extrication of oneself from existing relations of domination, and the unrelenting struggle against their objective forms.” (Moggach, 2006, p. 120) In this process, self-consciousness has no fixed end, it is rather autotelic. To Bauer's mind, infinite self-consciousness was defined by an opposition to any particular interests and especially, of course, all particular interest that pretends to be universal (as political or religious forms of tutelage), namely the positive. In dialectical terms, “the perspective of infinite self-consciousness is that of a determinate negation, grasping the implicit possibilities and contradictions at the heart of the present and undermining all efforts to sustain the existent against the progress of reason and freedom. It thus fulfils the millennial struggle for emancipation that is the essence of history.” (Moggach, 2006, p. 123) And that is why it corresponds to the realisation of freedom, and to concrete universality.

In *The Reaction in Germany*, Mikhail Bakunin as a young Hegelian, defended brilliantly that kind of position in the light of Hegel's logic. He tried to show that all compromise positions and all “juste milieu” postures were nothing more than tepidity and mediocrity, and that they were anti-dialectical insofar as they implicated the adoption of a median position, which might never be confused with an authentic dialectical mediation. As the compromise tries to deflate conflicts, it avoids the “oppositions” at work in history being unveiled in their purity and transformed into “contradictions” (see Angaut, 2007; García, 2018). In other words, compromise is inconsistent and inconsequential; it is just a manner of preserving the status quo. From this point of view, the problem with such a gesture is that it presupposes that the opposites are equal in right. It is true that both are unilateral — and that none of them, as such, can claim to be fully universal — but one of them is preponderant because, as the moving principle, it can bring to universality: the negative.
2.2. Theory of terrorism and politics of love

In one aspect, Bakunin and Edgar Bauer’s perspectives differed from Bruno Bauer’s. They consider that the struggle against the Reaction cannot be merely theoretical and intellectual. They were some of the most prominent supporters of a translation of Young Hegelianism into a real praxis — while Bruno Bauer still considered that theory was in itself “the strongest praxis” (Tomba, 2006, p. 94). Nonetheless, we can notice an interesting difference between Edgar Bauer’s and Bakunin's perspectives. The first one came to promote a theory of terrorism and the second one a politics of love.

The defence of the supremacy of the negative led Edgar Bauer to look upon the French Terror as a practical model. He gave, in that respect, a provocative answer to the moderate liberals that drew on the legacy of the National Constituent Assembly. Nonetheless, he also claimed: “it is certain that every principle that newly steps into world history is vandalistic. And it is vandalistic because it must stride forward to its most extreme form” (As translated by Luft, 2006, pp. 143–144). If Bruno Bauer alleged the destructive powers of criticism, his brother Edgar came to insist on the critical powers of destruction. “The new principle is Sansculottish. Truth is Sansculottish” (as translated by Luft, 2006, p. 144), he affirmed. And this principle could even justify violence. As Eric van der Luft put it, for Edgar Bauer, “any measure of violence or arrogance was permissible toward the goal of overthrowing the monarchy” (Luft, 2006, p. 144). However, Edgar Bauer's views remained marginal in the Young Hegelian movement; he was in all likelihood the only Young Hegelian to go so far. And note that, in fact, that apology of violence remained purely rhetorical. But it produced a real debate in the movement (see, e.g., the closing of Ruge, 1842).

Bakunin too considered, as a partisan of the negative, that destruction was necessary. The ending (and most famous) phrase of The Reaction in Germany is undoubtedly: “the passion for destruction is a creative passion, too!” (Bakunin, 1842/2015). But contrary to Edgar Bauer, Bakunin tried to think how politics without compromise could avoid being a politics of hate. For Bakunin, true love could only be experienced and practised by the partisans of the negative. Because of their defence of a unilateral position that could not become universal, consistent positivists were only capable of hatred against their opponents. The
only kind of love they could feel was entangled in the false universal provided by
Christianity. But Bakunin alleged that it was not real love because it was not
rooted in the pursuit of universality. The partisans of the negative, as they
belonged to a party — and, as such, took a unilateral position — could feel hate
too. But as they were to realise the concrete universal, they could experience a
true and regenerated love which could even be a love for their enemies. The
partiality to which the negative was forced tended unceasingly to abrogate itself
in the universality of love. That's why for the partisans of the negative, love was
more than a feeling, it was a duty.

3. Conclusion

More or less consciously, we often represent radicalism in a topological way;
on the political spectrum, the radical dwells in the extremes. Even if we use a
more complex analytical perspective on political ideologies, we tend to define
core and peripheral groups, concepts, etc. And the radical is always situated on
the outskirts. In that respect, we can easily attest that Young Hegelianism was
born of a radicalisation process. Young Hegelians, nevertheless, provided a
completely different representation of radicalism, putting it at the centre of the
historical process.

For Young Hegelians, and to quote Marx's “Contribution to the Critique of
Hegel's Philosophy of Law. Introduction”: “To be radical is to grasp the root of
the matter” (Marx, 1844/2010, p. 182). In that way, radicalism is not just a
“topological” situation, it is a matter of truth; and obviously, from a Hegelian
point of view truth does only make sense in history. Hence, the historical
necessity of rejecting all kinds of compromise. But here appears a major
theoretical issue: that of the relations between philosophy (and its truth) and
politics. It is hardly surprising, therefore, that one of the major theoretical
problems faced by Young Hegelians as they get involved in politics was that of the
end (or more precisely the exit [Ausgang] from) philosophy.

References

dehors. Lyon: ENS éditions.


ENACTING LEVINAS’S INFINITE RESPONSIBILITY AS AN ETHICO-POLITICAL COMPROMISE

IMPLEMENTANDO A RESPONSABILIDADE INFINITA DE LEVINAS COMO UM COMPROMISSO ÉTICO-POLÍTICO

Julio. A. Andrade*

julesjnr@gmail.com

Abstract. Levinas’s work does not offer us an ethical theory but seeks rather to describe a pre-originary ethical encounter with the other. Within this face-to-face encounter with the other, my subjectivity is held hostage because of an originary asymmetry between us. This ethical asymmetry produces an infinite responsibility to and for the other, in order that the singularity of the other be preserved. In order to moderate such a demanding position Levinas introduces the third party who restores justice by permitting ethical calculation. This marks a move from ethics to politics. Nonetheless, there remains a lacuna between ethics and politics. I argue for a reading of Levinas’s claim that the third party is an incessant correction of the asymmetry of proximity in order to posit infinite responsibility as the compromise of ethics with politics. I discuss some implications for business ethics, in particular CSR, in light of these findings.

Keywords: Levinas, ethics, politics, infinite responsibility, business ethics, CSR.

Sumário. A obra de Lévinas não nos oferece uma teoria ética, antes procura descrever um encontro ético pré-originário com o outro. Neste encontro face a face com o outro, a minha subjectividade torna-se refém por causa de uma assimetria original entre nós. Esta assimetria ética produz uma responsabilidade infinita pelo e para o outro, de modo a que a singularidade do outro possa ser preservada. Para moderar uma posição tão exigente como esta, Lévinas introduz um terceiro que restaure a justiça ao permitir o cálculo ético. Isto assinala uma passagem da ética para a política. Ainda assim, permanece uma lacuna entre ética e política. Defendo uma leitura da pretensão de Lévinas segundo a qual o terceiro é uma correcção da assimetria de proximidade de modo a afirmar a responsabilidade infinita como o compromisso da ética com a política. Discuto algumas implicações para a ética dos negócios, em particular CSR, à luz destas descobertas.

Palavras-chave: Levinas, Ética, política, responsabilidade infinita, Ética dos negócios, CSR.

* Department of Philosophy, University of Stellenbosch, South Africa.
0. Introduction

Perpich (2008; 3), following Bernard Williams in *Ethics and the Limits of Philosophy* (1985), describes the Enlightenment’s moral question ‘what ought I to do?’ and Ancient Ethics’ moral questions “How should one live?’ and ‘What is the best life for human beings?’, as the “end points of a continuum along which normative ethical enquiry may run, depending on whether it is individual actions or the shape of a whole life that is most at issue.” As a straightforward answer to these ethical questions, Emmanuel Levinas’s notion of an infinite responsibility to, and for the other makes little sense. Levinas does not seek to outline an ethical system or offer any normative guidelines, but rather the essence of the ethical relation in general, or what Critchley (1999; 3) has called the “primordial ethical experience”.

This primordial ethical experience is to be found in the face-to-face encounter with the other, wherein the other holds my subjectivity hostage because of an originary asymmetry between us (Levinas 1969). This ethical asymmetry produces an infinite responsibility to and for the other, in order to preserve the singularity of the other. However such a demanding ethics needs moderating. Levinas (2001; 165-6) himself acknowledges that if only the other and I existed then

I would have had nothing but obligations! But [...] there is always a third party in the world: he or she is also my other, my fellow. Hence, it is important to me to know which of the two takes precedence [...] must not human beings, who are incomparable, be compared? Thus justice here, takes precedence over the taking upon oneself of the fate of the other”.

Justice is, for Levinas (1988; 171), “a calculation, which is knowledge, and which supposes politics; it is inseparable from the political. It is something which I distinguish from ethics, which is primary.”

Levinas’s insistence on the primacy of ethics is to safeguard the singularity of the other in this process of reducing the Other to the same (which he calls ‘totalisation’) in order that the other can be compared with other others and a response calculated. Critchley (1999; 223) remarks that for Levinas, “ethics is ethical for the sake of politics – that is, for the sake of a new conception of the organization of political space”. In other words, being infinitely responsible for the other means that we must strive to preserve the singularity of the other in the
very processes and spaces where the singularity of the other is in danger of being effaced. What follows from this is that “between ethics and politics there is a ‘hiatus’, a ‘lacuna’ and yet ethics calls for politics [...] ethics and politics call for each other on the basis of their non-relation” (Raffoul 1998; 280).

In this paper I will attempt to make the case for an understanding of Levinasian infinite responsibility as the compromise between ethics and politics. Central to the idea of a compromise is the necessity of concessions. Levinas (1998; 158) says that “the relationship with the third party [the political] is an incessant correction of the asymmetry of proximity [i.e. the face to face encounter]”. I will argue that this correction can be understood as the concession the ethical must make to the political in order to safeguard the singularity of the other in the process of recognizing and respecting the singularity of the other in the presence of the third party.

The paper proceeds as follows: I start by outlining Levinas’s ethics, focusing on his notion of infinite responsibility as a response to the ‘proximity’ of the other. The section thereafter examines how Levinas’s introduction of the third party to the face to face encounter facilitates his move to politics. I will then argue for infinite responsibility as the compromise between the ethical and the political.

I conclude the paper in a section that transposes the discussion to the field of business ethics. In particular I discuss the recent ‘political turn’ in CSR (corporate social responsibility).

1. The other and infinite responsibility

Traditional morality has it that answering standard ethical questions such as ‘why be good?’ or ‘what is the right thing to do?’ requires reasons and arguments to support any moral claims made. Arriving at these reasons requires, inter alia, an impartialism between two or more moral agents and their means. Impartialism in turn amounts to “a perspective of impartiality, impersonality, objectivity, and universality” (Alford 2001: 149) between moral agents. Levinas (1969; 39) characterizes the impartialist perspective as an “imperialism of the same”. What Levinas means by this is that impartialism seeks to reduce everything that is different, or Other, to the same, so that they can be systemized into a body of knowledge and spoken about. Levinas calls this reduction of the
Other to the same a ‘totalization’. What is unique and singular about another individual is lost when we try to ‘thematize’ or categorize the Other. Levinas refers to the singularity of the other as the alterity of the other. Levinas argues that the only way to maintain the alterity of the other in the totality of knowing, that is, to avoid the reduction to the same, is to posit the other as infinite; hence the title of Levinas’s major work, *Totality and Infinity* (1969). The subtitle, ‘An Essay on Exteriority’, locates the Other, as infinite, outside of this totality. How is this possible? Levinas (1969; 49) adapts Descartes’ third meditation, zeroing in on the idea of infinity, which is exceptional, he argues, in that “its ideatum surpasses its idea”. Levinas continues: “the distance that separates ideatum and idea constitutes the content of the ideatum itself”, and as such, the other “is the sole ideatum of which there can only be an idea in us; it is infinitely removed from its idea, that is, exterior, because it is infinite” (ibid.).

Levinas takes this structure as a model for the alterity of the other. The other’s mode of presenting herself, “the exceeding of the idea of the other in me”, Levinas calls ‘face’ (50). Levinas goes on to say that “the being that presents himself in the face comes from a dimension of height” (215). In other words, the face to face encounter is marked by a radical asymmetry. The reason for this asymmetry is that any attempt by me to represent the other would be a first step towards her totalisation. I can never be certain that my representation of the alterity of the other matches her presentation of her alterity as face. I can never be certain what the other requires of me. This asymmetrical relationship I have with the other, “who cannot be resolved into ‘images’ or be exposed in a theme”, Levinas labels ‘proximity’ (1989; 89).

In *Substitution*, Levinas (1989; 90) ties proximity to exteriority (and *Totality and Infinity* to *Otherwise Than Being*):

> The relationship of proximity cannot be reduced to any modality of distance or geometrical contiguity, nor to the simple ‘representation’ of a neighbour; it is already an assignation [...] an obligation anachronously prior to any commitment. This anteriority is ‘older’ than the a priori [...] The relationship with exteriority is ‘prior’ to the act that would effect it.

---

1 ‘Other’, with a capital, will refer to otherness in general, or alterity; while ‘other’ will refer to the personal other, the other person.

2 An ideatum is the external object around which an idea (or concept) forms. In the case of the idea of infinity, the ideatum is the paradoxical object which must both include and exceed itself.
But how is proximity already an assignation? How can the subject be affected “without the source of the affection becoming a theme of representation”? (ibid.). Perpich (2009; 87) rephrases these questions as “how does the face of the other command me or concern me?” We can also ask, ‘How can the face signify the ethical demands of the other?’ Levinas (1985; 85) answers that “access to the face is straightaway ethical”, “the face is signification, and signification without context” (86); that is, without categories and representative themes. Elsewhere, Levinas (2001; 48) says that the face “is an appeal or an imperative given to your responsibility: to encounter a face is straightaway to hear a demand and an order”.

This summons to responsibility by the face just means that my subjectivity is hostage to the other: “the I is a privilege and an election [by the other]” (1969; 245), which is in turn, one more formulation of Levinas’s definition of ethics as “the calling into question of my spontaneity by the presence of the other”(43). Does this mean, as Alford (2004; 153) asks, “that ethics is better than freedom?” No, what this means is that my autonomy is validated in responding to the other (ibid.). Levinas calls this, ‘freedom invested’ (1969; 302-4), which Perpich (2008; 91) explains as follows: “my freedom is put into question and ‘invested’ as goodness.”

I am because I am called by the other, and thus I am infinitely responsible to the other. Infinite responsibility arises because of the asymmetry between the other and I in the face to face encounter. I try to faithfully represent the other and her demands without reducing her into a totality of like and same but only the other can present her alterity, as face. But infinite responsibility goes further, it also means that “I am responsible for a total responsibility, which answers for all the others and for all in the others, even for their responsibility. The I always has one responsibility more than all the others “(1985; 99).

At this point, Levinas will concede that these are “extreme formulas which must not be detached from their context” (ibid.). The compromise with politics is put forward: “in the concrete, many other considerations intervene and require justice even for me. Practically, the law sets certain consequences out of the way” (ibid.). It is, however, important to note that this compromise with politics does
not mean that Levinas gives up his “utopian and, for an I, inhuman conception”; he still holds that “subjectivity as such is *initially* hostage; it answers to the point of expiating for others” (100). (Emphasis added).

Next, I explore the parameters of the intervening practicalities and then expand on the Levinasian ethico-politico compromise.

**2. The third party and ethico-politico compromise**

To recall the citation in the introduction, Levinas (1998; 104) acknowledges that we “don’t live in a world in which there is but one single ‘first comer’; there is always a third party in the world: he or she is also my other, my fellow.” The import of this is that, in the words of Alford (2004; 156), “the introduction of the third, saves us from being consumed by the infinite need of the other.” This is because I am too an other; I am other to another other. I can restrict the infinite demands made on me by the other by pointing to the other other (including myself) who requires my attention, failing which I fail to respect their alterity. However, we are still in a quandary, because while “theoretically my obligation to each is infinite; practically that cannot be. How to parse the difference?” (164).

Alford remarks that Levinas will not answer this because that would require some kind of universal ethics (or impartialism) (ibid). Levinas himself has said his “task does not consist in constructing ethics; I only try to find its meaning” (1985; 90). Levinas’s terminology can be confusing at this point; by ‘constructing ethics’, he means normative guidelines, which in return requires calculation. Critchley (1999; 3-4) notes that Levinas sometimes speaks of this distinction, between constructing ethics and ethics’ meaning, in terms of the difference between the ethical and the moral (although he is not consistent on this point), where the latter refers to ‘the socio-political order of recognizing and improving our human survival’ which is itself founded upon the *prima philosophia* of an ethical responsibility to the other.

In *The Ethics of Deconstruction*, Critchley (ibid.) stipulates that he will use a “distinction between ethics and *politics*”, where politics refers to the moral (or ‘constructing ethics’). I will follow Critchley’s convention.

The central way in which Levinas ‘parses the difference’ between infinite responsibility in theory and practice is through temporality. Levinas (1969; 213) says that the third does not arrive *after* the other in the face to face encounter but
is already there in the encounter with the other: “It is not that there first would be face, and then the being it manifests, or expresses would concern himself with justice; the epiphany of the face quae face opens humanity”. That is to say, each face, representing the singularity of the other, recalls the singularity of every other other, including myself. As such, remarks Jordaan (2009; 97), “every person I come across is both the other and the third to me. Every person I come across is both a general and equal other with whom I stand in a political relation as well as a specific other who summons me to responsibility in the ethical relation”. The Other and the third are coterminous and contemporaneous.

And yet, as noted before, I can only respond to any one demand at a time, I can only face one other at a time. Jordaan (ibid) reframes this dilemma we have already spoken about:

When faced by the other, I can respond politically, seeing the other as my equal, restricting my responsibility to him, insisting on a reciprocity and equality between us, and asserting my rights against his; or I can respond ethically by being concerned and assuming responsibility for him beyond what is required by our political equality and reciprocity. Do I relate to the other ethically or politically? I am constantly faced with this choice. (Emphasis added)

While Jordaan in the main captures what Critchley (1999; 225) calls the “double structure of community”, his use of the disjunction ‘or’ can be misleading because, Critchley continues, the “passage from the ethical to the political is not a passage of time, but rather a doubling of discourse whereby a response to the singularity of the Other’s face is, at the same time a response to community [as] a commonality”, that is, as a group of equals.” To put it simply, as Wolff (2011; 21), “ethics necessarily passes into politics, or in fact, that ethics has always already passed into politics”. Politics can be said to be the enactment of the Levinasian primordial ethical relation while ethics is the thinking of the Levinasian primordial ethical relation (Woermann 2016; 133). Politics can thus be characterised, in the words of Critchley (2004; 178), as “the art of a response to the singular demand of the other”. That is to say my response to the demand of the other will of necessity also be a response to the third in that my response will have been a political calculation that takes into consideration its simultaneous effect on both the other and the third.

If this so then, argues Wolff (2011; 25), “‘political responsibility’ is a pleonasm.” My infinite (ethical) responsibility for the other is also a finite
(political) responsibility for the third, or rather my political responsibility for the third is contained in my ethical responsibility for the other, so that “[p]eace, or responsibility, to the near one, the neighbour, is peace to the one far off, the third party” (Critchley 1999; 223). The third, argues Levinas (1998; 157) “is of itself the limit of responsibility and the birth of the question: What do I have to do with justice?” This is why Critchley (1999; 223) claims that “politics provides the continual horizon of Levinasian ethics”. Politics as the limit of responsibility does not imply “ethics [as] necessarily a restraint on power – i.e. one where politics comes first and then ethics evaluates it later – [rather] politics is the machinery through which the ethical demand can be responded to” (McMurray et al. 2010; 546).

It is at this point that I want to make the case that politics as the ‘art of a response to the Other’, as the ‘machinery’ of the ethical, as the ‘horizon and limit’ of ethics, can be made sense of by a particular reading of Levinas’s (1998; 158) remark that “the relationship with the third party is an incessant correction of the asymmetry of proximity.” (Emphasis added). By now it is a quick step to note that ‘the relationship with the third party’ just refers to the political, while ‘proximity’, to recall, refers to my relationship with the Other, “who cannot be ...exposed in a theme” (1989; 89), that is, the primordial ethical relationship of the face to face. Substituting these phrases into the citation above then produces the formulation ‘the political is the incessant correction of the ethical’. It is this ‘incessant correction’ that I want to characterize precisely as a compromise.

The OED defines a compromise(n) as both a settlement reached by making concessions, and ‘an intermediate state between conflicting alternatives’. As the (Levinasian) ethical calls for an infinite responsibility, such a call can only be realized in practice as a compromise, as a concession of my resources and abilities between the Other and the third. Levinas (1969; 244) himself says “infinity of responsibility denotes not its actual immensity, but a responsibility increasing in the measure that it is assumed”. Perpich (2008; 89) takes this to mean that

the better I accomplish my obligations, the more demands I find addressed to me. It is not a matter of the actual number of demands increasing, but a matter

---

3 See https://en.oxforddictionaries.com/definition/compromise
of my sensitivity increasing so that the demands and injustices of which I was formerly unaware now come to press and weigh on my conscience.

In other words my infinite responsibility is incessantly being corrected by my response-ability to the other, which is also a simultaneously a response to the other other.

One citation in the introduction had it that between “ethics and politics there is a ‘hiatus’, a ‘lacuna’ (Raffoul 1998; 280). It is my reading that infinite responsibility should be understood exactly as this lacuna between ethics and politics. A hiatus/lacuna can in turn be understood as an ‘intermediate state’, that is, as a compromise (forming the second part of the definition of compromise used above). Levinasian infinite responsibility can thus also be understood as an ethico-politico compromise.

