



Ethics, Politics & Society

A Journal in Moral and Political Philosophy

Number 1- 2018

ARTICLES

A Multidimensional View of Misrecognition

Douglas Giles

Combining Democratic Equality and Luck Egalitarianism

Lars Lindblom

Norms for the Public Remembrance of Nonhuman Animals

Matthew McLennan

**8TH BRAGA MEETINGS ON ETHICS
AND POLITICAL PHILOSOPHY,
WITH GUSTAF ARRHENIUS**

SYMPOSIUM: HAYEK-POLANYI

ETHICS, POLITICS & SOCIETY

A JOURNAL IN MORAL AND POLITICAL PHILOSOPHY

Number 1- 2018

A Multidimensional View of Misrecognition

Douglas Giles

Combining Democratic Equality and Luck Egalitarianism

Lars Lindblom

Norms for the Public Remembrance of Nonhuman Animals

Matthew McLennan

8th Braga Meetings on Ethics

And Political Philosophy,

With Gustaf Arrhenius

Symposium: HAYEK-POLANYI

Title

Ethics, Politics & Society. A Journal in Moral and Political Philosophy (nº1 2018).

Editor-in-chief

João Cardoso Rosas.

Editors

Alexandra Abranches, David Álvarez, João Ribeiro Mendes, Roberto Merrill.

Guest editors

Alexandra Abranches and Eze Paez for the “8th Braga Meetings on Ethics and Political Philosophy, with Gustaf Arrhenius”; António Batista and Pedro Teixeira for the “Symposium: Hayek-Polanyi”.

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, Jonathan Wolff, Maria Xosé Agra, Samuel Scheffler, Seyla Benhabib.

Edition

April 2018. Centre for Ethics, Politics and Society University of Minho, Braga, Portugal.

ISSN

2184-2582

DOI

<https://doi.org/10.21814/eps.1.1>

Website

<http://www.epsjournal.ilch.uminho.pt>

TABLE OF CONTENTS

Introductory Note

The Editors _____ 7

A Multidimensional View of Misrecognition

Douglas Giles _____ 9

Combining Democratic Equality and Luck Egalitarianism

Lars Lindblom _____ 39

Norms for the Public Remembrance of Nonhuman Animals

Matthew McLennan _____ 63

8TH BRAGA MEETINGS ON ETHICS AND POLITICAL PHILOSOPHY

Introduction

Alexandra Abranches, Eze Paez _____ 85

The democratic boundary problem reconsidered

Gustaf Arrhenius _____ 89

Are moral functionalism's moral a priori commitments really a priori?

Ashley Lane _____ 123

Gender Identities and Feminism

Josh T. U. Cohen _____ 139

Are we post-justification? Stout's case for self-knowledge, political justification and public philosophy

Deven Burks _____ 175

Hobbes: a voluntarist about the permissibility of state enforcement?

Daniel Guillery _____ 203

Basic Liberties, the Moral Powers and Workplace Democracy

Stephen K. McLeod _____ **231**

From Natural Law to the Golden Rule: Aquinas Revisited

Damiano Simoncelli _____ **261**

SYMPOSIUM: HAYEK-POLANYI

Hayek and Polanyi for the 21st century – an introduction

António Baptista, Pedro A. Teixeira _____ **279**

***O credo liberal e os seus críticos: pauperismo e rede de
segurança***

António Baião, J. A. Colen, Pedro Moreira _____ **305**

***The Double Movement in Polanyi and Hayek: Towards the
Continuation of Life***

Filipe Nobre Faria _____ **329**

***Mito, utopia e democracia em A Grande Transformação de
Karl Polanyi***

Patrícia Fernandes _____ **351**

***Polanyi's Economic Embeddedness, Countermovement, and
Republican Political Economy***

Bru Lain _____ **369**

***O debate na grande transformação: Ludwig von Mises,
Friedrich Hayek e Karl Polanyi sobre capitalismo, socialismo e
mercados***

João Rodrigues _____ **403**

INTRODUCTORY NOTE

Ethics, Politics & Society is a new open access academic journal with double blind peer review dedicated to the publication of high level contributions in the fields of Political Philosophy and Theory, as well as Normative and Applied Ethics. Although it is open to all themes and approaches in these areas of knowledge, the journal focuses on issues related to theories of justice, democracy and recognition, as well as on ethical issues connected to scientific and technological development and their social and environmental impacts. *Ethics, Politics & Society* accepts the submission of originals both with a direct contemporary approach, or using the History of Moral and Political Philosophy to shed light on relevant problems in our time.