3. Some implications for Business Ethics

After the above explication of the compromise of ethics and politics in Levinas, I now move to considering some implications such an understanding might have for the sub-discipline of business ethics. The compromise required between ethics and politics is succinctly posed in the question addressed to business ethicists by Martin Parker (2003; 198) in ‘Ethics, Politics and Organizing’. He ask business ethicists to consider “why they don’t call their discipline ‘business politics’?” instead of business ethics. In not so doing argues Parker (189), they justify themselves as “properly concerned with ‘ethics’ and [so] avoid the troubling pro-managerial hegemony of the wider discipline”. In turn they use the term ‘ethics’ to frame questions and answers in generally individualistic ways which tend to exclude any sustained examination of the organizational contexts and power structures individuals find themselves operating in (ibid). Thus, Parker (197) continues, “because business ethics asks ‘how can people in organizations be good?’ it does not often ask ‘how can business

---

4 ‘Applying’ Levinas to business ethics is a popular activity, however, paper after paper within this literature includes the inevitable caveat that doing so is a fraught enterprise because, inter alia, applying ethics is just another form of totalisation (see for example the articles in the 2007 special edition on Levinas and business ethics in Business Ethics: A European Review, 16(3)). As the move from ethics to politics I discuss in this paper is considered to be a rubric to ‘applying’ Levinasian ethics, I believe I have heeded the necessary caveats.
be part of a good polis?” I will briefly explore this concern through considering a central business ethics issue – corporate social responsibility (CSR).

Discussions of CSR typically start by contrasting the views of Milton Friedman on the one hand and proponents of stakeholder theory on the other (Jones et al. 2005; 97). The former’s views can be directly gleaned in the title of his seminal 1970 essay ‘The social responsibility of business is to increase its profits’; while the latter grouping is usually represented by Edward Freeman (1984) who argues that business has a responsibility to a wider set of stakeholders beyond the shareholders, and thus profit maximization is not the only, or even most important responsibility of business.

Friedman’s position is often misrepresented and taken out of context, because although advocating for business to make as much money as possible, he emphasizes that this must be done “while conforming to the basic rules of society, both those embodied in law and those embodied in ethical custom” (1970; 33). This common misrepresentation notwithstanding, Jones et al. (2005; 99) nonetheless give a very charitable reading of Friedman. They argue that Friedman does not reject legal and ethical responsibility because he thinks it unimportant, but because business operates within a particular context and “that context – civil society and the political institutions it generates and ethical expectations it produces – is not and should not be a matter of corporate control or governance” (ibid.).

The CSR literature is vast, with several splits occurring over the decades since Friedman’s article. It is not my intention to critically analyse this literature or to even provide an overview beyond the two opposing positions outlined above. Instead, I touch on what Scherer et al. (2016) call the ‘political turn’ in the CSR literature during the past decade. Part of the reason for this interest is the increasing scope of globalization which has in turn increased the power of the MNC (multi-national corporation) both within and across national borders (Scherer et al. 2009). As such the societal context in which business operates, and the ethical expectations this produces, has changed dramatically since Friedman wrote about business in 1970. Business provides more and more public goods

---

5 For a comprehensive overview see Agle et al. (2008).
(water, communications, education) and shapes public policy – Facebook and media for example. A corporate responsibility to ensure that the privacy concerns of its customers is honoured has turned into a political responsibility to ensure that nefarious agents do not subvert democratic processes. The recent Facebook imbroglio is just the latest and most urgent exemplar of why PCSR (political CSR) is needed.

Scherer et al. (2016; 275) note that the term PCSR brings together two highly contested concepts – the ‘political’ and ‘CSR’ – and as such PCSR should be understood as an ‘umbrella concept’. Scherer et al. (ibid.) cite Frynas and Stephens (2015; 485) definition of PCSR “as activities where CSR has an intended or unintended political impact”. However, their definition is rejected precisely because they fail to define the ‘political’. It seems then, that what the political entails cannot be bracketed from CSR, even under an umbrella concept such as PCSR. Scherer et al. (2016; 276) argue that “public deliberations, collective decisions, and the provision of public goods” are key features of the political and as such define PCSR as “responsible business activities that turn corporations into political actors, by engaging in public deliberations, collective decisions, and the provision of public goods or the restriction of public bads” (*ibid.*).

My intention is not to challenge such a definition of what constitutes the political and thus to demarcate a sphere of enquiry that would circumscribe PCSR, if indeed such a concept is theoretically feasible. I note rather that public deliberations and collective decisions constitute a calculation, an attempt to compare and weigh certain categories and themes with others. As such, Scherer et al.’s notion of the political is congruent with what I have argued in this paper is the claim of the third party in conjunction with the other in the face to face encounter. In order to do justice to both the other and the third the I must deliberate and choose, even as the I’s freedom to do so is resisted by the asymmetry of proximity. I have described that dynamic as the ethico-politico compromise that constitutes infinite responsibility. If the claim made in the previous section by Wolff (2011; 21), following Levinas, that “ethics necessarily passes into politics”, holds true, then a fortiori so does the claim, business ethics necessarily passes into business politics. As such CSR will also necessarily pass into PCSR.
4. Conclusion

This paper has put forward the claim that Levinas’s notion of infinite responsibility can be usefully understood as an ethico-politico compromise, where the ethical describes the pre-originary, and asymmetrical encounter with the other, and the political the contemporaneous presence of the third within that encounter which necessitates the need for calculation and comparison. I have indicated just one instance – the business ethics field, in particular the concept of PCSR – in which this finding can be put to use. Future research might productively use this insight for other applied ethics problematics in the field of environmental or information ethics for example, which in both cases will require calculation and compromise.

References


Enacting Levinas’s infinite responsibility as an ethico-political compromise


A REACTIONARY FASCINATION: EMIL CIORAN AND JOSEPH DE MAISTRE

Abstract. The aim of this paper is to examine Cioran’s interpretation of Maistre’s reactionary thought. Cioran judges the philosophical work of De Maistre as a useful instrument to investigate the Twentieth century ideological debate on Revolution, with specific reference to the issue of the engagement of philosophers in politics. Through the analysis of Maistre’s criticism of revolutionary thought at the time of the Enlightenment, Cioran proposes an insightful deconstruction of the ideological dichotomy between Reaction and Revolution. Indeed, the paper will show that Cioran’s understanding of De Maistre implies both a radical criticism of Sartre’s existentialism and an original interpretation of Schmitt’s Political Theology.

Keywords: Cioran, De Maistre, reaction, revolution, ideology.

In 1957, one year since Soviet troops overrun Hungary, Emil Cioran writes an essay where he examines the ideological question of Revolutions. Oddly enough, his writing is called An Essay on Reactionary Thought and is dedicated
to Joseph De Maistre, the most important reactionary thinker at the time of the Enlightenment.

This paper aims to understand why Cioran decides to examine the philosophical works of De Maistre to deconstruct the ideological dichotomy between Reaction and Revolution. Moreover, the paper will show that Cioran’s understanding of De Maistre implies both a radical criticism of Sartre’s existentialism and an original interpretation of Schmitt’s *Political Theology*.

Emil Cioran is a Romanian philosopher who spent most of his life as a stateless man in Paris, where he became one of the finest French writers of the Twentieth century. Anyway, he was almost unknown by the general public when he died in 1995. All his French essays are focused on the necessity of being skeptical, that is, of not being engaged in any kind of philosophical or ideological apology. Cioran is quite clear on this point. If a philosopher is supposed to defend a political or religious system, he is not allowed to doubt its ideological premises: therefore, he would easily fall into a dogmatic position (Cioran, 1995; Di Gennaro, 2014). In this sense, Cioran argues that skepticism is synonymous of disengagement — i.e., of the ability of not seriously believing in any kind of truth. In the *Temptation to Exist*, Cioran defines his own skeptical attitude toward philosophical or political ideologies as a form of futility, while also specifying that “futility is the most difficult thing in the world. (…) Each time I catch myself assigning some importance to things, I incriminate my mind, I challenge it and suspect it of some weakness, of some depravity” (Cioran, 2011b, p. 210-211).

Cioran ironically suggests that futility is an unattainable ideal since faiths, dogmas and certitudes represent a physiological necessity of human existence: indeed, human beings would prefer to “starve to death” rather than to renounce their own certainties and beliefs (Cioran, 2015, p. 34). From this point of view, Cioran’s apology of futility appears to be the confutation of Sartre’s understanding of ideologies. According to Sartre, the more a philosopher is

---

1 Cioran first publishes *An Essay on Reactionary Thought* as a preface of a book dedicated to Maistre that Cioran himself has edited in 1957 (*Joseph de Maistre. Textes choisis et présentés par E.M. Cioran*. Monaco: Le Rocher). The preface is re-printed by the publisher Fata Morgana in 1977, with the title *Essai sur la pensée réactionnaire. Joseph de Maistre*. Finally, in 1986, the Essay will also appear as a chapter of the last book by Cioran: *Exercices d’admiration* (the text is now available in the Critical Edition of Cioran’s *Œuvres* published by Gallimard in 2011). We will follow the English translation made by Richard Howard in *Anathemas and Admiraions* (2012b).
seriously engaged in a ‘good’ ideology, the more he would contribute to the
realization of freedom (Sartre & Kulka, 2007). In other words, Sartre states that
freedom could not be conceivable in the absence of an ideal which establishes the
conditions of possibility of freedom itself. However, human beings are easily
fascinated by the existence of some ‘negative’ ideologies which claim for the
abolition of individual liberty. In this sense, it is fundamental to develop the
capability of recognizing the right kind of ideologies, and to engage themselves in
the fight against the wrong kind of belief systems. According to Cioran the
problem is much more complex, because the ‘goodness’ of an ideal doesn’t
guarantee the goodness of its historical realization. It has often happened that an
ideal of freedom and justice becomes the source of violence and intolerance at the
very moment of its political affirmation. Therefore, the problem is not about the
political/ethical values defended by an ideology; it is about the mechanisms
which allow the passage from a demand of a radical change to the taking of power
of this radicalism (Guerrini, 2002).

In accordance with this problem, Cioran refuses the solution proposed by
Sartre. In a chapter of A Short History of Decay called On an Entrepreneur of
Ideas, Cioran compares the engaged philosopher to a “conqueror” who

(...) has but one secret: his lack of emotion; nothing keeps him from dealing with
anything, since he does so with no accent of his own. His constructions are
magnificent, but without salt: categories swell with intimate experiences, classified
as in a file of disasters or a catalogue of anxieties. Here are ranged the tribulations of
man, as well as the poetry of his laceration. The Irremediable has turned into a
system, even a side show, displayed like an article of common commerce, a true mass
product of anguish. The public delights in it; the nihilism of the boulevard and the
bitterness of the café feed on it (Cioran, 2012a, p. 352-353).

In a manuscript version of the French text, we discover that Cioran has
initially entitled this chapter Le Cas Sartre, revealing us that his portrait of the
“entrepreneur of ideas” is a caricature of Sartre’s existentialism (Cioran, 2011a, p.
1342). The reader has the feeling that the engagement of this entrepreneur is as
false as his existential anguish: a mere product of a philosophical vogue,
generated by the liberal illusion of humankind ruled by the laws of Reason and
Progress (Guerrini, 2002). In this context, “the tribulations of man” are nothing
but a deviation from the path of unlimited social improvement. Cioran judges this
trend as dramatically dangerous, in so far as it makes impossible to see the risks
linked to the ‘utopia’ of humankind freed from any chain. In a letter to his friend
Wolfgang Krauss, Cioran writes that such philosophers as Sartre are not able to understand that an ideal of extreme pureness can justify acts of extreme violence, as it happened during the Chinese and Russian Revolutions. In this sense, he says that such apologists of absolute freedom often turn out to be the most fanatic apologists of the “organized terror” (Cioran, 2014, p. 244-245).

In his Essay on the Reactionary Thought, Cioran wants precisely to explore how an ideal of revolutionary improvement of the world can be overturned into a justification of a dogmatic vision of reality. In its very essence, revolution is a negation of the status quo — a negation of authority. But what happens when a revolution reaches its political goals? What happens when the authority of the old government is substituted by the authority of a new revolutionary government? If a revolution obtains political power, how can it keep its contesting force against the status quo? To answer these questions, Cioran dedicates an essay on Joseph De Maistre, the controversial apologist of political absolutism at the time of French Revolution; the theorist of theocracy at the time of the Enlightenment’s critique of religion. But why?

Actually, Cioran’s choice is not provocative; it rather follows a theoretical reason. According to Cioran, the French revolution represents a “paradigm” to understand the mechanisms of all modern revolutions (e.g., the Communist revolution). Consistently, he also thinks that the failure of the Enlightenment ideals can be a key of interpretation of the failure of contemporary revolutionary ideologies. Maistre is the most vigorous critic of the contradictions of the French Revolution, the one that had predicted the fanatical epilogue of the cult of Reason. His metaphysical interpretation of the revolutionary event is judged by Cioran as a useful instrument to approach the impasses of our political age. Moreover, Cioran undoubtedly finds in De Maistre a kindred spirit, due to his life and his style. On the one hand, Maistre is a thinker who had “the appetite and the genius for provocation”; who loved to exasperate the reader with “paradoxes” and “enormities” (Cioran, 2012, p. 39). On the other hand, given the political consequences of a war, De Maistre spends most of his life in exile experiencing the sense of loneliness and failure, as Cioran does (Cavalcanti, 2005; Paseyro, 2005).
Indeed, De Maistre is a French Catholic nobleman who becomes a member of the Savoy Senate in 1787. After the invasion of Savoy by the French Revolutionary armies in 1792, he begins his lifelong exile in Switzerland and in Russia. As an envoy of the King of Sardinia, he remains at the Russian court in St Petersburg for 14 years, until 1817. But he has very little power and very little influence: a marginal personality far from home. His masterwork, *St Petersburg Dialogues*, is written during this period and it reflects this feeling of rage against the disorders provoked by the Revolution. Basically, De Maistre accuses the French *philosophes* of having undermined the basis of social order by subjecting tradition, religion and authority to the criticism of reason. An overturning of the natural hierarchy of Creation that has triggered the anger of the Creator; an insane overturning of reality whose conclusion is the blood-drenched guillotine. But the Revolution and the Terror are not only the logical consequence of permitting the spread of heretical ideas; they are also an instrument of God to keep the integrity of the French reign until a new Catholic King would regain the throne (Boffa, 1989). In his *Considerations on France* De Maistre asserts the scandalous thesis according to which the Revolution is a divine *miracle* showing that even the freedom of speech is just a word of the Providence monologue. He argues that God created the Revolution to demonstrate that the *superstition* of Christian law is better than the *tyranny* of philosophical reason (Boffa, 1985)

Such a paradox is typical of De Maistre, described by Cioran as an author of pamphlets which destroy the creators of pamphlets, as an *anti-philosophe* who exploits the dialectical weapons of philosophers against themselves (Zobel-Finger, 1975). Cioran says that De Maistre’s “style evokes the image of an Old Testament prophet and of a man of the eighteenth century” (2012b, p. 43). And he observes that:

Toward the end of the [Nineteenth] century, at the height of the liberal illusion, it was possible to indulge in the luxury of calling [De Maistre] the “prophet of the past,” of regarding him as a relic or an aberrant phenomenon. But we — in a somewhat more disabused epoch — know he is one of us precisely to the degree that he was a “monster”; it is in fact by the odious aspect of his “doctrines” that he lives for us, that he is our contemporary (Cioran, 2012b, p. 40).

Cioran recognizes two “monstrous” (or anti-modern) doctrines proposed by De Maistre that should be read as a sign of his modernity. The first one is about a pessimist view of human nature, the second one is about the mechanisms of
conservation of political power. Both are fundamental to understand De Maistre’s criticism of revolutions.

As mentioned above, De Maistre is a Catholic senator who knows perfectly the European juridical tradition and its relation to the law developed by the Church during the Middle Ages. He is not only a witness of the passage from the medieval law to the modern one; above all, he is one of the finest interpreters of the aporias regarding the new paradigm of law. According to modern philosophers such as Thomas Hobbes or John Locke, society is not the creation of a transcendent Creator but the effect of a mutual contract between human beings who originally lived in the State of nature (Fisichella, 2005; Piovani, 1961). De Maistre refutes this hypothesis, because he postulates an absolute identity between State of nature and Society. Human beings are originally created by God as social beings; therefore, there could not exist a State which comes before Society. On the one hand, and from a theological point of view, Society is as natural as Creation is. On the other hand, human beings are as social as sinful. Due to this fact, they are naturally predisposed to do evil things to their neighbors and to ruin common life. The law of God – which keeps humans away from Evil and establishes what is good and what is bad – is the only guarantee of the existence of society and of its historical permanence (Fisichella, 2002).

Cioran observes that De Maistre takes advantage of the doctrine of the Fall to propose a dogmatic and unbearable apology of the Papacy as the institution which represents the law of God on Earth. Nonetheless, he infers two important principles from De Maistre’s theological premises. In the first place, the concept of politics always implies a transcendent dimension, and it does not matter if we call it ‘religion’ or ‘tradition’. In both cases politics represents a reality that is not created by human beings; a reality that could not be an effect either of social resolution or rational planning. In the second place, if religious tradition is the basis of any possible legislation, society could not exist without some ‘sacred’ dogmas about its origins—that is, without some mysterious superstitions about the genesis of Power. From this point of view, De Maistre criticizes the French revolutionists because they claim to do tabula rasa with their past, as if they could create a rational Constitution released from History: a new constitution founded on reason – and not founded on mystery and tradition (Kluback & Finkenthal,
1997). In the French revolutionary perspective, the more a Constitution is rational, the more it is erected on abstract principles: principles which should be as clear and consistent as a mathematical equation. De Maistre blames this statement because Constitutions are a matter of history, and not a matter of mathematics, thus it would be a tragic mistake to judge them according to the principles of pure logic. As a good reader of David Hume, he states that an abstract constitution is unstable and dangerous since it is grounded upon an historical void and exposed to any kind of criticism. In fact, a legislation which is built by reason can also be torn down by reason. In the absence of a strong tradition, any rational dissenter can disavow such a legislation and instigate disobedience or social disorders (Fisichella, 2005).

The last sentence needs clarification. De Maistre is not affirming that constitutions are unchangeable or exempted from historical variations; he is affirming that such variations must not be conceived as a delegitimization of the holy tradition. In this sense, he accuses the Enlightenment of being a ‘satanic’ phenomenon because its rational criticism of authority implies a fragmentation of the unity of power, unity which constitutes the conditio sine qua non of the continuity of tradition (Maistre, 1994). To avoid this fragmentation, De Maistre postulates the infallibility of power: to emphasize his stance, he even compares the Enlightenment to a “universal solvent” which notably dissolves the certainties of tradition in the name of a doubting reason (D’Addio, 1975, p. 59). Notwithstanding these criticisms, De Maistre does not conclude with a condemnation of reason. According to his reactionary philosophy, he deduces that tradition must be the judge of reason, and not vice versa. Reason can be a useful instrument of interpretation of human society. However, human society is founded on tradition, which in its turn is founded on some religious principles. Consequently, De Maistre infers that without religion there can be neither a tradition nor a society. In the tenth chapter of the Dialogues, this idea is emblematically exemplified by the definition of “superstition” made by the Chevalier, one of the characters of the book. After saying that superstition “is something beyond legitimate beliefs”, Chevalier notices that “in this world that what suffices is not sufficient” (Maistre & Lebrun, 1995, p. 305). And he explains:

To jump a ditch, you must always fix your eyes well beyond the bank or fall in. In short this is a general rule; it would be very peculiar if religion were an exception to
it. (…) I imagine, my good friends, that honour could not displease you. Yet what is honour? It is the superstition of virtue, or it is nothing. In love, in friendship, in loyalty in good faith, etc., superstition is pleasing, even valuable, and often necessary; why should it not be the same with the respect to piety? I am led to believe that the outrages against the excesses of the thing originate with the enemies of the thing. Reason is no doubt good, but it is not necessary for everything to be ruled by reason (ibid.).

This apology of religion does not correspond to a cult of irrationality composed by an ante litteram Romantic philosopher. De Maistre is not a ‘Romantic’, he is a philosopher who elaborates a conception of rationality different from the one of the Enlightenment (Fisichella, 2005). In the Maistrian thought, religion becomes the symbol of a universal reason representing a depository of centuries-old norms and habits which guarantees the existence of civilization. De Maistre (1984a) distinguishes the “universal reason” of religion from the “individual reason” of philosophers, by pointing out that the last one would never allow the creation of a peaceful society. Instead of producing shared beliefs and common goals, philosophy produces different opinions and countless occasions of conflict. Therefore, philosophical reason represents an enemy of political life. Against the Enlightenment’s philosophical principles, De Maistre maintains that the authority of facts must be substituted by the fact of authority. At a certain point in time, both religion and politics demand faith and obedience with no rational discussion because reason itself shows that sometimes we must stop reasoning (De Maistre, 1984b, p. 39).

Now, we can understand why De Maistre criticized rational abstraction in politics. Unfortunately, he exploited his criticism to justify a theocracy. Cioran would never accept such a theocratical ideal, but he explains why it is important to reflect on De Maistre’s reactionary thought. De Maistre proves that authority, “to maintain itself, [must] rest upon some mystery, some irrational foundation” (Cioran, 2012b, p. 79). According to Cioran, this political radical truth destabilizes the philosophical dichotomy between ‘left’ and ‘right’, between revolution and reaction, as it shows that “any order that seeks to last succeeds in doing so only by surrounding itself with a certain obscurity, by flinging a veil over its motives and its actions, by generating an aura of the ‘sacred’ that renders it impenetrable to the masses” (ibid., p. 79). In this passage, Cioran points out that any existing government claims to be an infallible order: for this reason, it must be surrounded with obscurity (since the less you can see about authority, the more
you must follow its rules). As soon as it becomes fallible, a government risks being replaced by another form of power. Cioran (ibid., p. 81) observes that this apology of obscurity is a “truism” defended by reactionary thinkers and rejected by democratic ones, though the democrats know that “reaction often translates their hidden thoughts, [...] many bitter certitudes of which they can give no public account”.