This first issue of *Ethics, Politics & Society* includes an open section and two closed sections that have been nevertheless subject to the same standards of peer-review. The first open section is composed with original articles submitted by Douglas Giles, Lars Lindblom and Matthew McLennan. The section devoted to the 8th Braga Meetings on Ethics and Political Philosophy opens with the keynote lecture on “The Democratic Boundary Problem Reconsidered” delivered in Braga by Gustaf Arrhenius. This outstanding contribution is followed by a dossier of selected papers presented at the Meetings by Ashley Lane, Josh T. U. Cohen, Deven Burks, Daniel Guillery, Stephen McLeod, and Damiano Simoncelli. This section has been guest edited by Alexandra Abranches and Eze Paez.

The third section consists of a *Symposium on Friedrich Hayek and Karl Polanyi*. This timely and provocative book discussion that combines readings of *The Road to Serfdom* and *The Great Transformation* has been guest edited by António Baptista and Pedro Teixeira, with contributions by António Baião, José Colen and Pedro Moreira, Filipe Nobre Faria, Patrícia Fernandes, Bru Lain, and João Rodrigues.

Ethics, Politics & Society encourages prospective authors to submit their manuscripts in English, Portuguese or Spanish through the journal website, together with the statement that the submitted piece has not been published before and elsewhere. All the papers submitted to *Ethics, Politics & Society* are

subject to the evaluation of at least two reviewers in the corresponding scientific domain. Papers are sent anonymously to blind referees, who are asked to write a review according to the evaluation form adopted by the journal, which includes the following aspects: adequacy to the journal publication standards; adequacy of the paper subject to the scope of the journal; substantive relevance; originality; relevance of the critical methodology; clarity of presentation; arguments and relationship between initial hypotheses and final results.

With the appearance of *Ethics, Politics & Society*, its editors believe that a clear and urgent lacuna is filled in for an international journal that works as a high-level global forum imprinted with the perspective of Southern Europe.



A MULTIDIMENSIONAL VIEW OF MISRECOGNITION

UMA VISÃO MULTIDIMENSIONAL DO NÃO-RECONHECIMENTO

Douglas Giles*
dmgile@essex.ac.uk

Abstract. *Following Axel Honneth, I accept that recognition is integral to individuals' self-realization and to social justice and that instances of misrecognition are injustices that cause moral injuries. The change in approach to misrecognition that I advocate is to replace a macrosocial top-down picture of misrecognition, such as Honneth's typology, with a fine-grained phenomenological picture of multiple dimensions in misrecognition behaviors that offers greater explanatory power. This paper explains why a multidimensional view of misrecognition is needed and explores the various ways that engagement with pathological norms or disengagement from individuals lead to injustices of misrecognition and how understanding behaviors in terms of these two dimensions—norms and individuals—illuminates causes of injustice. The multidimensional view of misrecognition replaces Honneth's binary view of misrecognition as the contrary to recognition without replacing Honneth's conceptions of the value of recognition.*

Keywords: *recognition, injustice, Honneth, social justice, misrecognition.*

Sumário. *Seguindo Axel Honneth, aceito que o reconhecimento é parte integrante da auto-realização dos indivíduos e da justiça e que os casos de não-reconhecimento são injustiças que causam ferimentos. A mudança de abordagem para o não-reconhecimento que eu defendo é substituir uma imagem macrossocial de alto nível de desconhecimento por uma imagem fenomenológica de dimensões múltiplas de comportamentos de reconhecimento misto que oferece maior poder explicativo. Este artigo explica por que é necessária uma visão multidimensional do não-reconhecimento e explora as várias maneiras pelas quais o engajamento com as normas patológicas ou o desengajamento dos indivíduos conduzem a injustiças de não-reconhecimento e a compreensão de comportamentos em termos de duas dimensões - normas e indivíduos - ilumina causas de injustiça. A visão multidimensional do reconhecimento misto substitui a visão binária de Honneth do não-reconhecimento como contrário à noção de reconhecimento, sem substituir as concepções de Honneth sobre o valor do reconhecimento.*

* Associate Fellow, School of Philosophy and Art History, University of Essex.

Palavras-chave: *reconhecimento, injustiça, Honneth, justiça social, não-reconhecimento.*

0. Introduction

Following Axel Honneth and others, I accept that recognition is integral to individuals' self-realization and to social justice and that instances of misrecognition are injustices that cause moral injuries. The change in approach to misrecognition that I advocate is to replace a macrosocial top-down picture of misrecognition, such as Honneth's (2003; 2007a; 2008) typology or Emmanuel Renault's (2011) institutional approach, with a fine-grained phenomenological picture of multiple dimensions in misrecognition behaviors. A multidimensional view of misrecognition is needed because recognition requires both engagement with positive recognition norms and engagement with other individuals. The multidimensional view of misrecognition in terms of the dimensions of engagement with norms and with individuals replaces Honneth's binary view of misrecognition as the contrary to recognition without replacing Honneth's conceptions of the value of recognition. In this paper, I explore how engagement with pathological norms and disengagement from individuals are forms of injustice. I offer a multidimensional view of misrecognition that illuminates misrecognition behaviors and provides greater explanatory power of both recognition and misrecognition by addressing the importance of individuals recognizing each other *as* individuals.