During the Enlightenment, De Maistre denounces that les philosophes use the concept of reason in such an obscure way. They are so extreme in their confutation of dogmatism, that they deny the very possibility of being irrational. By doing so, however, they reveal themselves as sneakily dogmatic: in the name of reason, they assume their undoubted rightness. This is a modern aspect of De Maistre’s philosophy, since it discloses how we can be dogmatic by exploiting an anti-dogmatic ideal. And it is precisely in this context that Cioran recalls De Maistre to deconstruct the dichotomy between revolution and reaction. His thesis is that we can be reactionary in the name of revolution, in so far as we absolutize the ideal of revolution. In a central passage of his essay, Cioran writes:

Revolutions start in order to give a meaning to history; such meaning has already been given, replies reaction, we must submit to it and defend it. This is exactly what will be maintained by a revolution that has triumphed; hence intolerance results from a hypothesis that has degenerated into a certitude and that is imposed as such by a regime — from a vision promoted to the rank of truth (Cioran, 2012b, p. 105).

It is important to emphasize that Cioran does not imply that there are no differences between revolution and reaction. Actually, he is strongly convinced that this world needs to be improved, and he does not give up on the idea that the truths of reaction are the last words of our hopes (Cioran, 2015). However, he ironically ascertains the fatality of this improvement: the natural degeneration of the hope in a different world into the static certitude of a well-defined reality — a reality with no escape routes. Shortly before the passage we just quoted, Cioran suggests that “only the pre-revolutionary condition is truly revolutionary”. And he provocatively affirms that “every anarchist conceals, in the depth of his rebellions, a reactionary who is awaiting his hour, the hour of taking power, when the metamorphosis of chaos into . . . authority raises problems no utopia dares solve or even contemplate without falling into lyricism or absurdity” (Cioran, 2012b, p. 102).
This last sentence recalls a passage from Carl Schmitt’s *Political Theology* (originally published in 1922), where the German philosopher examines the absolutist character of power and the question of sovereignty. It is interesting to note that in 1950, seven years before writing his essay on reactionary thought, Cioran sends two letters to Schmitt where he also discusses De Maistre, an author that both Cioran and Schmitt have profoundly studied. In these letters, Cioran thanks Schmitt for having sent to him the manuscript of his last book *Ex Captivitate Salus* and a copy of his essay on the counterrevolutionary thinker Donoso Cortés (*Donoso Cortés in gesamteuropäischer Interpretation*). Besides, he confesses his astonishment for having found so many affinities between his own skepticism and Schmitt’s philosophy. Especially, Cioran mentions their passion for Max Stirner and their obsession with Kleist’s suicide, and he also observes that Schmitt’s distinction between “friend and enemy” should be a founding principle of any political theory (Bojanić, 2009; Pozzi, 2016). Moreover, Schmitt defends an idea which will be central in Cioran’s thought, namely that the exception is more important than the general norm (Nicoletti, 1990). At the end of the first chapter of *Political Theology*, Schmitt writes:

> Precisely a philosophy of concrete life must not withdraw from the exception and the extreme case but must be interested in it to the highest degree. The exception can be more important to it than the rule, not because of a romantic irony for the paradox, but because the seriousness of an insight goes deeper than the clear generalizations inferred from what ordinarily repeats itself. The exception is more interesting than the rule. The rule proves nothing; the exception proves everything: it confirms not only the rule but also its existence, which derives only from the exception. In the exception the power of real life breaks through the crust of a mechanism that has become torpid by repetition (Schmitt, 2005, p. 15).

Indeed, Schmitt defines sovereignty as the decision on the state of exception; a decision which can be made only by the sovereign (ibid., p. 5). The problem is that such a decision appears to presume an absolutistic character of sovereignty (an idea notably contested by democrats and liberal philosophers). For this reason, in the fourth chapter of *Political Theology*, Schmitt focuses his attention on De Maistre and his theory of absolutism. Given that the “relevance of the state rested on the fact that it provided a decision, [and] the relevance of the Church on its rendering of the last decision that could not be appealed”, De Maistre establishes a perfect reciprocity between “sovereignty” and “decision” which entails a conformity between secular power and religious power (Schmitt, 2005, p. 55). In order to be efficient, the decision must be infallible; but to be
infallible, it must be absolute. According to Schmitt, De Maistre is stating that *every sovereignty is infallible because every sovereignty is absolute* (ibid.). Obviously, this statement should be applied to any kind of government, regardless of its political structure (democracy, aristocracy, monarchy, etc.), Schmitt suggests that even

(...) an anarchist could pronounce [this sentence] verbatim, even if his intention was an entirely different one. (...) All the anarchist theories from Babeuf to Bakunin, Kropotkin, and Otto Gross revolve around the one axiom: “The people are good, but the magistrate is corruptible”. De De Maistre asserted the exact opposite, namely, that authority as such is good once it exists: “Any government is good once it is established”, the reason being that a decision is inherent in the mere existence of a governmental authority, and the decision as such is in turn valuable precisely because, as far as the most essential issues are concerned, making a decision is more important than how a decision is made. (...) In practice, not to be subject to error and not to be accused of error were for him the same. The important point was that no higher authority could review the decision” (ibid., p. 55-56).

Starting from the same principle of absolutism of power but with the aid of “his axiom of the good man and corrupt government”, Schmitt explains that the anarchist draws the opposite practical conclusion of De Maistre: namely, “that every government must be opposed for the reason it is necessarily a dictatorship” (ibid., p. 66). But this radical opposition to any kind of authority forces him to decide dogmatically against decision because every claim of decision is Evil. Therefore, decision must be forbidden. A conclusion which necessarily entails a contradiction, resulting “in the odd paradox whereby Bakunin, the greatest anarchist of the Nineteenth century, had to become in theory the theologian of the antitheological and in practice the dictator of an anti-dictatorship” (ibid.).

This paradox noticed by Schmitt is one of the cornerstones of Cioran’s essay, and it is the reason why the Romanian philosopher esteems De Maistre as a fundamental key of interpretation of the Twentieth century ideological debate. While examining “the metamorphosis of chaos into... authority”, Cioran must face the same problem encountered by Schmitt in his analysis of the state of exception, namely the problem of decision and of conservation of power. Schmitt has underlined that neither the anarchist nor the reactionary can avoid the responsibility of decision in the moment of conflict. In his turn, Cioran shows that if a revolution triumphs, it could not be anymore an exception from political rules since it stops to be a negation of power: it becomes power, and it will operate to legitimate its own authority, as a reactionary regime would do.
In this sense, Cioran talks of the reversibility of ‘the revolutionary sentiment of time’ into ‘the reactionary sentiment of time’. Revolutionaries believe that the absolute perfection will be realized in a forthcoming future, whereas the reactionaries believe that this perfection was realized in a remote past. Nonetheless, Cioran remarks that this distinction would be quite superficial if we do not focus our attention

(...) on the most curious thing of all: the revolutionary idolizes Becoming only up to the instauration of the order for which he fought; subsequently, for him, appears the ideal conclusion of time, the Forever of Utopias, an extratemporal, unique, and infinite moment, provoked by the advent of a new age, entirely different from the others, an eternity here on earth that closes and crowns the historical process. The notion of a golden age, the notion of paradise pursues believers and unbelievers alike (Cioran, 2012b, p. 77-78).

The notion of paradise also pursues revolutionaries and reactionaries alike. It is precisely for this reason that Cioran reflects on De Maistre philosophy: to explain better how both a revolutionary hope and a reactionary hope can convert themselves into a fanatical conviction. Furthermore, Cioran finds in the Maistrian thought a proof of the urgency of being skeptical — that is, of being disengaged. Differently from Sartre or from other engaged philosophers, Cioran confesses the disappointment that revolutions “generate in all who have believed in them with some fervor” (ibid, p. 104). However, this confession does not mean that Cioran is inviting us to think as a dogmatic reactionary. On the contrary, it is an invitation to reflect on the anthropological optimism that has allowed the proliferation of modern utopian illusions: an optimism originally developed during the Renaissance age, and then exploited by the Enlightenment to justify its conception of human history as an history fulfilling the requirements of a pure reason.

In History and Utopia, Cioran explains that all modern utopias are marked by the heretical doctrine of Pelagius, the theologian who had denied the effects of the Fall and who had theorized that human beings are born good and free, with no trace of an original corruption:

It is hard to imagine a doctrine more generous and more untrue; this is a heresy of the utopian type, fruitful by its very extravagances, by its absurdities which were rich in futures. Not that the authors of utopias took their inspiration from it directly; but it is incontestable that in modern thought there exists, hostile to Augustinianism and to Jansenism, an authentic current of Pelagianism—the idolatry of progress and all revolutionary ideologies will be its conclusion—according to which we constitute a mass of the virtual elect, emancipated from original sin, infinitely malleable, predestined to the good, capable of any and every perfection (Cioran, 2015, p. 204).
At the beginning of the Nineteenth century, De Maistre denounces the fanatical risks connected with this optimistic vision of human reason. One century later, Cioran develops De Maistre’s criticism to show that a revolution which dogmatically believes in its own goodness and perfection could represent an ideological justification of political violence. Cioran criticizes that kind of philosophical optimism which judges the revolution as a good in itself, forgetting that every single revolution should be judged in the light of its political effects (or reactions). At the same time, his analysis of De Maistre’s philosophical inconsistencies proves that he is not a reactionary thinker. Cioran knows that this world needs to be improved, and that revolutions could represent the instrument of this improvement. Nonetheless, Cioran is a skeptic and he could not ignore that the history of revolutions is a history of political failures. Thus, he tries to examine both revolutionary and reactionary doctrines in order to better understand the contradictions of any ideological movement claiming the creation/restoration of a perfect world. Cioran is aware that ideologies comply with an anthropological need: the need for certainties. In his essay on reactionary thought, he highlights that we should learn to doubt of our own certainties since there does not exist anything more dangerous than an undoubtable certainty. We could also suggest that Cioran is trying to bring together skepticism and ideology, to see if the revolutionary hope in a better world could be imagined as a skeptical confutation of this world. It is thus not by chance that, in the last page of History and Utopia, he compares the quest for Revolution to the quest for Paradise since both could entail the same melancholic epilogue: “No paradise unless deep within our being, and somehow in the very heart of the self, the self’s self; and even here, in order to find it, we must have inspected every paradise, past and possible, have loved and hated them with all the clumsiness of fanaticism, scrutinized and rejected them with the competence of disappointment itself” (ibid., p. 223).

References


Sartre. J.P. & Kulka, J. (2007). *Existentialism is a humanism (L’Existentialisme est un humanisme); including, a commentary on The stranger (Explication de L’Étranger)*. New Haven: Yale University Press.


WHAT DOES BEING AN “ARISTOTELIAN” REALLY MEAN?

O QUE SIGNIFICA REALMENTE SER UM “ARISTOTÉLICO”?

Sinan Kadir Çelik*
sinancelik@atauni.edu.tr

Abstract. The present paper presents two main arguments: 1) The meanings of terms like “(neo-) Aristotelian” or “Aristotelianism” have become extremely ambiguous in the present literature of ethics and political philosophy. These terms have even become confusing rather than being descriptive or explanatory. The following questions seem to have no proper answers: Who is actually “Aristotelian,” or “neo-Aristotelian,” to what extent and for what reasons? What does “(neo-) Aristotelian” really mean? 2) In order to give some clues to properly answer these questions, as its second argument, the present paper attempts to define the essential methodological characteristics of Aristotelian ethical/political exploration. To be called as an “Aristotelian,” a research should start from the methodological peculiarities of Aristotle’s practical philosophy that make a research “Aristotelian” rather than “Kantian” or “Hegelian.” In the second part of the paper, these peculiarities are defined as methodological prudence and medical dialectics, which are characteristic aspects of Aristotelian way of inquiry regarding ethics and political philosophy.

Keywords: Aristotle, Aristotelian, Aristotelianism, ethics, political philosophy, methodology.

Sumário. Este artigo apresenta dois argumentos principais: 1) os significados de termos como “(neo-) aristotélico” ou “aristotelismo” tornaram-se extremamente ambíguos na actual literatura de ética e filosofia política. Estes termos tornaram-se confusos, perdendo o seu carácter descritivo ou explanatório. Parece que para as questões seguintes não há uma resposta adequada: quem é que, de facto, é um “aristotélico” ou “neo-aristotélico”, em que medida e por que razões? O que significa realmente “(neo-) aristotélico”? 2) o segundo argumento, para fornecer algumas pistas que permitam responder adequadamente a estas questões, tenta definir as características metodológicas essenciais de uma investigação ética/política aristotélica. Para ser designado como “aristotélico”, o investigador deve partir das peculiaridades metodológicas da filosofia prática de Aristóteles que fazem dela uma filosofia prática “aristotélica” e não “kantiana” ou “hegeliana”. Na segunda parte deste artigo estas peculiaridades são definidas como prudência.

* Associate Professor at Ataturk University, Faculty of Literature, Department of Philosophy, History of Philosophy Section Erzurum, Turkey.
Metodológica e dialética médica, que são os aspectos característicos da forma aristotélica de investigação em ética e filosofia política.

Palavras-chave: Aristóteles, aristotélico, aristotelismo, ética, filosofia política, metodologia.

0. Introduction

It would be no exaggeration to say that in the contemporary literature of moral and political philosophy, one’s mind can easily be confused when he comes to realize the abundance of ways available for conducting a research on the current problems of ethics and/or politics. There are so many currents, which could be divided into their rival sub-currents that it seems almost impossible to draw the entire genealogical tree illustrating all of the divisions in a detailed and an indisputable manner. While the literature especially as regards the normative claims about the nature of and the relations between human beings has propagated more and more, intensive efforts have been shown to differentiate all these diverse approaches from each other. These attempts have eventually created their own literature of meta-ethics.

One of the ways often taken in distinguishing scholars and their views regarding ethics and/or political philosophy from each other is to highlight the name of a philosopher who is pursued by them in principle. That is why existing viewpoints in political philosophy are usually denominated as “Aristotelian,” “Spinozist,” “Kantian,” “Humean,” “Hobbesian,” “Hegelian,” etc. Moreover, most of the scholars typically add prefixes like “neo” or “post” in front of these terms and employs notions like “Neo-Aristotelianism,” “Neo-Kantianism,” “Neo-Humanism,” and so on. This leads to a widespread usage of dichotomies in the literature of meta-ethics, such as “Kantian/Aristotelian,” “Kantian/Hegelian,” “Kantian/Humean,” etc. Though their meaning and the validity of their utilization have become a subject of an ongoing debate, all these distinctions have notoriously been employed for categorizing various sorts of inquiries on current problems of ethics and political philosophy. Among them, the most commonly used one is the “Aristotelian/Kantian” distinction, which roughly divides the present literature of moral and political philosophy into two predominant camps. As the prominent figures of the Kantian camp one could give the following short list: John Rawls, Jürgen Habermas, Otfried Höffe, Karl Otto Apel, Robert

1. **Who is an“Aristotelian” and How?**

Although it is still a subject of a relatively wide-spread and ongoing debate that to what extent and in which sense these diverse philosophical views attributed to the above listed scholars can be called as “Kantian” or “Aristotelian,” one might say that the list more or less reflects the general view in the literature. Let us note that terms such as “Aristotelianism” and “Kantianism” have closely been associated with political philosophy rather than metaphysics or epistemology (Benhabib 1988, 38). For instance, Richard H. King holds that there are left and right wings in both Aristotelian and Kantian circles. In his classification, King places Rawls to the left wing and Nozick to the right among Kantians; he sees Richard Berstein closer to left Aristotelians and counts Richard Rorty as one of the liberal Kantians. He also makes use of the terms like “democrats” for Walzer and Sandel (King 1990, 104). In employing left/right dichotomy, Bernard Yack takes a similar path. His notion of “left Kantians” has veritably an entirely different meaning (Yack, 1990). It must be noted that political appraisal of Kant’s works engendered to radical interpretations of his practical philosophy, according to which there are in fact close affinities between Kant and Marxism.¹

Likewise, if we look to the Aristotelian camp, we can see that Aristotle is either regarded as a passionate defender of participatory democracy or as a detrimental enemy to it.² Nussbaum at this point seems right in holding that the works of Aristotle actually gave inspiration for supporting a variety of modern political thoughts like Catholic social democratic views (Jaques Maritain).

---

¹ See (Linden 1988) and (Karatani 2003).
² For an interpretation presenting Aristotle as a defender of democracy, see for instance, (Bates 2003) and (Frank 2005). For an argument typically depicts Aristotle as against the democracy, see (Wood & Wood, 1978).
conservatism (Finnis), Catholic communitarianism (MacIntyre), humanist Marxism (David A. Crocker) and British liberal social-democratic tradition (Ernest Barker)(Nussbaum 2001, xix). In parallel with Nussbaum’s remark, John R. Wallach argues that “Aristotle may, depending on which neo-Aristotelian you read, augment or condemn Rawlsian liberalism, revive or undercut Straussian naturalism, rehabilitate or discredit the tradition of Thomism, support or oppose humanistic Marxism” (Wallach 1992, 615). There is a dispute widely known as the liberalism versus communitarianism debate, in which both Aristotle’s and Kant’s works have been appropriated by various rival political projects and have even played a significant role in developing them.3

What can be inferred from such a complicated picture at first glance is that there are various competitive political projects generated within both the “Aristotelian” and the “Kantian” circles. These political interpretations indicate the existence of sundry competing perspectives, which have been in principle inspired or even derived from Aristotle’s and Kant’s studies on practical philosophy, in order to support or oppose certain political stances. It could however be contended that this trend is actually something closely related with politics in its ordinary sense, not with methodologically adhering to Kant’s or Aristotle’s original works. Understood this way, adopting these political terms as canonical criteria for distinguishing a number of Aristotelian and Kantian currents hardly provides us a proper key to grasp what makes an inquiry an “Aristotelian” one rather than a “Kantian.” Let us note that the meaning of political notions like right/left distinction is in itself contentious. Although these distinctions always refer to a certain political affiliation, they are still vague terms. Secondly, the political philosophies of Aristotle and Kant were originally neither liberal nor Socialist. Being an Aristotelian or a Kantian should not be, first of all, treated as a matter of supporting a certain political view. It seems that attributing our modern political terminology to their political philosophies would rather be perplexing due to its derivation from misused retrospective evaluation. The tendency towards appealing to the categories briefly reviewed above may of course be convenient in sociological mapping of the current tendencies in the literature under consideration, though this would not possibly be explanatory in

3 For general lines of this debate, see (Mulhall & Swift, 1992).
explicating the methodological peculiarities of aforementioned ways of research conducted in the fields of ethics and political philosophy.

Alternatively, grappling with the question of “what does (‘neo-) Aristotelian’ philosophically mean?” one may at the very beginning prefer to embark on an inquiry into the methodological characteristics of Aristotle’s practical philosophy. In the relevant literature, contrasting Aristotle with Kant has nevertheless become a conventional option to deal with this question. On the other hand, there has also been an inclination towards introducing a disjunction for meta-ethically categorizing the salient contentions. Some of these are: teleological ethics versus deontological ethics, eudaimonist versus deontological ethics, practical ethics versus procedural ethics, virtue ethics versus duty ethics and so on. Then, the well-known Aristotelian/Kantian division is not a singular enterprise. Within the frame broadly drawn above, the present Aristotelianism in political philosophy literature is in no doubt generally associated with being teleological, eudaimonist, substantial and practical. In other words, contrary to Kantian ethics, it has been advocated that the Aristotelian point of view privileges the concepts of human good, happiness and virtue over the concepts like right, duty and obligation. In this sense, virtue ethics is often thought to be a kind of ethics which “[p]uts primary emphasis on aretaic or virtue-centered concepts rather than deontic or obligation-centered concepts” (Crisp & Slote, 1997, 3). Defining Aristotelianism as virtue ethics, thus, has turned out to be another consensus, characteristically exemplified in R. Anna Putnam’s words: “[v]irtue ethics is what Aristotle did” (Putnam, 1988, 379). Likewise, when Phillip Mantague identifies two foremost currents in the contemporary literature under consideration as virtue ethics and duty ethics, his perspective represents this common view (Mantague 1992).

At first glance, it seems reasonable to accept the virtue/duty ethics distinction and to interpret Aristotle as the canonical example of virtue ethics. Again, such taxonomical endeavors have for the most part appeared to be more and more confusing rather than being explanatory as the debate progress. Some scholars, notably Nussbaum, have drawn our attention to the point that “virtue ethics” conceived as one of the premier approaches in contemporary debates does not solely belong to Aristotelianism. Likewise, Nussbaum does not hesitate to direct her enthusiastic criticisms against the propensity for opposing virtue ethics
to Kantianism and Utilitarianism. Her critique fundamentally rests on the assumption that both Kant and Utilitarians actually developed a considerable account of virtue. This is the rationale behind her preference to consider Kantian and Utilitarian views as mere exemplars of virtue ethics. In the accompanying literature of ethics, she further enunciates that there seems to be at least two distinct trends, which can be classified under the category of virtue ethics. The first viewpoint, according to her argument, basically undertakes as a duty to reconcile Aristotelian and Kantian perspectives. Scholars who are assumed to belong to the first group, in this sense, “are best understood as motivated by a dissatisfaction with Utilitarianism (...) These ‘virtue theorists’ are likely to turn to Aristotle, or a certain reading of Aristotle, to elaborate their picture of a deliberative political life. They are not hostile to Kant, and they may even desire a synthesis of Aristotle and Kant (Nussbaum 1999, 168). Along with herself, she posits scholars like John McDowell, Iris Murdoch, Nancy Shermen, David Wiggins into this camp. The second group, on the other hand, would be identified with anti-Kantianism. For Nussbaum, anti-Kantians like Annette Baier, Simon Blackburn, Philippa Foot, Alasdair MacIntyre and Bernard Williams “are hostile to universal theorizing in ethics, and they are likely to have some sympathy with cultural relativism, although they do not all endorse it”(Nussbaum 1999, 169).

After her detailed discussion of these trends, Nussbaum’s argument proceeds in the following way: along with Aristotelians, if both Kantians and Utilitarians appeal to a certain understanding of “virtue” in their ethical inquiries, then, there seems no need to employ a conceptual instrument like “virtue ethics” so as to differentiate these schools from each other. In the frame drawn by Nussbaum, that is to say, while “virtue ethics” turns out to be a misleading category, the old notions like “Neo-Humean,” “Neo-Aristotelian,” “anti-Utilitarian,” “anti-Kantian” would thus be sufficient to depict the significant common ground shared by and the dissimilarities between rival theories in ethics (Nussbaum 1999, 201).

Nonetheless, Nussbaum’ account of virtue ethics as a misconception appears to be excessively inclusive as Sean Mcaleer convincingly argues in his recent article “An Aristotelian Account of Virtue Ethics: An Essay in Moral

---

4 See Nussbaum (1999).
What Does Being An “Aristotelian” Really Mean?

Taxonomy” (2007). One of the objections in Mcaleer’s study is plainly directed against Nussbaum’s categorical rejection of virtue ethics. Nussbaum’s proposal introduced in her article entitled “Virtue Ethics: A Misleading Category?” is, for him, absolutely derived from a false postulation that every ethical system that somehow mentions virtue, could plausibly be called as “virtue ethics.” In other words, Nussbaum puts totally irrelevant attitudes into the same basket without explicitly defining what “virtue ethics” is or should be. Kant and Mill, no doubt, declared their ideas about the concept of virtue and even considerably grappled with it in their ethical writings, but this does not simply mean that they can safely be regarded as virtue theorists. In order to settle on the point that whether a theory of ethics can be subsumed under the category of virtue ethics or not, one should first answer the following question: within a theory’s conceptual economy, “are the virtue concepts foundational or derivative, primary or secondary, explicantia or explicanda?” (Mcaleer 2007, 210). Without appealing to this or a similar set of criteria, one would certainly face a great difficulty in determining the crucial distinctive feature which makes a theory virtue ethics. But, for Nussbaum, employing a conception of “virtue” in this or that way seems to be the necessary and sufficient condition for a theory to be labeled as virtue ethics. That is why her understanding of virtue ethics may be seen as too inclusive as effectively argued in detail by Mcaleer.

Further, for Mcaleer, there exists another tendency in the relevant literature, which is engaged with defining virtue ethics in a deficiently inclusive way. For instance, in terms of Gary Watson’s account of virtue ethics, even Aristotle cannot be counted as a virtue theorist. Mcaleer 2007, 210–215). The confusion regarding the meaning of virtue ethics, therefore, brings about an emergent need for further elaborations, rather than a clear meta-ethical taxonomy of the current theories of ethics. For instance, from a significantly disparate angle, Stephen Buckle in his recent article “Aristotle’s Republic or, Why Aristotle’s Ethics is Not Virtue Ethics,” persuasively supports the assertion that Aristotle’s practical philosophy cannot be appraised as a virtue theory in the modern sense. Buckle draws our attention to the point that if the relation between Aristotle’s Ethics and his Politics is carefully inspected, one cannot equate Aristotelian ethics with the modern virtue ethics; for, Aristotle had always insisted on dealing with the question of virtue as part of much more broader
project like Plato’s *Republic* (Buckle 2002). The contemporary virtue ethicists in sharply separating moral and political spheres from each other very much like the Kantians do, actually fail to realize the peculiarities of an Aristotelian way of inquiry. A similar view is suggested by Peter Simpson:

> Viewed in the light of the Politics, Aristotle’s ethical theory is inseparable not only from the opinions of gentlemen, but also from the politics of gentlemen. Virtue exists fully in aristocratic regimes, and elsewhere only in isolation. Since contemporary virtue ethicists have no intention whatsoever of trying their theory to gentlemanly opinions, let alone gentlemanly politics, their theory is not, and could never be, Aristotelian. The ‘neo’ in their title destroys the ‘Aristotelian’ to which it is attached (Simpson 1992, 523-524)

The views examined so far are sufficient to show that the meanings of notions like “Aristotelian,” “Kantian,” “virtue ethics” are in dispute. From one point of view, Kant might be a virtue ethicist, but seen from another, he is not. This is also the case for Aristotle. His practical philosophy could on the one side be interpreted as the well-known exemplar of virtue ethics without hesitation; but on the other side, it is likely to be in contrast to our modern conception of virtue ethics. That is to say, contemporary interpretations of Aristotle and Kant are so interwoven that the notions of “Aristotelian” and “Kantian” could even be regarded as misleading. This seems to be a natural result of the general tendency to undermine the validity of Aristotelian/Kantian distinction by accentuating the putative close affinities between their practical philosophies. Among all others, Otfried Höffe’s argument exhibits one of the most passionate defenses of this position (Höffe 1997). In his article entitled as “Outlook: Aristotle or Kant—Against a Trivial Alternative,” he argues that four types of re-aristotalisation of ethics in the current literature can be distinguished. Opposing Aristotle to Enlightenment thought and liberalism by means of highlighting the elements against universal conception of morality in Aristotle’s political philosophy is, according to Höffe, the first kind of re-aristotalisation, generally known as “communitarianism.” Drawing attention to the importance of Aristotelian “substantial morality,” and at the same time contrasting it to the formality of Kantian ethics is the second version of re-aristotelisation. The third version is called by Höffe as “virtue ethics,” the central concern of which is essentially the development of one’s character excellences and human flourishing. The fourth

---

5 Some of well-known examples of the aforementioned tendency are: (Hursthouse 1999), (Korsgaard 1996), (Sherman, 1997), (O’Neill, 1989) and (Herman 1993).
What Does Being An “Aristotelian” Really Mean?

one is founded upon the assumption that Kantian ethics is deontological and this is the reason why it gives a crucial importance to the concept of “duty.” Aristotle’s ethics, however, would rather concentrate on the concepts like happiness and practice instead of notions like *categorical imperative* and “duty.”

All these versions of re-aristotelisation in the literature of ethics and political philosophy, for Höffe, primarily stem from reductive and misleading interpretations “which are carried out partly on Aristotle, partly on Kant and frequently even on both thinkers, permit rhetorically brilliant polemic and also prophetical exhortations to moral and moral philosophical *metanoia*” (Höffe 1997, 19). Against “re-aristotelisation,” Höffe hence endorses the following claims:

1. In the intention of ethics as a practical philosophy Kant is an Aristotelian.
2. In the basic elements of his ethics Aristotle is a universalist.
3. Where Aristotle allegedly surpasses Kant, namely respecting judgment, he presents an analysis which Kant not only essentially acknowledges, but also continues moral philosophically.
4. With regard to the theory of acting some of the Aristotelian analyses have implications beyond their own theoretical point of departure, i.e. the principle of mere pursuit.
5. With the theory of happiness Aristotle develops a concept, on which Kant’s thesis of the conceptual indefiniteness expresses a skeptical opinion, but which is actually an objective, amazingly broad and also well-defined concept (Höffe 1997, 18).

Höffe’s argument briefly reconstructed above, like many others, no doubt borders upon blurring the defining characteristic features of Aristotelian and Kantian ways of inquiry. Nevertheless, in reading Kant, one could hardly fail to see that his practical philosophy is essentially a strong challenge against Aristotelian political philosophy. This is simply why contrasting Aristotelian ethics and political philosophy to that of Kant seems to me a much more reasonable than laying stress on the common denominator shared by these competing perspectives. But this does not necessarily mean that there cannot in any case be some good reasons to compromise their practical philosophies. For instance, Nussbaum states that her “Aristotelianism has increasingly been influenced by the idea of John Rawls and of Kant” (Nussbaum 2001, xx). Let us note here that there exists an evident disparity between Höffe’s argument and Nussbaum’s statement above. Whereas Höffe endeavors to weaken the validity of Aristotelian/Kantian distinction categorically, Nussbaum, by referring the following facts for clarifying the motives behind her change of mind regarding the relation between Aristotelian and Kantian ways of thinking, states that the
supreme ideal of dignity of humanity is absent in Aristotle’s practical philosophy. Aristotelian approach originally is also blind to the problem of gender, that is, subordination of women to men. Indeed, Aristotle justified slavery without hesitation. Second, Aristotle was always not ethically and politically interested in the people who lived outside his own city-state. Moreover, he did not recognize men as *kosmopolitai* (all human beings in the world). And lastly, in Nussbaum’s words: “Aristotle lacks an essential element of a good modern political approach: a robust conception of protected areas of liberty, of activities with which it is wrong for the state to interfere.” Nussbaum 2001, xxi).

One might agree in principle with Nussbaum. The Aristotelian standpoint must be revised in accordance with the novelties that we have been confronted. In our modern conception, it is plain to the eye that justification of slavery or of inequality between men and women are often regretfully condemned as backward ideas. This is the key lesson that we should learn from the Cynics or the Stoics. However, this does not mean that one is justified in harmonizing practical philosophies of Aristotle and Kant by presuming their methodological proximity. As exemplified in Höffe’s argument, focusing on finding a common ground shared by these two distinct political philosophies could easily result in the ignorance of their radical differences that cannot be reconciled in principle. Further, methodological reconciliation of the two attitudes at first requires to be supported with a detailed meta-ethical assessment. But if one prefers to combine the strong sides of Aristotle and Kant without giving a detailed meta-ethical account, this picture easily turns out to be a patchwork. On the other hand, to be called an “Aristotelian” does not mean repeating and accepting every single word of Aristotle. Rather, if a survey is conducted by following Aristotelian methodology, this must mean that the uniqueness of Aristotle’s methodology is to be pursued. That is to say, the research should start from the very peculiarities that make a research “Aristotelian” rather than “Kantian” or “Hegelian.”

Therefore, it is clear that some scholars call themselves Aristotelians or are denominated by others as Aristotelians. But, what does Aristotelian really mean? In order to answer this question properly, let us try to underline two essential methodological characteristics of Aristotelian exploration.

2. Aristotelian Quest for Prudence in Political Philosophy
What Does Being An “Aristotelian” Really Mean?

In the eighteenth century, inquiries conducted mostly under the umbrella of philosophy faculties, were frequently labeled as “science of man,” “moral science,” “moral physics,” or “sciences morales et politiques.” Mathematics and natural sciences in this epoch constituted crucial reference point for the research regarding politics and social sciences. Concerted efforts were made to unfold natural principles and laws of human nature in place of supernatural agencies’ putative properties. This resulted in the development of elaborate systems of moral duty and political obligation. For example, natural law theories like that of Thomas Hobbes and Samuel Pudendorf were founded upon the postulation that social reality can only be explained in terms of permanent features of human nature, such as the concern for self-preservation. In order to explain how society is formed and functions, what nature of society or politics is, most of the moral and political treatises of that time usually employed mechanical metaphors, organic analogies and an image of the world as a well-ordered machine.

From about 1770 to 1830, especially in France, treatises on moral and political issues were denominated as “social mathematics,” “social mechanics,” “social physics,” and “social physiology.” Within this intellectual climate, philosophical inquires regarding the socio-historical realm in a sense intimated the methodologies of the natural sciences and mathematics. As Bernand le Bovier puts it, “the geometrical spirit is not so attached to geometry that it cannot be taken and applied to other knowledge. A work of morals, politics, and criticism, perhaps even of rhetoric, would be improved, other things being equal, if written by a geometer.”

This is an intellectual current that has been a subject of long-lasting dispute, in respect of which one might pose the following questions: is there a single methodology that encapsulates all facets of reality? Is it really plausible to presuppose that methodology of natural sciences or of mathematics can directly be applied for attaining the purpose of investigating moral and political issues?

On the one side of the debate, one can argue that the application of the methodology that belongs to mathematics and natural sciences into the social sciences or researches dealing with ethical and political issues would at least yield

---

6 Cited in (Gaukroger 2006, 12).
some misleading conclusions. As Alasdair MacIntyre holds, “[t]he eighteenth-century moral philosophers engaged in what was inevitably an unsuccessful project; for they did indeed attempt to find a rational basis for their moral beliefs in a particular understanding of human nature, while inheriting a set of moral injunctions on the one hand and a conception of human nature on the other” (MacIntyre 1984, 53) (italics mine). This is surely a charge that requires a rigorous justification. One still assert that a certain conception of “laws” or “principles” peculiar to an unchanging human nature has more and more been adopted as an ultimate premise in the relevant philosophical texts. This is in effect consciously or unconsciously, implicitly or explicitly, usually coupled with the inquiries regarding ethics and politics. It seems that this trend has turned out to be a prevailing stream in the literature of moral and political philosophy after the 18th century.

One of the motivations behind the variety of the doctrines generally termed “contemporary (neo) Aristotelianism” becomes clear especially when this intellectual current is carefully taken into consideration. As a matter of fact, Aristotle himself would have raised the following questions with regard to aforementioned frantic efforts: do we have a guarantee that the laws of human nature can be defined exactly? Are they essentially similar to that of physics or of mathematics? Is morality or politics just consisting of a series of laws that should be obeyed in any given case? That is to say, following the strategy of “finding out the ultimate laws first and deriving the solutions of specific ethical and/or political problems” would yield no solution to the problem under consideration, and this might even be a vain hope.

Here we encounter a serious epistemological puzzle: is it really possible to identify ourselves and reveal the permanent laws inherent to our nature? If this is possible, is there a guarantee that we can resolve every single problem regarding ethics and politics by directly appealing to these laws? These challenges are quite vexing and always deserve a rigorous philosophical interrogation. But there seems to be a practical way out of this alleged dilemma. This is the Aristotelian solution: in order to investigate the relations between human beings as they are and would possibly be, we are not obliged to take anchor from the putative existence of the unconditional principles similar to mathematical or the
physical laws. It could be losing time to try to derive the ultimate principles regarding ethics and/or politics in a deductive manner either from the laws of human nature like Hobbes or from the “pure practical reason” like Kant. In short, what differentiates Kantian or a Hobbesian inquiry from the Aristotelian is their common presumption that there is a certain human nature, which they derive by observing an “abstract individual.”

2.1. Aristotle’s Challenge: From Abstract Metaphysical to the Concrete Social Individual

To evaluate a human being as an “abstract individual” means that there are some ontological properties, certain characteristics of human beings, which do not change over time and in altered circumstances. This conception, in other words, implicitly or explicitly refers to the putative existence of basic set of intrinsic characteristics of human behavior or of morality which do not change in time and place. It finds its root on the view that human beings are always governed by a set of objective laws. J. S. Mill typically exemplifies this viewpoint when he states that “all phenomena of society are phenomena of human nature, generated by the action outward circumstances upon masses of human beings; and if therefore, the phenomena of human thought, feeling and action, are subject to fixed laws, the phenomena of society cannot but conform to fixed laws” (Mill 1974, 835) (emphasis mine).

One can, however, raise the objection that this or that set of the properties ontologically attributed to a certain conception of individual are not as empirically obvious as claimed. Simple observation would not be sufficient to perceive, for example, man’s natural tendency to engage in a war against every man as Hobbes depicts. If so, it would then be unnecessary even to talk about moral ideals, which would be beyond what is actually going on in social reality. On the other hand, as in Kant’s challenge, it seems possible to hold that investigating only the content of “pure practical reason” enlightens the a priori ground of the human nature, which provides the real morality for all of us. Thus, there would be no need for looking at the empirical world, to our everyday life to know what we ought to do. Nevertheless, both approaches still hinge on an abstract description of what the human being is.
Seen from Aristotle’s perspective, there are no essentially self-seeking, ontologically distinct (or isolated) individuals who come together by social contracts to give birth to a community as Locke or Hobbes argued. It is impossible to observe an autonomous individual exempt from all social relations who could also be endowed with “pure practical reason.” What was discussed up to here is just a brief outline of the major ontological presumption of “methodological individualism.” An Aristotelian way of inquiry radically deviates from the aforementioned apprehensions. In the Aristotelian approach, there is no ontologically autarchic individual. Instead, we always find ourselves as conditioned by and dissolved in the ensemble of social relations. That is the very reason why MacIntyre suggests that an Aristotelian exploration ought to grapple with questions like “of what story or stories do I find myself a part?” instead of the question “who really am I?” In his own words:

> I am someone’s son or daughter, someone else’s cousin or uncle; I am a citizen of this or that city, a member of this or that guilt or possession; I belong to this or that clan, that tribe, this nation. Hence what is good for me has to be the good one who inhabits these roles. As such I inherit from the past of my family, my city, my tribe, my nation, a variety of debts, inheritances, rightful expectations and obligations (MacIntyre 1984, 220).

Thus, Aristotle did not take for granted an abstract conception of individual in its modern sense: “unlike moderns, [Hobbes, Locke and Rousseau] Aristotle ... does not use nature to establish the pre-condition and necessary conditions of politics. He treats human nature, instead, as both a measure of polity and as itself a question for politics” (Frank 2005, 19). The individual as such in Aristotle’s conception is not, then, a given entity having a definite nature. He rather “seeks a richer vision of human experience by taking our own commitments seriously and inquiring into what they imply about the human condition” (Smith 2001, 19). The Aristotelian project, therefore, finds its ground on the following starting point: as we gain experience, and attain knowledge about our moral psychology, our history, biological nature and about politician constitutions, we can have further opportunities to think and act in accordance with the knowledge of politike.

7 For a detailed support of the argument presented above see: (Ball 1978) and (Hinchman 1984).
9 Politike, for Aristotle, is he peri ta anthropeia philosophia, which can be translated as “the philosophy of human affairs.” Its range is much broader than what we understand from the term “political science.” It covers the whole subjects of psychology, anthropology,
2.1.2. *Politike* as Inexact Practical Science.

Aristotle, in the *Nichomachean Ethics*, argues for the supremacy of *politike* over all other theoretical sciences and other practical disciplines (NE 1094a26-1094b6). Aristotle's conception of *politike* as the most authoritative science is the one of the most striking indications of how he thinks practically rather than theoretically in his *Ethics* and *Politics*. When Aristotle hierarchically arranges types of knowledge in accordance with hierarchy of goods (1094a1-1094a26), and when he places *politike* at the apex, he appeals neither to epistemological nor to metaphysical criteria in the modern sense. He orders various kinds of knowledge with respect to practical concerns in a *polis*. This is a kind of taxonomy that is based on one single criterion, which is the value of knowledge in our lives.

Aristotle exposes three substantial reasons in defining *politike* as the most authoritative and cardinal science. First, he argues that *politike* rules and should administrate all other sciences, simply because it prescribes which sciences should be learned in a *polis*, by whom and to what extent (1094a29-1094b1). Mathematics, for instance, is the most exact theoretical science, though mathematics cannot decide whether it is better for every citizen to learn geometry or not. Medicine is definitely very beneficial and useful for all, but this does not mean that every citizen should be trained as a doctor. It is also necessary to have recourse to *politike* when one seeks answer to questions like “how many professional mathematicians, engineers, soldiers and doctors should be accommodated in a *polis*?” It is *politike* that should adjudicate to what extent and why sciences like mathematics and medicine are useful for us, not *vice versa*. There is no further need to emphasize the vital importance of medicine. Mathematics might supply the most exact and structural knowledge that we could ever reach, however, this does not mean that the mathematical method ought to be pursued in order to assess its own usefulness for achieving the goal of living happily in a *polis*. Mathematics cannot mathematically prove its own usefulness. Exactitude belongs to the sphere of theoretical sciences, which aspire to ascertain economics, history, sociology, ethics, law and political science in our modern conception. All these are separate disciplines for us, subject matters of which are remarkably different from each other. But for Aristotle *politike* is the only discipline, an artful and practical science that encompasses all questions related with these modern academic disciplines. In the texts of Aristotle, in particular cases, it might also mean “practical philosophy” or “art of politics.” There is a danger that its meaning in Aristotle’s nomenclature can be confused with our contemporary understanding of the terms like “political science.” That is why I prefer not to use English “political science.” Cf. (Wallach 1992, 616-618).
the unconditional truth for its own sake. But, the significance of truth, the ways of its appropriation and its essential function in our life are all the subject matters of politike. If the life in a polis is fundamentally organized through politics and if we are “political animals,” then politike must be the master of all practical concerns, unlike mathematics, logic or physics.

Secondly, seen from Aristotelian point of view, reversing the hierarchical order of politike and other theoretical or practical disciplines means that knowledge partial in its scope is blindly put over a much more comprehensive one. That is why all practical disciplines like medicine or the art of commandership ought to be subordinated to politike (1094b2-3). In other words, the scope of politike is already much broader than all other sciences and disciplines. Reversing their positions in the frame that Aristotle has drawn would be both against Aristotle’s teleological orderliness and would also bring about putting an impediment to accomplish the most honorable purpose of life, namely eudamonía (1097a25-1097b27). For example, if the subordination of politike to a practical discipline such as generalship is accepted by people, then it can immediately lead us to justify the existence of an unjust social system, e.g. dictatorship in which most of the citizens do not have a chance to be happy as could be inferred from the Politics. The purpose of politike is rather to embrace and to govern all practical disciplines and theoretical sciences in accordance with Aristotle’s teleology.