1. Why a Multidimensional View?

Honneth provides us with the valuable insight that justice requires mutual recognition. Recognition is not sufficient for justice, but it is necessary for it. I cannot receive legal rights unless I am first recognized as a person. If I am not recognized as a member of the ethical community, I will not receive the respect that persons deserve. Recognition norms inform us what conduct is proper when we encounter certain types of people in certain types of situations. Engaging with recognition norms is part of our acceptance of our moral responsibility, and lack of a sense of moral responsibility for another is a significant factor in injustice. By

Honneth's definition, recognition behavior is a tracking of specific positive properties in others as measured by means of social recognition norms (Honneth, 2008, 153). For example, society's recognition norms teach me that I should respect honesty and that when I know someone who is honest in his dealings, I recognize that individual as honest and treat him accordingly. Given Honneth's picture of misrecognition as the contrary of recognition (Honneth, 2003; Honneth 2007a), it would seem that misrecognition behavior can be defined as occurrences in which recognition norms are not engaged and applied. If I meet someone who is honest, but I either do not care about honesty or do not recognize the other's honesty, I am not valuing the other appropriately. If I do not consider a right to be important or do not acknowledge an individual's entitlement to that right, I will not behave as required by that right toward the individual, and that is misrecognition.

Honneth's account gives the impression, intended or not, that if recognition norms are in place and engaged with, then there is justice in recognition relations. For two reasons, I think the connection between recognition norms and misrecognition is not a binary one. First is that the norms themselves may be unjust, so following them would lead to injustice. Second is that engaging with the norms themselves is not always sufficient to achieve justice. We need to take these two aspects of recognition into account and extend Honneth's insights into the importance of recognition in justice by clearly indicating all that is required in recognition relations to achieve justice. A closer analysis of recognition norms and misrecognition behavior reveals their complex interrelationship. First, I will look at the ways that individuals do not engage with norms (a lack of vertical recognition), and then I will explore ways in which individuals can engage with norms that, in practice, perpetrate misrecognition on others. It is counterintuitive to think that misrecognition behavior could maintain engagement with recognition norms, but it is the case in some misrecognition behaviors.

Honneth does not develop this idea adequately, but he said that embedded within a discussion of individuals' socialization into their society's recognition order, "subjects acquire the capacity to move about within the normative structures of their social network by treating each other in accordance with the

specific kind of recognitional relationship they maintain with each other (Honneth, 2011, p. 396). I think this points to a core aspect of recognition—that it is behaviors within the constellation of relationships between individuals that must be constantly maintained by the individuals involved. I think that this thought opens up a new emphasis on recognition relations that I wish to expand. Unlike Honneth, I do not think that “a just society requires no more than that subjects learn the various patterns of mutual recognition ‘well enough’” (Honneth, 2011, p. 395). I understand his point—that we need not place a moral burden on individuals to excel at the “art” of recognition. However, injustice occurs when recognition relations are dysfunctional, and recognition relations can be dysfunctional in ways beyond individuals not learning the patterns of mutual recognition—in other words, misrecognition is more than recognition norms not being applied.

To recognize another is to see him or her as a member of the ethical community and to acknowledge his or her experiences as real and worthy of consideration. Stephen Darwall (2010) has argued that the authority to demand respect from others and hold them answerable if they do not provide it are second-personal reasons that operate within a circle of mutually involving concepts of authority and accountability. Similarly, Christopher Zurn (2015) observes that only other recognizing agents can engage in the mutual interactions of recognition and participate in our claims of normative behavior expectations. Mutual recognition and normative behavior require intersubjective involvement with other human beings as human beings. Because recognition is related to certain aspects of another individual, it is a specific response to a specific individual. Because recognition is *by* someone *of* someone, recognition is a relation between individuals. It is reasonable then to consider misrecognition as a dimension of social interaction in which recognition relations between individuals lack reciprocity. Recognition relations are complex, and we need to delve into what is involved in occurrences of misrecognition to craft a more expanded and finer-grained account than Honneth’s account of misrecognition as the contrary of recognition. We need to clarify the relations of recognition norms to injustice and identify which norms contribute to injustice, and we also need to clarify the nature of our relations with individuals who are the recipients of our recognition.