Third, politike makes use of all types of knowledge obtained from all of the other sciences concerned with action and puts them into practice by laying down rules and laws. This again shows us what is right and what is wrong to do (1094b4-5). In today’s terminology, the task of politike is to collect the relevant data from psychology, anthropology, sociology, history, law, biology, etc. in order to adjust the required regulations properly and execute them within a society. Thus, politike is the discipline that appeals to all sorts of knowledge for specific purposes.
Aristotle, on these grounds, affirms that *politike* is and ought to be superior to all other sciences (1094b6). *Politike*, in short, aims to achieve the highest good of human being by employing all sciences. Let us note that when Aristotle speaks of the mastery of the *politike* over all other sciences, he does not appeal to the internal characteristics of knowledge, such as consistency or certainty. Similar to the practices in a society, for Aristotle, sciences are and must be ordered in accordance with the existing hierarchy of goods, and here *politike* occupies the highest place. This is also fulfillment of the human function (*ergon*) (1097b25), being happy and doing good, not to define what the virtue, happiness or the good in itself is (1097b11, 1095a4-6, 1103b26-30, 1179b1-4). As Salkever says:

> [a] misleading approach to Aristotle is to read him as if he were attempting to establish some sort of Archimedean point, an absolute perspective from which the accuracy of ethical and political choices could be guaranteed by reference to some rule or system of rules. This kind of reading gets its plausibility from the way it reflects modern expectations about philosophy. We assume Aristotle is a ‘philosopher,’ and we expect philosophers to be in the business of replacing practical doubt with theoretical certainty - an expectation that seems adequately met by, say, Hobbes, Kant and J.S. Mill (Salkever 1990, 4).

We should here note again that when Aristotle sorts various kinds of sciences and practical disciplines, he does not try to extend his method developed in the *Metaphysics* into *politike*. Theoretical sciences have their own aim and principles:

> But in each science the principles which are peculiar are most numerous. Consequently it is the business of experience to give principles which belong to each subject. I mean for example that astronomical experience supplies the principles of astronomical science: for once the phenomena were adequately apprehended; the demonstrations of astronomy were discovered (Post. An. 46a17-19). 10

Yet, practical sciences have practical purposes. It is worth underlining the point that Aristotle’s distinction between theory and practice does not resemble our modern understanding. In Aristotle’s terminology, the distinction between theory and practice “is not a distinction between the general and the particular but between knowing what is right and wrong on the one hand, and actually doing the right and avoiding the wrong on the other. The modern distinction passes over Aristotle’s distinction and ignores his practical concern.” (Simpson 1992, 505). In one sense, practical sciences are essentially action-oriented and depend on the improvement of experience. Principles by themselves have no meaning in

---

10 Translation of “Prior Analytic” and other works of Aristotle not related with his practical philosophy are from (Aristotle 1984).
**politike.** The real difficulty is always to find out how to substantiate them. Arriving at the right judgment without executing is meaningless. The definitive purpose of politike, for this reason, is the act itself, not contemplation. This contention rests on the assumption that knowing the general principles about right acts is something different from committing the act in a particular context. In Aristotle’s view, abstracted from the agent who commits an action in a particular context, an act itself has no meaning. For Aristotle, the essential problem is not to ascertain a procedure that secures arriving at the right moral judgment; rather it concerns constitution of the character in a certain manner and the context within which this character behaves. Aristotle does not see moral principles as universal recipes that can be applied in every situation; rather he prefers to seek a way to be, for instance, trying to be a generous person who acts generously in a spontaneous manner instead of performing the right act by obeying some higher principles. This confirms that politike is essentially inexact and necessitates prudence (phronesis). As Simpson states, Aristotle’s theory of virtue or his Ethics and Politics in general does not seem to be a moral theory at all, at least in the sense of moral theory that we standardly recognize. What we want from a moral theory is some over-all account of moral goodness and badness which we can then use to show why this or that particular act is right or wrong. So the Kantian explains the right and wrong in terms of agreement with the categorical imperative, and the utilitarian in terms of promotion of the general welfare. Aristotle indeed has a general account of virtue, that it is a mean between two extremes, and so on. This general account, however, cannot be used to show that something is an act of virtue or something else an act of vice. The truth about such particulars is not shown by theory; it is perceived by prudence (Simpson 1992, 512).

Hence, from the Aristotelian perspective, trying to approximate the level of certainty in mathematical knowledge in politike would be imprudence. Then, the first and the last task we should undertake in Aristotle’s doctrine are to become skilled at how to think and act prudently.

**2.1.3. Methodological Prudence as a Remedy for the Inevitable Inexactness of Politike**

According to Aristotle, one should first of all be aware of the fact that it is impossible to reach the same level of exactness in politike as the one attained by mathematical proofs. This is one of the salient motives behind Aristotle’s criticisms when concerning the incompatibility of Plato’s theory of forms with politike. Since there is no single universal idea of the good for all sciences...
(theoretical and practical), the search for it is futile (NE 109615-35). Further, even if one could reach a universal vision that encompasses all practical sciences, this would be irrelevant to action and useless for the purposes of politike (1096b30-1097a15).

Thus, seen from the Aristotelian point of view, Plato goes to an extreme when grounding his argument for the ideal state on axioms about the good similar to those of geometry. Plato was wrong, according to Aristotle, in following a mathematical method for the purpose of providing permanent solutions to the political problems of a polis, and in assuming the preexistence of some universal laws pertaining to all kinds of social practices. The Aristotelian way of inquiry does not ground ethics on such unshakable foundations. Its goal is not to characterize what morality is or is to find out indubitable facts about human nature.

Secondly, it should not be considered a mere coincidence that Nichomachean Ethics begins and ends with the theme of happiness (eudaimonia). For Aristotle, activities in political life, in general, may properly be defined as the pursuit of happiness, which is the highest good in human life. Since a life with happiness is that which is most complete, self-sufficient and choiceworthy (hairetos); happiness is and ought to be the primary life-goal governing all of our other (secondary) goals in life. Hence Aristotle seems to think that the task assigned to politike is to strive at attaining a kind of knowledge, which is by definition, or necessarily useful for happiness. Furthermore, Aristotle argues that genuine ethical arguments should “be useful for the conduct of life,” “harmonize with facts of human life ... and so encourage those who comprehend them to live by them” (1172b5-8).¹¹

However, it would then be very difficult for us to uncover one single way for being happy. The search for happiness is a long and a resolute journey; and we should first be trained to be very careful in our steps. At the opening of NE, Aristotle emphasizes the importance of prudence in the search for truth in politike:

Our discussion will be adequate if its degree of clarity fits the subject-matter; for we should not seek the same degree of exactness in all sorts of arguments alike, any

¹¹ All quations from *Nicomachean Ethics* is from the following translation: (Aristotle 1985).
more than in the products of different crafts. Moreover, what is fine and what is just, the topics of inquiry in political science, differ and vary so much that they seem to rest on convention only, not on nature ... Since these are the sorts of things we argue from and about, it will be satisfactory if we can indicate the truth roughly and in outline; since [that is to say] we argue from and about what holds good usually [but not universally], it will be satisfactory if we can draw conclusions of the same sort (1094b3-23).

This is evidently not an invitation to the pursuit of certainty. Aristotle here suggests that his objective is to discover “what holds good usually.” A student of politike, thus, should not expect to attain exact knowledge at the end of investigation. This is not his only remark; throughout NE, Aristotle continually to insists on the essential inexactness of politike:

Each of our claims, then, ought to be accepted in the same way [as claiming to hold good usually], since the educated person seeks exactness in each area to the extent that the nature of subject allows; for apparently it is just as mistaken to demand demonstrations from a rhetorician as to accept [merely] persuasive argument from a mathematician, each person judges well what he knows, and is a good judge about that; the judge in a particular area is the person educated in that area, and the unconditionally good judge is the person educated in every area (1094b23-1095a1).

What is indispensible for a student of politike is to be aware of the peculiar nature of the subject matter being investigated. While theoretical sciences are concerned with the nature of unchanging beings, politike deals with the sources and the products of human activity. Then, the nature of questioning life itself or illuminating questions like “how can we be happy?” should not be confused with the aim and subject-matter of scientific investigations:

What science is evident from the following, if we must speak exactly and not be guided by [mere] similarities. For we all suppose that what we know scientifically does not even admit of being otherwise; and whenever what admits of being otherwise escapes observation, we do not notice whether it is or it is not, [and hence we do not know about it]. Hence what is known scientifically is by necessity. Hence it is eternal; for the things that are by unconditional necessity are all eternal, and eternal things are ingenerable and indestructible (1139b19-24).

In contrast to theoretical sciences, then, politike is concerned with what is up to us, that is, what can be otherwise, not with necessities and essentially unchangeable. It is de facto associated with the deliberation which “[c]oncerns what is usually [one way rather than another], where the outcome is unclear and the right way to act is undefined. And we enlist partners in deliberation on large issues when we distrust our own ability to discern [the right answer]” (1112a8-11). In this sense, the subject matter of politike is always open to debate and further deliberation, for it is concerned with particular and contingent cases, namely
human actions and interactions, which belong to the field of what is up to us, not necessity.

Studying *politike* is therefore not merely a matter of attaining the most accurate knowledge that we can ever reach. The relations between numbers or physical objects cannot be treated in the same way with the relations between human beings:

[W]e must also remember our previous remarks, so that we do not look for the same degree of exactness in all areas, but the degree that fits the subject-matter in each area and is proper to investigation. For the carpenter’s and the geometer’s inquiries about the right angle are different also; the carpenter is confined to the right angle’s usefulness for his work, whereas the geometer’s concern what, or what sort of thing, the right angle is, since he studies the truth. We must do the same, then, in other areas too, [seeking the proper degree of exactness], so that digressions do not overwhelm our main task (*NE* 1098a27-35, see also 1103b34-1104a9).

What can reliably be inferred from these passages is that applying the method of mathematical investigation in *politike* (like Plato) would most probably yield results incompatible with its ultimate aim. An investigation regarding life has also an essential practical side and is in fact genuinely practical. On that account, all pursuits to attain accurate knowledge about what happiness is without ultimately striving at being happy seem to Aristotle to be nonsensical. For Aristotle, assuming the existence of a higher separate reality (like the world of Ideas), which governs or should control our actions, is not only a theoretically implausible assumption, but also practically futile. If one investigates life itself as it is and as it ought to be, then the subject matter of this study must both be the improvement of a single character, its acts, behaviors, and interactions among people living a *polis*, all of which are in the terrain of what is up to us, things that can be otherwise. Instead of these actions and interactions themselves, it would seem implausible for Aristotle to begin the investigation by suggesting the putative existence of some universal laws underlying every action of human beings. The *methodos*, the way of inquiry that ought to be followed in these incommensurable realms like mathematics and politics, should conform to their peculiar natures and characteristic aims. Sources of human actions and their results, the conventions and beliefs about them cannot be analyzed as if they were physical objects in front of us. From the very beginning, as students of *politike*, we find ourselves within a wide span of contradictory views about what is better for us, how we can achieve happiness and act in a right way in a particular context.
It is not a coincidence that the starting point of Aristotle’s inquiry concerns conventional beliefs on happiness. The first and the last lesson that should be derived from Aristotle is to learn not to stick to an idea without questioning it. Aristotle warned against thinking imprudently, and advised prudence in the research concerning *politike*. One of the most striking peculiarities of Aristotle’s methodology regarding *politike*, one essential aspect of his argument for the elimination of false beliefs and unhealthy desires from our souls, is the medical dialectic.

### 2.2. *Politike* as Medical Dialectic

To repeat, the ultimate purpose of *NE* is at least in one sense to present the general knowledge of what virtue and vice are in the realm of what is “up to us.” Aristotle does not dictate us some general rules like “you should do X in the context of Y.” To do so would contradict the nature of *politike*. It is very difficult to recognize, in particular contexts, the best way to act. For instance, people might get angry at their close friends when they think that they are treated by them unjustly. Anger in certain situations can be a mere expression of excessive love, as exemplified in the sayings “cruel is the strife of brethren” and “they who love in excess also hate in excess” (*Pol.* 1328a10-17). An action, then, might be provoked by anger, but the problem is not to assess what the anger is as such. A Kantian is prone to reject anger categorically for the reason that a right action should not originate from an emotion. From this perspective, motivation behind an ethical judgment ought to be a rational decision making procedure. In contrast, Aristotle does not assign a particular meaning to anger. He neither categorically confirms nor rejects it as such. Rather, he analyzes the context within which anger emerges. Anger is a feeling that might also stem from incontinence. In this case, anger seems to be associated with the immaturity of a character. But in another case, anger could emerge as a response to an apparent form of injustice. Moreover, becoming angry in particular instances can be inevitable and even an expression of justice (*NE* 1135b25-30). In some other situations, to settle on “the way we should be angry, with whom, about what, for how long” would not be easy (1109b 14-18).

---

12 All citations regarding Aristotle’s *Politics* are taken from the following translations: (Aristotle 1995), (Aristotle 1997), (Aristotle 1999) and (Aristotle 2005).
The problem is, thus, to find out the right means, the most proper way to behave in certain circumstances. This is, for Aristotle, a puzzle that cannot be solved in a purely theoretical way. Aristotle deals with this problem by assigning himself the task of finding out the right means for the virtue of prudence. This virtue can be defined as a kind of eye that “operates in the here and now, for it decides what is the virtuous thing to do here and now; and judging the here and now is the work of perception ... In other words, prudence does not reason about virtue, it directly intuits it” (Simpson 1992, 510). This is why only students who have an inclination to be prudent can learn politike; it is by means that he improves his capability of being prudent and acting prudently. Openness to develop one’s own character in the manner of prudence is thus both a precondition and an ultimate aim for the student who wants to be trained in politike, and only he can benefit from this knowledge.13 This is associated with experience. But, learning to reason in mathematics, and learning how to respond to the peculiarities of the contexts in which we act are remarkably different from each other. Cleverness would be sufficient for learning theoretical sciences, in which young people can easily be successful:

[To understand the difficulty and importance of experience] we might consider why a boy can become accomplished in mathematics, but not in wisdom or natural science. Surely it is because mathematical objects are reached through abstraction, whereas the origins in these other cases are reached from experience. Young people, then, [lacking experience], have no real conviction in these other sciences, but only say the words, whereas the nature of mathematical objects is clear to them (1142a 16-21).

Here, for Aristotle, we should note that being experienced or mature does not solely depend on the age, but mainly refers to maturity of character (1095a6-7). Nonetheless, young people, who are generally inclined to give up anything that is not pleasant easily (Pol. 1340b15-16) are prone to be driven by their sensuous pursuits, not benefiting from politike (NE 1095a8-11).

Aristotle repeatedly emphasizes how important it is to eliminate false beliefs, prejudicial inclinations and dispositions from the constitution of a

13 For a detailed analysis of Aristotle conception of virtue, see (Bodéüs 1993, especially 27-38).
The variety of ways of life, what people generally do and why, are the chief problems that wait be resolved by finding out what could be better for us to do and to be. At this stage, one must not start with principles taken for granted in political inquiry. A mathematician can base his inquiry on self-evident axioms. A student of politike, in contrast, should always be ready to face puzzles and obscurities throughout his investigation. There is no a ready solution to a problem regarding what is better for us. Therefore, the outstanding task of a student of politike is (i) first, “to set out the appearances” [i.e., setting out what appears to people as right]; (ii) second, “to go through the puzzles [i.e., exposing the appearances to the test of reason for consistency and compatibility with the circumstances and requirements of human life]; (iii) to leave behind prejudices and take the rationally tested beliefs as basis for further investigation in political inquiry (1145b3-9).

That is to say, Aristotle’s method of inquiry aims discovering the origins (archai) of human affairs by observing the very practice (praxis) embedded in a certain web of social relations and by critically evaluating the common-beliefs (endoxa) adopted by the majority of people. The paramount task of the Aristotelian inquiry is to assess significant philosophical views so as to teach live in a certain way than another. This inquiry, to repeat, should make use of the dialectical method by which one would perhaps arrive at general principles. At the end of the research, we can have an opportunity to formulate these principles that are correct usually, but not unconditionally. Furthermore, what is more important than formulating these principles is to be trained in how to be an excellent person:

who judges each sort of thing correctly and each case what is true appears to him. For each state [of character] has its own special [view of] what is fine and pleasant, and presumably the excellent person is far superior because he sees what is true in each case, being a sort of standard and measure of what is fine and pleasant” (1113a30-35).

This is a very important point that requires to be emphasized: modern readers who become accustomed to a Hobbesian or a Kantian methodology in searching for unconditional laws of morality or of politics should not expect to

14 For a detailed exposition of this argument, (Allard-Nelson, 2004).
15 Cf. NE 1176a16-20.
give a definite demonstration \((apodeixis)\) of the Aristotelian dialectical inquiry.\(^{16}\) It is not a mere analogy that while Aristotle contrasts ethical inquiry with mathematics, he often resembles it to medicine. Along with scholars who conceive this “medical analogy”\(^{17}\) as a key to comprehend Aristotle’s methodology, I think that this analogy really explicates the spirit of the Aristotelian way of inquiry in a lucid manner. If \(politike\) is a practical and an inexact science and if its primary goal is improvement of character in the manner of prudence, the medical analogy makes it clear that Aristotle’s conception of \(politike\) is both context-sensitive and non-relativistic.

Let us note that there is no requirement for a mathematician to encounter a concrete case. For a physician, however, facing a tangible case, such as a disease, is inevitable. The task of the physician is both to grasp the general principles of medical science and to improve his medical experience. In this sense, a doctor’s responsibility is not only to acquire theoretical knowledge of medicine, but also to be involved in the improvement of practicing it. Unlike mathematics, both medical and ethical knowledge have a practical goal, which is curing diseases. While medical treatment is the remedy for the illness of the body, \(logos\) (understood as speech and/or argument) is the treatment for illnesses of the soul. The function of \(logos\) is, in other words, to cure the soul, namely the diseases stemming from our false thoughts and unhealthy desires. So as to become happy, one should enhance his practical wisdom by eliminating his false beliefs and by constantly curing diseases stemming from his unhealthy desires. This process resembles how a physician gains his occupational experience. In short, both \(politike\) and medical science have theoretical (knowledge of general principles) and practical (experience of complex particular unique cases, having a practical goal, responsiveness to context, gaining experience, comparing several competing views to resolve a particular problem) dimensions.

Suppose that as a doctor, I prescribe a wrong medicine to a patient. There may be two reasons behind such an error: (i) the current knowledge of medical science might not have developed yet to cure this disease. The best way I can

\(^{16}\) Cf. \textit{Top.} 100a27-30. For recent and lucid treatments about the dialectical nature of Aristotelian political inquiry, see also (Smith, 2000) and (Jacobs 2004, especially 26-34).

\(^{17}\) For detailed analysis of this analogy, see (Jager, 1957) and (Lloyd, 1968).
choose in this case is to trust my intuitions and experience. I may naturally do something wrong, and as in the case under consideration, there could be no other alternatives. (ii) I may not be experienced and/or educated enough to decide which medicine would be useful for curing this particular disease. Similarly, as someone who is confronted with a complicated situation, which requires giving an ethical judgment, if I deviate from the way of virtue and misjudge, this may stem from similar reasons: (i) current theoretical knowledge of ethics may not be sophisticated enough to cope with such a complexity. (2) I may not be experienced and/or educated enough to arrive at a proper ethical judgment when I am faced with such circumstances. Thus, very much like the medical science, improvement of ethical theory and experience are open-ended processes. By cultivating ethical theory alongside with gaining experience we can train ourselves in such a way as to immediately find the most virtuous way, even when we are encountered with complex situations.

The following seem to follow from the medical analogy: first of all, just like the medical science defines what health is and searches for the means for being healthy, ethical science or politike aims at attaining a clarified and articulated view of proper ethical life. Secondly, Nussbaum seems to be right when she argues that Aristotle was the first philosopher who has shown an explicit awareness to the fact that “while medical treatment is a form of bia, of external causal intervention, argument [on which ethical treatment based] is something apparently gentler, more self-governed, more mutual” (Nussbaum 1994, 69). Since the success of an ethical treatment depends on the adoption of the ethical ideal by the pupil, the “intellectual asymmetry” and the “authoritative and one-sided quality of logoi” on which medical treatment is based cannot be extended to ethical inquiry. Ethics is and should be “less one-sided, more ‘democratic’ than medicine: the benefits of its logoi require each person’s active intellectual engagement [on the sides of both the teacher and the pupil].” Nussbaum 1994, 70). Thirdly, it is one of the basic tenets of Aristotelian ethics that ethical treatment should conform to ethical and political experiences (or “appearances”). If the diversification of these experiences across different places and times is taken into account, the Aristotelian ethical treatment should consider relativity of values in the realm of human affairs. Yet, this does not mean that the Aristotelian view rests on a relativistic ground; rather, it has a non-relativistic
conception of good human functioning. In line with this, general ethical principles must not be considered as clear-cut recipes directly applicable to all actual situations, but as guidelines for recognizing what is actually there in order to determine what is the right thing to do in any particular context.

3. Concluding Remarks

Let us highlight Aristotle’s genuine conception of ethics and politics by exposing some remarks concerning how one should read the Politics. In interpreting Politics, one should always keep in mind that Aristotle always derives a set of generalizations from the previous studies possibly made by his pupils, at least under his supervision. These studies were on the histories, constitutions and other main characteristics of 158 city cities or political organizations. This could be a reason why Aristotle gives abundantly many examples from histories of Greek city states to demonstrate a specific conclusion. Principles in Aristotle’s practical philosophy can be conceived as the generalizations that had already or could in future have certain exceptions. In Aristotle’s logics, the statement that “this proposition is false” amounts to saying that this proposition is unconditionally false. However, in Politics asserting that “this is false” means that “but perhaps this is not wholly correct” (Pol. 1282a14). This might be the reason why Aristotle frequently employs the word “might be” or “perhaps” in Politics (Robinson 2005, ix). Aristotle’s inquiry in the Politics, in this sense, is consistent with context-sensitivity and non-relativism discussed in the previous section of the paper. Since the subject matter of Politics has not an unalterable nature like mathematics, one should always be cautious about the scope and degree of certainty of the statements regarding politike.

References


COMPROMISE AND/OR INTEGRITY? SOME NORMATIVE REMARKS

COMPROMISSO E / OU INTEGRIDADE? ALGUMAS OBSERVAÇÕES NORMATIVAS

Beatrice Magni*
beatrice.magni@unimi.it

Abstract. In a compromise parties make a strategic calculation that a sacrifice of some their commitments is the best tactic to advance their other values and interests over their current positions. But if we try to figure a compromise as a way to get to agreements that can minimize harm and promote mutual respect between individuals, and could increase the level of cooperation between them, we can consider a compromise shaped in that way as an agreement that can translate at its best some of the individual’s deliberative ideals, as fairness and mutual respect. This kind of compromise will be different from simple negotiation, and more demanding to each one of the parties. It will require first and foremost a reciprocal recognition of moral fairness. But can we ever compromise on matters of ethical principle without compromising our integrity?

Keywords: Compromise, Disagreement, Integrity, Pragmatism, Reconciliation, Hypocrisy.

Sumário. Num acordo, as partes fazem um cálculo estratégico de que o sacrifício de alguns dos seus compromissos é a melhor tática para avançar os seus outros valores e interesses sobre as suas posições atuais. Mas se tentarmos chegar a um compromisso como forma de chegar a acordos que possam minimizar danos e promover o respeito mútuo entre indivíduos, e aumentar o nível de cooperação entre eles, podemos considerar um compromisso moldado dessa forma como um acordo que pode traduzir no seu melhor alguns dos ideais deliberativos do indivíduo, como justiça e respeito mútuo. Esse tipo de compromisso será diferente da simples negociação e mais exigente para cada uma das partes. Exigirá, antes de tudo, um reconhecimento recíproco da equidade moral. Mas será que podemos fazer compromissos envolvendo princípios éticos sem comprometer a nossa integridade?

Palavras-chave: Compromisso, Desacordo, Integridade, Pragmatismo, Reconciliação, Hipocrisia.

* Università degli Studi di Milano.
Introduction

The challenge of dealing with compromise touches on a range of deeply contested yet essential concepts in contemporary political philosophy (Gallie, 1956), such as the nature of politics and ethics. The ambivalent term “compromise” is increasingly used in two quite different senses (Cejudo, 2010). We can use ‘compromise’ - in a first and more positive sense - to address a kind of agreement between different points of view, reached through a balance or mediation (e.g. a negotiation made by mutual concessions among the parties). Proposing, and accepting, compromises means in this first case to accept the presence of a trade-off between personal desires and commitments and various possibilities2 to agree, to disagree, and to agree on our disagreement.

Otherwise we can ‘compromise’ ourselves – in a second and more negative sense – in particular situations where we have to make concessions to something considered harmful or wrong, in other words to something damaging or prejudicial3 to our very deep beliefs. In this second sense the idea of some ‘moral compromise’ could seem an oxymoron: if a moral compromise occurs every time you deal with choices or actions that can put at risk values and ethical principles with whom you strongly identify, how could you accept a compromise and keep yourself morally undiminished? The paper aims to propose an answer to this question, firstly considering the normative evolution of the idea of integrity, and, secondly, investigating what public role integrity and hypocrisy can and should

---

1 This is one of the main argumentations of Fumurescu (2013): compromise is today understood in a positive or “commendable” sense, and in a negative or “condemnable” sense, and the origin of this radical split can be found in the way – as Fumurescu notes – this concept evolved in the two very different contexts of Britain and France between the late Renaissance and Enlightenment periods. “The positive sense refers to the view of compromise as a political virtue that enables two distinct entities to resolve disagreements without resorting to force and violence, and it was in this sense that the concept emerged in Britain. The negative sense refers to the view of compromise as the violation of the essential integrity of one’s inner self or self-conception, and it was in this sense that the concept evolved in France” (review by R.P. Hanley).


3 See Cejudo (2010), p. 301: “However, the literature on ethics has not displayed a similar interest on the issue, in spite of the recent work of Bellamy (2002), Nachi (2004), and Margalit (2010). Possibly this is because compromising and bargaining are not valued as part of honest behavior if we are concerned with values or duties. A compromiser should refuse to compromise when her moral values are at stake, or at least consider refusing... in that case a compromise will be wrong when the concession goes too far because we risk our dignity, do not respect our own principles, or do not comply with our duties”. See also Day J.P. (1989), “Compromise”, Philosophy, 64 (250), 471-485.
Compromise and/or Integrity? Some Normative Remarks

precisely play in politics, and how individuals can preserve the former and to incorporate the latter in the constitution of a public framework.

But now and then we could face some individual or institutional situations where a sort of compromise becomes necessary. The necessity to consider compromise is motivated by the existence of a number of reasons that contribute to increase disagreement within the institutions, and in society in general: individual interests, the scarcity of resources, the lack of adequate information, contingency, uncertainty and moral complexity are altogether elements of disagreement. A compromise may be a way to examine possible solutions to some of these difficulties.

In the first part of the paper I will present the proposal of a conception of compromise that – if we assume the fact of pluralism and disagreement – aims to be more than a mere strategic or ad hoc concession among two individuals or groups (or between a public authority and the citizens). Then I will argue that such a model of compromise could represent a valid method to face value conflicts in politics. We should use the term “pragmatic”, that in my argument doesn’t mean only the result of a virtuous balance between different kind of interests, but it will include a deeper commitment in finding the conditions to offer at least some normative accounts of how institutions ought to respond to value conflicts in politics. In the second part I will bring together the usually separate discussions on the ethics of authenticity/integrity and ethics of artificiality/hypocrisy. I will investigate – in a very brief analysis - what public role integrity and hypocrisy can and should precisely play in politics, and how the idea of pragmatic compromise could enable individuals to preserve the former and to incorporate the latter.

An overall conclusion is that – once integrity and hypocrisy are redefined as consistent concepts in politics – pragmatic compromise is positively related to the reliability of an individual and the accountability of an institution. Consequently, compromise and integrity can not only stand side by side, but also they can and should work together in the constitution of a public framework.

The problem of disagreement

Disagreement represents a peculiar feature of the political life of contemporary liberal and democratic societies. Citizens of liberal societies
disagree not only about which policies should be implemented and how, but also about matters of principle, namely on what are the best reasons to justify the adoption of one policy or the other.

Furthermore, in liberal societies there is an unstated consent on the opportunity that non-public believes should not be thoroughly expressed in a public field: everybody accepts limits - when not prohibitions - to the public expression of one’s personal convictions. But if respect (respect of personal believes) is an unavoidable bond for the stability of a democratic system, immediately we have to face a number of problems and difficulties, both practical and institutional, as follows:

- How can we find deliberative processes or procedures able to give voice to every citizen?
- How can we articulate a sort of neutral general public will, in case of disagreement?
- When we found this deliberative procedure, how could we then make it responsive to the content of individual moral conceptions?
- How can we, finally, tame the dissent that may nevertheless linger on in liberal societies?

Bioethical controversies have become in the last few decades crucial issues in public discussion of western societies (Benjamin, 1990). They are perhaps the best example of a persistent existence of a high level of public disagreement. This kind of controversies is even more interesting as they show that many different fundamental accounts of our moral lives may arise. Political disputes, public disagreements – and their connected controversies – deal not only with which kind of politics could give the best and most desirable outcomes, but also which kind of outcomes can be defined as desirable for themselves. This is one of the most typical cases where issues of desirability and feasibility cross and clash. In other words, and to give just an example, the disagreement on the beginning of life relates to the moral understanding and vision considered when we have to decide which kind of good practice adopt in each case (namely and in a nut shell, to sustain pro-life and/or pro-choice positions). In this case we can have
disagreements that can be, therefore, categorized in - at least - two different levels:

- disagreements on political solutions for solving problems: first level of disagreement;
- disagreements on desirability of the reasons of these political solutions: second level of disagreement.

If we consider for instance disagreements associated with medically assisted procreation, those that I called first-level disagreements will deal with discussion on the legitimacy of turning to such techniques (for example, some consider m.a.p. itself an immoral act as it manipulates a natural event); while second-level disagreements will deal with discussion on the justifiable validity of reproduction using such techniques involving donors and surrogate mothers. The discussion on embryos’ treatment (in particular, those extra embryos that have not been implanted in the m.a.p. process) can be placed at even another level. It seems evident that, even at this general stage of the argumentation, the concept of disagreement needs the involvement of some theoretical approaches, to answer practical questions. The two main key issues here seem to be:

- What can we do when shared principles have divergent implications?
- What can we do when there aren’t any shared principles?

If it is true that one of the central tasks in contemporary political philosophy is to identify principles governing political life (Moreno, 1995) - where citizens disagree deeply on important questions of value and, more generally, on the proper ends of life -, it is also plausible that rationally irresolvable disagreements (as is the case with bioethics) can become the task before us.

When we deal with disagreement and the possibility of compromise, an additional meaningful critical duality at which we have to look carefully is that between moral disagreement and moral dilemma. Moral disagreement is a conflict of moral opinions among different subjects. There can be apparent disagreements, that can be solved with moral considerations, factual or semantic, and there can be genuine disagreements, that cannot be solved on the basis of moral, factual, semantic considerations. Moral dilemma, on the other hand,
affects a conflict internal to the individual, concerning directly the agent. It is moral disagreement that will be the main focus in my approach on compromise

Therefore, given pluralism as a fact, given the irreducibility of some, if not all moral disagreements, I would like to find out, in this work, among the many and different meanings available in political theory, a meaning of compromise that could be something more than a mutual concession open to use for instrumental ends, and eventually accepted by two conflicting parties. Namely I’ll try to propose an idea of compromise that allows the coexistence of searching compromises in public life - and in some political fields - and the loyalty to our deepest principles. To do that we need to proceed orderly, and therefore to start defining precisely what it means to me firstly compromise, and secondly integrity.

1. The necessity of compromise

The tension between individual values and public reasons cannot only be found in the bioethical field, but it is constant and recurrent in political life. This is the lesson of Sophocles’ Antigone, highlighted by the tragic conflict between Creon’s point of view – the point of view of polis – and Antigone’s vision, namely her individual calling for transcendent codes and laws. In Sophocles’ Antigone, Creon, the ruler of Thebes, and Antigone, his niece, are severely divided over the burial of Antigone’s brother Polyneices—who is an enemy of the state. For Creon, a proper burial would undercut national unity, stability, social order, and state authority. For Antigone, to leave Polyneices unburied would violate gods’ laws as well as the deep pull of kinship commitments.

Hegel represents this conflict as a contest between the law of the gods (and the family) versus the law of the state—both of which he argues, as expressions of justice, are partial and incomplete. He writes on this subject:

“Here, familial love, the holy, the inward, belonging to inner feeling, and therefore known also as the law of the nether gods, collides with the right of the state...[Creon] maintains that the law of the state, the authority of government, must be held in respect, and that infraction of the law must be followed by punishment. Each of these two sides actualizes only one of the ethical powers, and has only one as its content. This is their one-sidedness. The meaning of eternal justice is made manifest thus:

*I follow here Winslow and Winslow (1991), p. 313: “… when people have resolved in their own minds what course best fits their considered moral judgments, but who then find themselves in conflict with others whose best moral judgments have led them to an opposing position.”*
both attain injustice just because they are one-sided, but both also attain justice. (Hegel, 1990)

But Creon and Antigone represent also, according to a probably more persuasive Martha Nussbaum interpretation (Nussbaum, 2001), two practical universes strongly limited, that result in two different strategies of defending and simplifying one-way values. Nussbaum considers these two characters morally shortsighted not only because of their value monism (as Hegel already brightly pointed out), but especially because they are obsessively focused on avoiding conflict. Oversimplification of values put in action when we tend to eliminate conflicts and their obligation, fades away when we recognize to compromise a public constitutive function. A function that neither Creon nor Antigone take into consideration. As Nussbaum states:

"... Both Creon and Antigone are one-sided, narrow, in their picture of what matters. The concerns of each show us important values that the other has refused to take into account. On this issue Hegel’s famous and frequently abused reading is correct. Hegel erred, perhaps, in not stressing the fact that Antigone’s actual choice is, in the play’s terms, distinctly superior to Creon’s; but his general criticism of her neglect of the civic is not, as we have seen, incompatible with this recognition. Hegel, however, locates the deficiency of the protagonists in this narrowness or one-sidedness alone, not in their conflicting-avoiding aims. The elimination of conflict is, for Hegel, both an acceptable and a plausible aim for a human ethical conception... from our study of the two protagonists we might infer that to do justice to the nature or identity of two distinct values requires doing justice to their difference; and doing justice to their difference – both their qualitative distinctness and their numerical separateness – requires seeing that there are, at least potentially, circumstances in which the two will collide. Distinctness requires articulation from, bounding-off against. This, in turn, entails the possibility of opposition – and, for the agent who is committed to both – of conflict.” (Nussbaum, 2001, pp. 67-68).

If, as it seems in a very broad sense, human being are both capable of reason and vulnerable to luck, in need of a rich plurality of life-activities (the “human exposure to luck”, as in Nussbaum vocabulary), we can simply seek to limit those risks for the sake of stability of life and shape our lives and our principles in order to avoid in the ordinary experience most serious conflicts. One way to achieve this objective, namely to avoid the conflict, is to simplify the structure of one’s value-commitments (as in the case of Antigone). There is, in my opinion, an alternative option, which provides criteria for choosing between conflicting values in public space, without postulating the elimination of conflict, and for keeping pluralism5.

---

5 See also on that point Kekes J. (1993), p. 58: “What matter is that pluralism is a theory about good lives. According to it, good lives depend on both personal satisfaction and moral merit, and personal satisfaction depends on the realization of both moral and non-moral values”.
If our public commitment to a plurality of values always leaves open the possibility of conflict, many conflicts that at first seem intractable can themselves be surmounted with an idea of moral compromise that should be an element of assertion, instead of being seen as an element of threat of individual integrity. But individual interests, lack of resources, limited informed preferences, contingency, uncertainty, and moral complexity, can contribute to increase and raise the level of conflict and disagreement within institutions and societies: when you have to make shared decisions (Moreno, 1995) on controversial issues you will probably face a moral conflict. This can be due to different and often divergent beliefs and principles of the individuals involved in the decisions making process. In those cases it could be interesting to take a closer look at the notion of compromise.

2. A new way of compromising

Compromise has been traditionally considered a viable strategy for reconciliation. Indeed, compromise is usually connected to a process of negotiation, by which each of the parties in disagreement concedes something for the sake of something else considered of more value. However, the traditional interpretation seems to neglect certain crucial aspects which render the idea of compromise something more than a mere strategic concession. In a compromise, persuasion, reciprocal respect, and favorable dispositions towards mutual concessions seems to be crucial. In other forms of negotiations, power, threats, and other kind of pressures are on the contrary fundamental. The parties involved in a compromise are committed to consider others' arguments and to be in connection with them through various forms of discussion. They have to be competent and ready to give concession at least to reach the end of the decisional process.

If we take into account the basic definition of compromise, we can see that compromise is a way to respond to conflict with a deal that requires and implies a mutual concession, to get to a better condition of status quo ante. According to this first definition, compromise is a sort of agreement among different points of view, reached thanks to mediation (or a mediator) or thanks to reciprocal allowances. To embark on a compromise means to accept that a trade-off exists involving my wishes/values and different other possibilities. Then the idea of compromise can seem to have the pejorative implications of my second and
negative sense of compromise, and may inspire in some mistrust or even rejection, as though it inevitably fails to fully reflects the values rooted in our cherished ideals. A classic compromise is an agreement that parties reach because it serves their interests better than the status quo but also an agreement in which all sides sacrifices something and in which the sacrifices are at least partly determined by the other sides’ will. Critically, this sacrifice “involves not merely getting less than you want, but also, thanks to your opponents, getting less than you think you deserve,” that is, less than you feel you should have received if only your opponents had been more reasonable (Gutmann and Thompson, 2012, pp. 9-10). Thus, in classic compromise, parties make a strategic calculation that a sacrifice of some their commitments is the best tactic to advance their other values and interests over their current positions.

But if we try to figure a compromise as a way to get to agreements that – firstly – can minimize harm and promote mutual respect between individuals⁶, and secondly could increase the level of cooperation between them, we can consider a compromise shaped in that way as an agreement that can translate at its best some of the individual’s deliberative ideals, as fairness, mutual respect, and equality of opportunity, in order to influence the outcomes of the agreement. This kind of compromise will be different from simple negotiation, and more demanding to each one of the parties. It will require first and foremost a reciprocal recognition of moral fairness. It will represent a rational motivation to critically and socratically revise our automatic engagements towards the common world and to the typical features of political action, as its singularity, uniqueness, novelty, unpredictability, contingency, uncertainty (Arendt, 1998).

This new way of compromising in its broad form is based not on competitive self-interest, bargaining power, or internal reason about the merits of a dispute or the best way to solve a problem. Rather it aims to be a performative act that involves embedding conflict deeply within disputants and within the emotions and values they experience as constitutive of self and social relations. In this sense

---

⁶ See on that point also J. C. Smart, “Negative Utilitarianism”, *Boston Studies on the Philosophy of Science*, vol. 117, 1989, pp. 35-46: “The negative utilitarian principle is that we should minimize the amount of suffering and unhappiness in the world”.

compromise could become an expression of higher-order values such as respect, recognition, solidarity, or community.

3. Compromise and Integrity

So when and how real people locked in conflict are at times willing to revise and relinquish some their desires—be they material interests and/or matters of principled concern—for reasons of mutual respect, community, and other kinds of higher-order moral, political, and procedural ends?

In *The Theory of the Moral Sentiments* Smith introduces the notion of an *impartial spectator* and appeals to the reactions of such a spectator as setting the standard for our moral judgments: the impartial spectator, properly understood, sets a standard that endorses actions and institutions in proportion as they contribute to the public good or over-all happiness. In other words, for Smith in a public setting we enter into another person’s situation rather than into their feelings: impartial spectator is expected to become as well informed as possible about the circumstances at hand, and to remain as fair as possible in spite of his natural biases. Impartial spectatorship does not require a rawlsian veil of ignorance: the impartial spectator does not forget his own conception of the good life when he crosses the public floor, but he understands when a virtue (especially justice) calls for restraining his demand that others act in accordance with his conception. This result in a suspension of moral judgment of actions until after an imaginary exchange with the actor has been attempted. Crucially, sympathy depends on exposure to details, on familiarity with specifics, on the integrity of other people. On the other hand, if I begin to sympathize with another, and make an effort to place myself in her circumstances, this can serve as an independent motivation and encouragement for endowing her with the same rights I have. Exposure to context does not eradicate the possibility of moral criticism, does not render us incapable of deciding or acting; it just makes it more likely that we will decide and act carefully, pragmatically, reciprocally.

If we consider the sophisticated and attractive account of moral judgment that lies behind Smith’s specific substantive judgments and the standards he defends, we can conclude that it is an account that explains the emergence of our capacity to think in moral terms, mobilizing standards able to distinguish between accurate and inaccurate moral judgments.
Indeed, when we – for instance - discuss with someone whose opinions are radically different from ours, we can never be sure that we are dealing with ultimate and fundamental values that cannot be modified by that discussion. Not only we don’t know this of our opponent, but we don’t know this even of ourselves. Nobody can rely on some definitive and perfect map of one’s normative system.

The model of compromise on which I would focus requires – as Smith’s theory of moral judgments proves - considering the complex ways in which the terrain of human exchange is not fully captured by an oscillation among interests, reason, and rights. I follow on this point the important issues and questions raised by Martin Benjamin (Benjamin 1990, p. IX) who, in Splitting the Difference. Compromise and Integrity in Ethics and Politics asks if can we ever compromise on matters of ethical principle without compromising our integrity, and if can men and women who value their integrity commit themselves to the vocation of politics if, as a matter of course, it requires compromise on ethical issues.

A central claim of my argumentation is that both integrity and hypocrisy are inextricably connected with compromise on account of the peculiar character of political relationship. Political relations, ordinarily understood as power relations, can just as readily be conceived as relations of dependence (Grant, 1997). There are dependencies among people who require one another’s voluntary cooperation but whose interests are in conflict. “In such a situation, trust is required but highly problematic, and the pressures towards hypocrisy are immense. Because political relations are dependencies of this sort, hypocrisy – as we use to say - is a regular feature of political life, and the general ethical problem of hypocrisy and integrity is quintessentially a political problem” (Grant, 1997, p. 3). In public and political relationships, one must beware of trusting others. Political relationships are not true friendships, and what we have to learn about politics is that - if we follow Grant’s argumentation – because society requires

---


8 See again Grant: “Some of the most serious and most frequent political conflicts arise among people who share basic guiding principles but part company over when to take a stand on principle and when to accept a compromise. How can the distinction be made between a legitimate compromise and a sellout, idealism and fanaticism, statesmanship and demagoguery, or moderation and rationalization in defense of the status quo? How can we determine the moral limits of both moderation and moralism in politics?” (p. 3).
trust but men and women are not always trustworthy, deceit is inevitable. And because society requires morality but men and women are not always moral, hypocrisy is inevitable, and ironically the frequency of hypocrisy in politics testifies to the strength of the moral impulse in public life:

“Though political hypocrisy in many of its forms is morally reprehensible and politically dangerous, its necessity indicates something positive nonetheless. Hypocrisy only occurs where people try to appear better than they are. The pretense is only necessary where people need to be thought of as good and to think of themselves as good. Where there is political hypocrisy, there is a public moral standard and a significant moral impulse. The necessity of hypocrisy in politics is one indication of the enduring strength of that impulse in human life” (Grant, 1997, p. 53).

In other words, if we try to put hypocrisy (and integrity) first – as in the Judith Sklhar argument (Shklar, 1985) – we can conclude that hypocrisy is inevitable, even necessary, for (especially democratic) politics to function and for political actors to preserve themselves, and that – *vice versa* - extreme anti-hypocrisy can make people vulnerable to fanaticism and blind them to cruelty. Hypocrisy requires moral pretense, and that pretense I necessary because politics cannot be conducted solely through bargaining among competing particular interests. To argue that political hypocrisy is necessary is thus, as Grant states, to argue that moral cynicism as a public principle is impossible. On the contrary when integrity is viewed as purity, as a matter of fact, anything less may be condemned as unpardonable. As Hannah Arendt has convincingly demonstrated, this is the model that inspired Robespierre and Saint-Just. Fanaticism was the product of adopting moral integrity as the exclusive guide for political action.

On one hand, politics necessarily involves matters that are not only negotiable in the manner of interests – e.g. trust, loyalty, pride, honor, vanity, ambition, moral belief. On the other hand, we are still learning from Machiavelli that not only good men make good societies, but also bad men make good societies. We have to reevaluate – together with sympathy - in our definition of compromise the public role of hypocrisy aforementioned, if is it shared and accepted by both parties, if it helps to address imbalances and vulnerabilities, re-establish relationships, reconcile priorities that are worlds apart. Only in that

---

9 See Grant: “Machiavelli, so often understood as a theorist of “power politics”, rest his case for hypocrisy in politics precisely on this view. The necessity for hypocrisy arises whenever potential competitors depend on one another. And politics itself is understood as the situation where potential competitors depend on one another” (p. 55).
sense hypocrisy could perform some positive role in public, and could be a valid resource of freedom:

“So long as there are public moral standards, there will be, and sometimes even should be, hypocritical manipulation of them. There will be hypocrisy if only because there is necessarily a gap between the real and the ideal. Human beings cannot always practice what they preach. There will be hypocrisy too whenever there are justifiable exceptions to the rules and a simultaneous need to maintain public commitments to the rules. Most important, there will be hypocrisy because it works in situations where neither force nor honesty is a viable alternative” (Grant, 1997, p. 50)

As such, in the case of public moral disagreements problems may require multiple forms of political and social contestation so that people can re-conceptualize their interests and rights before reaching consensus—or, rather, before they engage in compromises that are irremediably shaped by social and structural constraints.

If we reconsider – according to Smith’s sympathy and Grant’s hypocrisy - the notion of integrity away from a strict fidelity to personal principles (McFall, 1987) and more related to one of social responsiveness to irreducible human plurality, we will have “a self that evolves in response to changing conceptions of the good over time, and when the self is viewed in this manner, what emerges is not an impermeable core of commitments, but rather a web of commitments that one must somehow weave together... Given a self that evolves and develops over time, moral compromise can be seen as an expected feature of one’s experience in a world of moral complexity... It is this critical perspective and reflection that allows one to draw upon and integrate the diverse commitments of the self. Loss of integrity therefore does not arise from a change of beliefs or values per se, but rather as a result of unreflective change.” (Goodstein 2000, p. 809). Moral complexity, in this case, can become an occasion for acknowledging the manner in which a moral pragmatic compromise can affirm integrity.

A compromise, in this sense, could basically represent our moral position more than our own moral choices do: rather than weaken our moral limits, a

---

10 McFall (1987): “Personal integrity requires that an agent (1) subscribe to some consistent set of principles or commitments and (2) in the face of temptation or challenge, (3) uphold these principles or commitments, (4) for what the agent takes to be the right reasons”(p. 9).

11 “Our concernin g is for the long run and our aim is to live and to have lived a good and optimally integrated life in conjunction with others whom we regard as in some sense equals and whose commitments, values, and principles will not always be ours” (May, L. 1996. *The socially responsive self*. Chicago: University of Chicago Press, p. 24).
moral compromise can make us more responsible in supporting these limits in the future.

Thinking about compromise in the way I propose here – pragmatic as in James’ terms - imply taking this interpretation into account, and entails thinking about the normative principles that bind people into communities, about experiences of commitment and about a model of Lockean positive altruism: political field is not our ultimate and unique perimeter of life (Locke, 1966), and take account to others means firstly respecting them. On this wake, Margalit’s decent society (Margalit 1996) offers a sustained, innovative and well-informed discussion of this issue. A decent society, in Margalit’s view, is a society whose institutions do not humiliate its members. He presents the logical, moral and cognitive reasons for choosing a sort of reasonable compromise: it is not justice that brings us to politics but injustice – the avoidance of evil rather than the pursuit of the good. In contrast to the elusiveness of the abstract notion of human dignity, the phenomenon of humiliation is tangible and instantly recognizable; so too is the notion of evil associated with it. Heading off evils and not the attempt to realize an ideal condition of justice should be the central focus of political thought and action. If we start from dealing with evils rather than striving for an ideal good—as Margalit propose – it is clear that in a decent society many types of injustice would be corrected; it is no less clear that remedying injustice is not the same as moving toward a condition of perfect justice. But his point is not that theories of ideal justice (such as those of John Rawls, for example) should be replaced by a philosophy that focuses simply on making the world less unjust—a position set out in Sen (2009). Margalit’s argument (Margalit 2010) is different and more radical: the struggle for a decent society requires compromise, including the willingness to accept a less just world where this is necessary in order to stave off greater and disrespectful evil.

The bond between moral compromise and integrity will be therefore considered in the case of pragmatic compromise from the point of view of an agent who makes commitments and intakes some (weberian) ethics of responsibility. This ethics is characterized as following: it takes into consideration the singularity of individuals, in practical contexts and circumstances; it aims to
Compromise and/or Integrity? Some Normative Remarks

responds to other’s needs; it is interested in the present. Precisely because compromise expresses higher-order or extrinsic values in this way, it is also distinct from reasoned convergence on the merits of the dispute. That is, compromise does not happen when parties come together in their deliberations on a right answer to resolve a conflict or the best way to solve a problem. It is more often an expression of value pluralism—a recognition that one’s opponents’ claims are also legitimate and therefore deserve a measure of accommodation. Such an accommodation, however, typically depends on an additional source of justification—for example, moral respect, community, obligation, paternalism, solidarity, love—that parties experience as transcending the dispute. As Henry Richardson puts it, for example, a “reasoner” does not compromise when she is convinced of the truth or soundness of another’s argument; rather she compromises when she agrees to change her position to one she “would not have accepted but for [a] kind of concern and respect.”

The kind of integrity and compromise that I’m trying to suggest attempts to answer to the following challenge: cognitive diversity should find public voice even when it affects moral and divergent beliefs. Pragmatic compromise can be seen as the more convincing possibility of cognitive confrontation between comprehensive, conflicting and controversial views of the world: in this sense, it may help to reduce disagreements and to shape a version of Weberian political integrity.

We can argue that rather than an expedient to manage the competitive assertion of self-interest, a pragmatic compromise is a moral response to the problems of reasonably competing values, perspectives, and resources that is grounded in a first principle of mutual respect.

4. Conclusion

12 H. S. Richardson, Democratic Autonomy: Public Reasoning about the Ends of Policy, Oxford, Oxford University Press, p. 148, 2002: “Compromise... is different from simply coming to appreciate the other’s arguments and modifying one’s view accordingly”.

13 Ibid.

14 Margalit (2012), p. 50: “Does compromise always require something like splitting the difference? Not quite. There is a notion that views the essence of compromise not so much in splitting the difference as in the willingness to accept a redescription of what is in dispute. For example, if Jews and Muslims would agree to redescribe their dispute about sovereignty over the Temple Mount in Jerusalem in terms of a dispute about the use of the place, then the main compromise step is already taken; splitting the use is trivial, whereas splitting sovereignty is extremely hard”.

On this new pragmatic and reasonable account, in the case of Sophocles’s tragedy a compromise would require Antigone and Creon to reconcile competing principles—which they would do, if at all, not only by rationalizing how the other’s principle accomplishes justice, but also by experiencing a measure of respect, even solidarity, for the overlapping communities of kinship and state they share. Perhaps Creon would be motivated by the love of his son, who is betrothed to Antigone, and through his son, feelings for Antigone; perhaps Antigone would experience a countervailing set of familial commitments to support her uncle in his efforts to consolidate the state. And through these transcendent values and commitments they would endure, if not resolve, feelings of loss. This kind of compromise—perhaps a decision to bury Polyneices in a remote location outside the state—resembles what Henry Richardson calls “deep compromise.” This is an agreement that happens when rivals revise and modify their ends—a possibility, Richardson argues, that “by definition, exists only for groups of people in which mutual respect, concern, or shared identity is strong enough for them to try to work together.” It also resembles what Avishai Margalit calls a “sanguine compromise,” an agreement that “involves painful recognition of the other side, a giving up of dreams, making mutual concessions that express recognition of the other’s point of view and that is not based on coercion of one side by the other” (Margalit 2012, p. 67). In this sense pragmatic compromise could be part of our democratic citizenship, because it will be helpful to overcome incompatible differences between individuals, respecting them without suppressing them.

Our conclusion is that one’s character is not solely measured by their ideals, but also by how one is willing to compromise. There are, of course, values that should not be compromised. But for the sake of peace, often we must compromise our upper hand even when we are certain of the truth. The art of compromise is far from easy and all of us have much to learn. So much more work needs to happen on the global, national, and interpersonal front. Our task is not merely pragmatic (spreading peace and regulating political disagreements concerning the justifiability of institutions, laws, and norms, which may cause social conflicts when not applied correctly) but also epistemic (learning to humbly see some truth in opposing positions). It is the lesson of John Locke: politics is significant, but not ultimate. Politics does enormous good, but it is not the only sphere in which to do good and the good it achieves is never ultimate.
The moral pragmatic compromise could therefore represent a paradoxical political paradigm: it is what makes conflict possible – conflict made perpetual, stable, cultivated. Since, as we know from Machiavelli’s political theory, common grounds can originate precisely only from division and conflict. In other words, if reconciliation (of values, opinions, and worldviews) is produced by way of rupture, the case of pragmatic compromise made explicit that what bonds individuals with different backgrounds together could be a sort of competence, expertise, towards conflict, the awareness that bonds originate in the discontinuity of a division that is shared.

**References**


Benjamin, M. (1990), *Splitting the difference*. Lawrence, University of Kansas Press.


Smart J. J. C., “Negative Utilitarianism”, *Boston Studies on the Philosophy of Science*, vol. 117, 1989, pp. 35-46:


THE VIRTUE OF COMPROMISE
A VIRTUDE DO COMPROMISSO

Yuval Eylon*
yuval@openu.ac.il

Abstract. Compromise is the virtue of political agents. This picture of the political is as common and familiar: politics is a realm where the reasonable, the compromising politicians get things done and the unreasonable and uncompromising are doomed to fringes. Thus, it is always right and reasonable make good compromises. The paper argues that under certain conditions, it is better and more effective to have non-compromising politicians. For example, think of a political party that every election moves towards the political center to maximize its chances of winning, but loses the elections at the cost of having the political center move further away from its original positions. If the process repeats itself, then a series of compromise would be disastrous, much as the considerations of the self-torturer are disastrous. Thus, there are systematic ways in which the reasonable compromises of a virtuous politician are sometimes (ultimately) unreasonable. Political virtue is all too often self-defeating, and therefore a curse in disguise.

Keywords: Compromise, self-torturer, virtue, political virtue.

Sumário. O compromisso é a virtude dos agentes políticos. Esta imagem do político é comum e familiar: a política é o domínio onde os políticos razoáveis e capazes de chegar a compromissos resolvem problemas, e onde os políticos pouco razoáveis e não dispostos a chegar a compromissos são relegados para as margens. Assim, é sempre correto e razoável chegar a bons compromissos. Neste artigo, argumento que, sob certas condições, é melhor e mais eficaz ter políticos que não fazem compromissos. Por exemplo, pensemos num partido político que a cada eleição se move mais para o centro do espectro político para maximizar as suas probabilidades de ganhar, mas perde as eleições e isto à custa do crescente distanciamento do centro político em relação às suas posições originais. Se o processo se repetir, então uma série de compromissos seria desastrosa, tanto quanto as considerações do auto-torturador são desastrosas. Assim, há algumas formas sistemáticas através das quais os compromissos de um político virtuoso não são (em última análise) virtuosos. Frequentemente, a virtude política é autodestrutiva, e é por isso uma maldição disfarçada.

Palavras-chave: Karl Polanyi, Friedrich Hayek, safety net, pauperism, liberal creed.

* The Open University of Israel.
1. Introduction: in praise of compromise

Compromise is the virtue of political agents. This picture of the political is as common as it is familiar: politics is a realm where the reasonable and the compromising rule and the unreasonable and uncompromising are doomed to fringes. This piece of conventional wisdom is famously summed up by Bismarek's often quoted dictum - "Politics is the art of the possible, the attainable — the art of the next best". The saying is an apparent truism: It is always right and reasonable make good compromises, because a good compromise is simply that which achieves the most from existing possibilities (May 2005). Thus, the virtuous politician is the politician that has mastered the art of the possible, the attainable, even if this means a willingness to compromise various principles, ideals, values, or even allies. Indeed it might come to be that the best course of action dictates adopting a compromise that is a far cry from what was originally sought, perhaps even proclaimed and publicly committed to.

My aim in this paper is to argue that sometimes, the road to hell is paved with good compromises. In other words, that (in certain types of circumstances) it would better and all things considered more effective to have non-virtuous and politicians. The claim is not the observation that sometimes even the best politicians fail, but rather that there are systematic ways in which the reasonable compromises of a virtuous politician are sometimes (ultimately) unreasonable. In other words, political virtue is all too often self-defeating, and therefore a curse in disguise.

The view that compromise is a political virtue, if not the political virtue, applies to different types of political activities and actions. From the straightforward deal struck in backrooms negotiating support for a proposed law, to developing a platform prior to an election campaign that will appeal to many voters, to striking deals with different voter groups. As a matter of fact, it is hard to conceive of any area of political activity in which compromise is not required.

---

1 May argues that compromises are never good simply because they are compromises, irrespective of their consequences. I will not take a position on this issue here. In what follows it is assumed that a good compromise is good because it is a preferable course of action irrespective of its being a compromise.
What does the virtue of compromise consist in? I shall not attempt a full response to this question here, but a preliminary sketch of a response might mention in addition to a willingness to make concessions regarding principles, ideals, values and allies also what is involved in identifying, making - as well as publicly justifying - reasonable compromises: sound judgment and assessment of risks and opportunities, matching available means to achievable ends, addressing public opinion, appealing to it and if necessary challenging it, and so forth. In particular, the virtuous politician has the ability to strike a deal, the right deal, at the right time. This is probably universally true, but even more clearly so in democracies. Democracies come in different sizes and shapes - they have different structures, different electoral systems, a variety of political cultures. But whatever the system, building coalitions and bargaining are always a key, plausibly the key, to political success at every level of government.

The description of the virtue of compromise above comes close to a rough characterization of reasonableness in the realm of politics. This is hardly surprising, because a reasonable or virtuous politician must possess the virtue of compromise. In line with this, it is worth noting that that ordinary usage tends to assimilate the two virtues - and the requirement to "be reasonable" usually implies "make a compromise" or "do not be so stubborn (or dogmatic or "such a purist")" rather than the opposite (note the initial implausibility of something like "be reasonable - don't compromise!").

For our purposes it is not necessary to undertake the difficult task of disentangling the two virtues. Reasonableness in politics is more general, and includes more than the disposition to make a good compromise and strike a good deal, but I think it is not contentious that compromise is often a significant component of reasonableness. Thus, we can maintain that it is typically reasonable to make a good compromise, and that a reasonable politician also knows when a compromise is not available. In what follows I will treat the two together.

2. The good and the Electable

The very ideas of virtue in politics in general and of political virtue in particular, give rise to various puzzles. One important puzzle concerns the relation between the qualities required for getting elected into office, and those
we expect from people who serve in office. Williams asked how can the good
enough can rule the world as it is, and presented the conundrum thus:

It is widely believed that the practice of politics select’s at least for cynicism
and perhaps for brutality in its practitioners. This belief, and our whole subject,
notoriously elicit an uncertain tone from academics, who tend to be either over-
embarrassed or under-embarrassed by moralising in the face of power. Excited,
in either direction, by the subject, they often take rather large-scale or epic
examples, such as the conduct of international relations by hostile powers, or
ruthless policies which may or may not be justified by history. I will touch
marginally on those kinds of issue at the end, but my first concern is more with
the simply squalid end of the subject, and with the politician not so much as
national leader or maker of history, but as professional. I shall defer the more
heady question of politicians being criminals in favour of the more banal notion
that they are crooks (Williams 1981, 55).

What concerns Williams in this passage is the wide gap, even schism,
between what it takes to get into office, and what we can expect - even reasonably
and realistically expect - from those holding office. Initially, the problem
inherited from Plato is how can the morally virtuous rule. And it is here that the
schism Williams mentions is at its widest. One way of narrowing the gap is of
course to allow for a conception of political virtue that falls short of moral virtue.
But as emerges later in his paper, even this resolution of the conflict will not
suffice - there is a real tension between what we expect - what we need - from
those who rule, and what it takes (sometimes) to rule.

The important thing to note is that as Williams highlights, we are dealing
with a structural problem - the need for procedures to select and elect politicians
in way that reflects popular will leads to the selection of different qualities from
those needed to rule, at least to some extent. The problem arises from the dual
requirement made of politicians - to get into office, and to actually hold it. The
difficulty in designing a system in which the former selects for the latter is a
persistent problem of political philosophy.

It is also clear that resolving, or more realistically alleviating, this problem,
requires a mix of institutional design and actual ongoing political engagement
and advocacy, as well as per Williams’ suggestion - changes in our political
culture. In particular, Williams suggests that we should recognize the fact that politics is a career, learn to acknowledge and respect the achievements of politicians whose career does not take them to absolute top, and recognize the legitimacy of personal interests and considerations, within reason, during the course of a political career. These suggestions are not merely practical (inasmuch as they are), but reflect an amended and more nuanced view of political virtue, which is richer because it encompasses more than the short questions of who can to get elected and who can rule well, and adjusts the conception of political virtue to suit political agents living a political life.

3. Chicken

A different problem with political virtue bears directly on the virtue of comprising. Game theory as well as common sense suggest that in some situations it is preferable to be perceived as unreasonable and non-compromising (Russell 1959). One well known such scenario is the game of "chicken", in which two parties are on a collision course, and the first to swerve away loses (but survives). In such a game, if one party creates the perception in the other party that it will never swerve no matter what, then this gives them an advantage and forces the other side, if rational, to swerve (Schelling 1980). Therefore, in chicken-like scenarios, there is clear advantage in being perceived as unreasonable.

If actually being unreasonable and non-compromising are conducive, or sometimes even necessary, for the creation of such a perception, then we have an argument to the effect that unreasonable politicians - even outright stubborn and irrational politicians - are preferable as office holders in situations in which chicken-like scenarios are likely.

This problem can also be seen is an complementing Williams' problem: compromise is typically required in order to get into office, but in chicken-like scenarios non-compromise is sometimes needed in order to effectively or better serve in office. This points to a further instability within the idea of political virtue. Whereas in Williams scenarios virtue got in the way of getting elected but was needed in order to rule, chicken scenarios suggest the opposite is also a problem - when it comes to compromise, it might be necessary to get elected, but can sometimes hinder the ability to rule. At any rate, it seems that the qualities
needed in order to excel as politicians in different circumstances vary greatly, and are arguably psychologically incompatible with each other - e.g., compromise and apparent unreasonableness.

The argument from the chicken-scenario to the conclusion that an uncompromising nature is superior in rulers was based on the assumption that typically, in order to create a perception of unreasonableness, genuine unreasonableness is required. An impasse seems inevitable, since the some aspects of office require reasonableness and compromise, and others unreasonableness and noncompromise.

However, all is not lost. The problem calls for an institutional solution: red limes can be backed by institutions designed to uphold them (Schelling 1963, Quinn 1990). A paradigm example would be an automated defense system programmed to launch a nuclear response if it detects any enemy attack. If such a system is known to be in place and it is known that response is guaranteed if a red line is crossed, then presumably a credible threat is created. Examples are not confined to automated systems and can include various institutions whose function and ethos are structured around response. Thus, a punitive system designed to respond when certain lines are crossed (with zero tolerance, absolute liability, etc,) can function as well to create the presumably required unreasonable and noncompromising perception of the penal system. The same holds for a military establishment. In such cases, the existence of institutions that "take care" of the creation of the required perception means that unreasonable rulers are not required (or not as required). This allows for a return to a more stable conception of political virtue that does involve reasonableness and compromise.

The significance of separation between the reasonable ruler and the unreasonable response is demonstrated by nuclear deterrence: arguably the threat of mutual destruction prevented war, although if war was to break out it would have been irrational to use nuclear weapons, even as a last resort. This tension is utilized by Arthur C. Clarke in the short science fiction story The Last Command (Clarke 1973). In the story, the crew of a nuclear space station, which is part of a second strike capability designed to ensure the enemy never launches a nuclear strike, receives the signal that nevertheless, war has broken out. The
leader is heard in a recorded message explaining that if they hear this message, then the country has been wiped out. The deterrent has failed. However, the leader orders the crew not to fire the missiles in retaliation, because destroying what is left of the world would be "unworthy of reasoning men", and instead surrender to the enemy (i.e. to the president of the United States).

In summary, thus far we have looked briefly at two paradoxical or quasi-paradoxical issues that threaten the conception of political virtue - Williams' claim that those fit to rule are unlikely to get elected, and the game theoretical scenario in which those unfit to rule are likely to rule more effectively than those who are fit to rule.

4. The Virtue of Compromise and Dirty Hands.

Let us set aside the problems discussed above, and return to the virtue of compromise and our starting point, which is the claim that the compromising politician is the politician that identifies reasonable compromises and is capable of making and publicly justifying them. We must guard against vicious circularity here: it is assumed that we have some grasp of what is involved in being reasonable which is not reduced to simply defining "the reasonable person" as someone who makes the reasonable or right compromises.

We can recognize in the compromising person modes of reasoning, temperament, non-dogmatic outlook, sensitivity to reasons in general, and sensitivity to the features of a particular situation as well as not being overly zealous in adhering to over-arching principles while remaining committed to them. The key is the ability to judge and act in a given situation, and not follow any pre-set pattern.

Broadly speaking, the virtue of compromise is a mean, flanked by the cynicism of the unprincipled political operative on the one extreme, and the overattachment to principles, ideals, or commitments of the purist or dogmatic ideologue at the other extreme.

Thus, at one extreme, we have the cynical politician. This person is all too familiar, and represents a temptation that politicians supposedly find hard to resist. There are two things to note here. First, that what defines the cynical politician is a lack of commitment to principles, ideals, promises etc. If and when
called upon to justify their actions, such politicians are liable either to adopt the language of necessity - claiming that they have made a necessary albeit painful compromise, or as Williams has pointed out, appeal to grand over-arching justifications (Williams 1981, 62).

The second important feature is that commonly such politicians are treated as if they prioritize their personal interests in advancing their careers over the public good. Arguably, this picture is exaggerated, and is the result of media focus on the personal rather than the ideological, and a tendency to depict political contests as sporting dramas. Contrary to this cynical view, it might argued, even in pursuit of what seem to be personal ends, politicians are typically committed to some world view and it is all easy to depict the compromises - the good and necessary compromises and not only the bad and opportunistic compromises - as cynical and unprincipled.

This objection to the depiction of politicians as an immoral and self-serving lot has some force. Nevertheless, even the staunchest advocate of present day politicians and political culture must surely concede that there is at least more than a kernel of truth here, and that pursuit of personal ends often takes priority over commitments to the public good. I will return to this tension between the good of politicians and (some versions of) the public good below.

At the other extreme we have over-commitment. This is the sin of being blindly committed to principles, ideals, previous commitments, etc. "at all costs" and ignoring contingencies and consequences. We can perhaps distinguish here between two related vices: dogmatism - an unyielding commitment to principle - and purism, which is an commitment to moral purity, or to clean hands, and is thus more self-directed.2

This sketch of the virtue of compromise highlights the tension between commitment to principles as well as ideals and some long term interests, and between consequences and in particular short term consequences (as well as some long term consequences). However, unlike the problem of dirty hands, the tension extends beyond moral principles to ideological commitments, ideals,

---

2 The idea of the politician as making a moral self-sacrifice is presented in Waltzer (1973). This possible view of dirty hands places the morality of the agent as the focus of the political vocation.
political principles, political promises, and alliances. All of these are liable to be compromised if and when the opportunity (or need) arises.

This also means that the vices of compromise, or the cost of compromise, does not always involve the kind of moral cost or even sacrifice that Waltzer (1973) associates with dirty hands. A political compromise can be costly, psychologically difficult to undertake, disappointing, and so forth - without being morally costly in the sense of being straightforwardly morally wrong or in violation of a moral principle.

5. The Self Torturer

Reasonable compromises can sometimes fail. There is no mystery here, and this is of course only to be expected, as even the best laid schemes of men often fail. But the problem I will present is that there are systematic ways in which the reasonable compromises of a virtuous politician are, all things considered (in a sense that will become apparent) - not effective and this sense unreasonable. The problem resembles the problem of the self torturer introduced by Warren Quinn (1990):

Suppose there is a medical device that enables doctors to apply electric current to the body in increments so tiny that the patient cannot feel them. The device has 1001 settings: 0 (off) and 1 ... 1000. Suppose someone (call him the self-torturer) agrees to have the device, in some conveniently portable form, attached to him in return for the following conditions: The device is initially set at 0. At the start of each week he is allowed a period of free experimentation in which he may try out and compare different settings, after which the dial is returned to its previous position. At any other time, he has only two options -- to stay put or to advance the dial one setting. But he may advance only one step each week, and he may never retreat. At each advance he gets $10,000. Since the self-torturer cannot feel any difference in comfort between adjacent settings, he appears to have a clear and repeatable reason to increase the voltage each week. The trouble is that there are noticeable differences in comfort between settings that are sufficiently far apart. Indeed, if he keeps advancing, he can see that he will eventually reach settings that will be so painful that he would then gladly relinquish his fortune and return to 0. 2 The self-torturer is not alone in his predicament. Most of us are like him in one way or another. We like to eat but
also care about our appearance. Just one more bite will give us pleasure and won't make us look fatter; but very many bites will. And there may be similar connections between puffs of pleasant smoking and lung cancer, or between pleasurable moments of idleness and wasted lives.

The problem of the self-torturer is a problem of practical rationality - it arises from what appear to be rational choices in individual cases, but which create a problematic pattern and lead to undesirable consequences.

6. Compromises

Now let us consider a political party facing an election. The current leader of the party, and current candidate, makes a tactical compromise in addressing the voters in order to improve their chances. The candidate yields some ideological ground, compromising some principles and perhaps abandons certain arguments and key concepts in justifying proposed policies. Thus, the candidate "shifts to center", under the assumption that by doing so they will attract more undecided voters, while retaining the party's base.

Despite the protests of apparent purists or dogmatic supporters, the party's candidate has a convincing argument: the proposed compromise is necessary in order to increase the party's chances of winning the elections, and is the best political maneuver available, given the parties values, ideals, and aims. Let us assume also that there is no substantial disagreement about what these values, ideals and aims are, or about their relative significance.

Suppose further that despite its efforts, the party fails to win the elections. However, in adopting the compromise it ceded political ground - important principles, key concepts and modes of justification - to its rivals, and the whole political map has shifted away from its core values towards the other end of the political spectrum. Thus, the imaginary center of the political spectrum has shifted as well.

In the next elections, this choice repeats itself and a pattern is created - a new candidate that replaces the failed candidate of the previous elections convincingly argues that in order to increase the party's chances of winning the elections it must shift its positions towards the (new) center in order to attract more voters while retaining its base, thus moving it even further from its core
values and principles. Again, the argument is convincing and the party runs its campaign according to the proposed strategy. And again, our party fails to win the elections, and again the compromise leads to a shift of the entire political spectrum and the political center have even further away from the party’s core values.

Suppose that this process repeats itself several times. Eventually, our imaginary party runs the risk of losing contact with its original values, and cedes the political map to its rivals. What this means is that a series of rational and justified compromises can lead to a loss of integrity - to giving up core positions and values. Furthermore, it can lead to an ultimate compromise that in advance is clearly unreasonable according to any of the interested agents. 3

At this point one might protest that if we take into account such long term effects as ceding political ground, then in cases in which they such effects are likely the compromise is not reasonable. A *compromise is reasonable if and only if the long term damage is outweighed by the short term advantage compromise just failed as sometimes if the case*, so if this outcome was likely then the compromise was unreasonable.

Here, however, an agency problem raises its head. The biconditional tacitly attributes the decision to compromise, and the virtue of political compromise, to an abstract agent that personifies the political party or political camp. Indeed, from the perspective of the party or of a spectator who is committed to the value and ideals of this party and taking into account the likely consequences, such a compromise would seem wrong.

But the decision to compromise and virtue of compromise do not belong to abstract political parties or to spectators, but to real agents - individual politicians, or arguably collective agents. And the interests - the *legitimate interests* - of individual politicians do not match those of the parties they belong to or lead. If we allow, as Williams implores us to, that politics is also a career and not only a calling, and that politicians can and should give some weight to their

---

3 Note that this does not require a change of preferences. We can postulate that the outlook of the party members has remained unchanged throughout the series. Nevertheless, each compromise seems reasonable at the time it is made. Thus this version of the problem resembles the self-torturer more than those presented by Elster (1984).
own personal interests, then the picture changes. The key claim is that it is reasonable enough for a politician that probably gets only one shot at winning an election, to compromise as in the abovementioned scenario.

Let us illustrate this claim with simple scenario. Suppose a candidate has a fairly small chance of winning an election, say around 20%, and making a compromise of the kind alluded to above will increase these chances to 30% or even 35%. For the candidate, it seems that there is no question what is the correct course of action. Furthermore, this is the course of action the candidate is expected to take, and that at least some party members expect.

Not so for actual party or political camp. Here, indeed, the long term harm may well justify rejecting the proposed tactic and compromise. This is particularly true if in the longer term, only a more principled and persistent approach could lead to victory. Thus, ceding ground here has long term consequences that matter to the party and to some party activists but less so to the candidate and perhaps other party activists. Whereas the short term consequences maybe matter a lot to the party - they are *everything* to the candidate.

The commitment of the candidate to the short term is a matter of institutional design and political reality - the candidate is supposed to be invested in the short term, and try and win the present elections. This is heightened by a political climate in which candidates typically get only one opportunity - they typically get to lead the party only once, unless they win. In fact, if they lose their political career is probably over - thus rendering passing over the compromise political self-sacrifice.

Thus, it is reasonable for the candidate to make the compromise described above, in order to maximize their chances. Not only is this the candidate's only chance, but also it is important to note that the candidate in this scenario is not undertaking some "scorched earth" tactic that will leave the party in ruins if they fail. Indeed, the present compromise might cause some damage to the party, but this damage is not irreparable. Furthermore, the long term risk to the party, even if likely, is far from certain and it will be up to future candidates to stop the descent or alternatively win an election. This means that our candidate is not responsible, at least not fully responsible, for the entire possible chain of compromises. Thus, our compromising candidate is not a complete cynic who
only interested in personal success, but a reasonable and virtuous career politician whose career and personal success do matter to him, and are on the line.

The key here is that we can distinguish, even within a political camp or party, between competing legitimate considerations and interests of different actors. Furthermore, arguably these might lead also to epistemically different perspectives, and thus e.g., to legitimate different perceptions of uncertainties.

7. Imperfect Virtue

If we allow that our conception of political virtue, and accordingly of the idea of a reasonable compromise, is sensitive to different interests of different stakeholders within a political body, then we risk that in relevant circumstances, reasonable compromises are prone to yield unacceptable consequences, and that thus reasonable politicians produce an unreasonable pattern of decisions.

When we encountered the related difficulty with the possibility that in scenarios that resemble a game of chicken, it might be overall more effective to have unreasonable leadership than reasonable leadership, it was suggested that the solution to the apparent difficulty lies in institutional design: creating mechanisms or institutions that enable decision makers to create red lines or deliver credible threats.

The crucial point here is that in the absence of such institutions, it is in the interest of the rulers themselves to create them and bring about a situation in which they exist, thus enabling them to rule more effectively. For example, this would mean that a ruler who is confronted with a crime wave and who believes that the threat of draconian punishments would help fight crime would have an interest to create a system that delivers severe punishments without discretion.

But our present predicament is more difficult to resolve, even in principle. The reason is that the problem in chicken-like scenarios is a problem for those who rule. They are at a disadvantage because they are perceived as reasonable and compromising. It is therefore in their interest to create mechanisms that change this perception - mechanism that can provide credibility to red lines and credible threats. A conflict between the comporting ruler and such a mechanism could arise only if the red line is infringed or the threat fails and the question of
carrying it out arises. Put differently, the interest to create a perceived uncompromising response is designed to guarantee such a response will not be required. A conflict of interests would arise only if the threat failed to achieve its designated end.\footnote{See Eylon (2009) for a discussion of threats and their function.}

Now let us suppose that in order to confront the problem of the serial compromises we introduce some measures such as a party constitution which lists its main principles, or a party court composed of its elders, or perhaps some other strong organs designed to embody and represent long term interests and more principled thinking. When election time comes around, these arrangements are liable, indeed likely, to stand in the way of what the party's candidates and those activists and groups that support them - interested in winning the elections and devoted to this task - will want to do.

This conflict of interests means that the candidates - the party's leadership - have a strong reason to work against such institutions: to oppose their creations if they do not exist, to oppose strengthening them if they are weak, to seek to weaken them if they are strong, and in general support their dissolution. This does not imply that some happy, or at least not too unhappy, balance cannot be maintained between candidates who are interested in short term success and institutions that cater for the long term and keep the candidates in check.

However, when such a balance or structure is lacking or is weak, it is not in the interests of the party leadership to introduce it or strengthen it. The case is quite the opposite. And since, \textit{ex hypotheses}, this is the predicament in which we are in, there is scant hope of recruiting the leadership to promote a solution. Thus, whereas in the case of a game of chicken the apparent paradox can be mitigated if not resolved, in our case the conflict of interests between the short terms interests of the candidates and the long term interests of the party are not easily reconciled by institutional design. In fact, when suitable institutions are not in place, then \textit{eo ipso} there is no recognizable agent that has an interest to bring them into being.

\textbf{8. The Tragedy of Greece}
Examples in which it is contended that political parties ceded essential ground by a series of compromise are fairly common, and I will not elaborate on such examples here. I want to mention a somewhat different case, one that demonstrates a different but related pitfall of a series of reasonable compromises. Namely, that of piecemeal erosion not so much of the ideology of a party or its avowed principles, but a similar process that led to the erosion of its motivation to act and political ability to do so. This apparent erosion involves and is the consequence of further structural factors, different from those that give rise to the need for compromise in our hypothetical elections example above.

Let us look very briefly at Yanis Varoufakis' *Adults in the Room* (2017): his account of the struggle during 2015 to save Greece from the austerity regime imposed on it. I will not question the account presented in the book, but accept it for argument's sake an accurate representation of events. The book recounts the events from January to late August 2015: from Syriza's winning the elections and forming a government with Varoufakis as Finance Minister on the basis of a promise to combat the debt settlements and austerity policies imposed on Greece, to the government's decision to ignore the clear outcome of the July referendum on the Eurogroup's proposal, and accept the proposal.

According to Varoufakis, his strategy in dealing with the debtors was based on the idea that a new and sufficiently improved debt agreement was preferable to withdrawing from the Euro, which in turn was preferable to an agreement resembling the old debt agreement and imposing further austerity of Greece. He believed that the only way to reach an improved deal was to credibly threaten to withdraw (Varoufakis 2017, 108).

In a nutshell, my claim is that the problem with the strategy was that during this period the Greek government faced a series of choices in which they either could quit negotiations and plausibly quit the euro, or continue despite being snubbed and thus also allow the forces that opposed quitting to gain ground. Invariably, the Greek government chose the latter at every junction, and thus undercut its ability to withdraw from the euro, which was an option they initially preferred to continuing with the original settlement.

The book, as Varoufakis himself states, is a testimonial - an account provided by a participant turned witness, who apparently us utilizing his skills
and training as an academic to provide an accurate and as objective as can be expected account of events. Importantly, Varoufakis himself views being in the position to provide this eye-witness account a possible justification of his own decision to enter politics.\textsuperscript{5} In this respect, the book is presented to us as redeeming - even if he failed to deliver Greece from its predicament, the attempt to do so enabled Varoufakis to expose the troika and perhaps lay the foundations for a more successful attempts in the future.

In addition to recounting the events and providing a glimpse into the workings of power in general, and the workings of the EU in particular, the book attempts to explain and assess the events that led to the rejection of the results of the referendum and ultimately to Varoufakis' resignation. Herein lies the book's real purpose, as an apologia. Ultimately, \textit{Adults in the Room} is the argument for the defense of what Varoufakis himself considers to have been a failed attempt to improve the terms of the debt relief imposed on the Greek people.

It is this context that Varoufakis turns to tragedy - in particular Shakespearean tragedy - to provide the moral framework of his account. Varoufakis' account of events turns, ultimately, on the personal weaknesses of his colleagues and partners in government. Whereas the actions of actors such as Lagarde (then at the World Bank) or Shwable (Varoufakis' German counterpart) or Social Democratic supposed prospective allies such as Hollande, are explained by various structural and circumstantial causes (p. 29), those of the Greek politicians and government are also accounted for, at least in no small part, by character and individual psychology. It is here that Varoufakis' many references to Shakespearean tragedy come to the fore - relating the unfolding events to the tragic flaws of various protagonists.

This in evident in particular the subsection entitled "Gazing into an abyss: comrades in retrospect" (pp 457-461) in which Varoufakis sums up both the fatal weaknesses of, e.g., Prime-minister Alexis Tsipras, as well as those character traits that account and justify having trusted him: "His desire to liberate Greece [...] was there. His intelligence and capacity to learn quickly were self-evident.

\textsuperscript{5} See in particular (Varoufakis 2017, 8), in which Varoufakis concludes a meeting with Larry Summers with the claim that if politicians fail him, he will turn whistle-blower.
His enthusiasm for the deterrent I had proposed and the debt relief [...] was real [...] When he had instructed me, in front of our cabinet, to fly to Washington to tell Christine Lagarde we were going to default, the enthusiasm had been authentic." (p. 457) It was only after what he describes surrender under pressure and embracing "big brother" (p. 461) that Varoufakis realizes that "My mistake was to miss other things that were bundled up with them: his back-up plan, which would inevitably annul my work; his frivolity; his tendency to melancholy; and lastly his intense desire to prove to a skeptical world that he was no shooting star." (p. 457)

Varoufakis' narrative is an ethical one and it places ethos at the center of human life, supposedly revealing the lack of character and backbone of his partners. First and foremost in this list of flawed characters is Prime Minister Tsipras. Greatly oversimplified, the narrative boils down to this: after the rejection of all offers to improve the deal, and following the referendum in which the Greek people unequivocally voted to reject the proposed settlement, the government caved in and capitulated. The reason for this decision lay in the lack of moral fortitude, and lack of respect to the verdict of the public. Thus, at the heart of a detailed account of the forces in play and of the economic, political, and ideological pressures and realities, we find a morality tale.

9. The Fault lies in Our Stars

In what follows I suggest there is a different possible lesson from the chronology of events as presented by Varoufakis, namely, that the problems Varoufakis' strategy ran into are primarily structural, and that these structural issues present a problem to the idea of the virtues of compromise and reasonableness. The problem is related to that of chicken-like scenarios, but mainly it is a problem of the erosion and processes that might occur as the result if a series of reasonable compromises.

As claimed above, Varoufakis' strategy was based on the following contention: there are three possible outcomes of the negotiations between the new Greek government and the Troika, ranking from best to worst:

A. An improved settlement

B. Quitting the euro
C. Accepting the current proposed settlement.

Varoufakis wanted to achieve A, and believed that A was achievable. His claim was that A could be achieved only if Greece could make a credible threat that if forced to choose, it would opt for B rather than C. If it could make such a threat, Varoufakis believed, the Troika would also prefer A to B as it was clearly in their respective interests.

This general scheme raises two problems we have been discussing. First, the threat to adopt B unless A has to be a credible threat. In particular, if it is perceived (by the troika) to be an unreasonable course of action - e.g., as politically suicidal, the threat would not deter anyone. Thus, some measures must be taken in order to create the perception that indeed Greece would opt for B rather than C.

The second problem is that if one fails to achieve A, then B must indeed be adopted in order to avoid C. However, it is possible that the strategy aimed at achieving A will create the conditions that ultimately led to adopting C over B.

First, the situation and strategy adopted by the Greek government rested on the assumption that taking option B would seem and appear to be a viable option. To this, certain preparatory steps were taken or supposed to be taken. It is important to note that these steps were not always completed, and at least in some cases they were actively frustrated. Thus, crucially, the type of mechanism that is required in order to create credibility was either completely lacking or at least fell short of grounding a credible threat. Arguably, the best way to create the perception of really preferring B over C and thus pursue A would have been to really begin a process of leaving the euro. However, according to Varoufakis himself, such preparations were constantly thwarted.

One important point that stands out in the account is that from the moment negotiations began, Varoufakis faced a lack of cooperation and repeated snubs by the troika. The relative lack of response to these snubs, could have produced a dual impact. First, it helped erode the perception that the threat to choose B over C if A were unavailable would be carried out.

Second, and more importantly, as each snub was countered by a further compromise and an attempt at making reasonable proposals and arguments, the
resolve to pursue B and choose it over C by the government was eroded. This is not a psychological assessment of the personal will of ministers, although this might have played a role as well, but a claim about the political will of the government. There are multiple factors that can come into play: cooptation of certain actors, political rivalries, political pressure, ideological pressure, fear of change. It is here that we can view the months between taking office and Varoufakis' resignation as a period in which the idea and resolve to pursue B eroded.

Thus, when it became clear that the troika will not yield, the government opted for a referendum on the proposal, under the assumption that the proposed settlement would be accepted. It is a testament of the either the erosion of political will, or an absolute reversal of priorities and preferring C to B. In either case, the it is clear that the government was so committed to remaining that it decided to ignore the results of the referendum it itself initiated, when the position it supposedly supported was accepted. If this is indeed the result of processes that took place while in office, and if assume that the original ranking of options was correct, then it follows that it was better to commit to B as soon as possible and avoid the attempt to pursue A.

10. Conclusion

We started from the idea that compromise is the virtue of political agents. I argued that if we look at the problem of electing worthy politicians, and at chicken-like scenarios, then it transpires that the in certain conditions virtuous politicians and political virtue become a political liability. Whereas the problem Williams presented was between moral virtue (even reasonable moral virtue) and political virtue, the chicken-like scenarios undermine the idea of political virtue. However, both call for an institutional solution.

Following this claim, I have argued that political virtue can encounter an additional problem, and one that is less friendly to an institutional resolution. Namely, that a series of reasonable compromises can lead to unwanted consequences. Put differently, that a series of reasonable compromises can turn out to be unreasonable.
This problem, if it arises, becomes especially stubborn because unlike in the chicken scenarios, the most powerful political agents have an interest to oppose any institutional solution, and might aim to aggravate the problem. This suggests that our conception of political virtue requires broadening to include politicians that are typically more principled than has been assumed in this article, or perhaps that a plurality of characters is an essential requirement of politics.

References


The Virtue of Compromise