Heikki Ikäheimo (2015) provides a helpful way of conceiving of differences in recognition relations. He first distinguishes between vertical and horizontal recognition. Vertical recognition occurs between persons and norms and institutions. Social norms and institutions exist if and only if individuals recognize them as authoritative (upwards vertical recognition), and social institutions such as governmental bodies recognize persons as possessing rights (downwards vertical recognition). Horizontal recognition occurs between individuals with Ikäheimo distinguishing between two forms: normatively mediated and purely intersubjective. The normatively mediated form of horizontal recognition is one individual recognizing another individual as a bearer of rights or entitlements stipulated by norms for which the recognition is obligatory. The second—purely intersubjective—is a recognitional response to another as an individual person independently of his or her rights and entitlements. Ikäheimo further identifies two modes of purely intersubjective horizontal recognition: conditional, in which concern for the other individual is instrumentally calculated in terms of one’s own interests, and unconditional, in which concern for the other individual is not conditioned by prudential considerations. Ikäheimo’s reason for making these distinctions is to argue that only unconditional purely intersubjective horizontal recognition—recognition that is not of another individual as a bearer of a normative status but as an individual irreducible to functional significance—can be called “love” and “respect.” When we are moved unconditionally by others, that is a genuine respect not mediated by a sense of obligation.

Ikäheimo’s discussion makes clear the real-world benefits of mutual recognition and the harm that comes from misrecognition. Humans are autonomous beings, but one of the essential features of the human life-form distinguishing it from animal life-forms is that humans are governed by social norms authorized by humans themselves. This means, Ikäheimo says, that to live a human life, human individuals must recognize “vertically upwards” some norms as governing their lives, and it also means that they must recognize some others horizontally both in the normatively mediated sense as bearers of the rights, duties, entitlements, and responsibilities prescribed by the norms, and in the purely intersubjective sense as having or sharing authority on those norms (Ikäheimo, 2015, pp. 32-33). Human freedom, then, is not a general

independence from others (which is impossible) or freedom from being determined by anything other than oneself. Concrete freedom is finding oneself affirmed as having authority by other individuals who we affirm as having authority in the unconditional mode of respect and this is the goodness of mutual recognition. Genuine freedom is therefore a practical question of the real-life capacity and propensity for individuals to have genuine respect recognition for each other (Ikäheimo, 2015, pp. 35-36).

How I interpret and apply Ikäheimo's conceptions is to understand that though we must vertically recognize that norms govern our social lives, there is more to applying a recognition norm than an awareness that applying it in one's interaction with another individual is the right thing to do. Recognition requires a set of norms and social institutions to guide it, but recognition also requires that we engage actively in intersubjective recognition relations. In many circumstances, only a particular way of engagement with the individual in his or her distinct circumstances is proper recognition of that individual. I need to tailor my response to the individual in front of me; my recognizing a particular individual is conditioned by his or her individuality. Justice demands that we consider what norms apply to the current situation, plus it demands that the application of those norms be tailored to suit the individuals involved.

Justice requires engagements with recognition in two dimensions, one vertical with recognition norms and one horizontal with individuals, meaning that nonengagement with either norms or individuals could lead to injustice. The demands of recognition, aside from a narrow set of legal relations, go beyond the conventionality of applying norms to groups of people. Justice requires that some forms of legal recognition apply equally to all individuals and, therefore, requires nonengagement with other individuals in their particularities. Basic human rights are invariant, but nearly everything else in social interactions is variant. Even a legal judge needs to apply the norms according to the individual circumstances. We need to modulate norms according to individual circumstances and a range of interpretations dependent on the individual. Within the need for engagement with other individuals, there are public interactions in which engaging with the individual is necessary but in which we do not need to engage the person as an *individual*. For example, in commercial transactions, we

are polite to those who serve us or to the individuals we serve, but we do not always need to know them personally—what Ikäheimo calls normatively mediated horizontal recognition. We can engage with individuals in ways that do not take into account their individuality, though we are still recognizing them as human beings. There are other interactions in which recognition requires engaging with another individual in a way that recognizes the individual *as* an individual. In personal relations such as between family and friends, and in mentoring or other care relationships, only unconditional personal engagement meets the needs of recognition and justice. Unconditional purely intersubjective horizontal recognition should not appear only in the most intimate relationships, however, because intersubjective recognition relations that are based on sincere care for others are sincere expressions of our humanity and are what, more than any other social activity, cultivates self-realization and autonomy, and, thus, justice.

Engagement is at the core of recognition because recognition always takes an individual as an object. Norms exert a constant influence on individuals, and the following of norms can be a nonreflective action. Recognition is different because it requires an intention—an unreflective following of the norms is not a recognition of another. Recognition norms are universal within a culture, but to be operative, the norms must be applied to an individual. Recognition norms are nonspecific and need to be made specific through an expression of intention. General recognition norms point to possible recognition behaviors, but this potential behavior must be made manifest through an individual expression directed at another individual. Because recognition is related to certain aspects of another individual, it is a specific response to the way a specific individual is. Recognition is recognition only if it is the expression of a purposeful intention, not incidental or accidental, but directed and specific.

Recognition is a matter of caring about others in terms of significance, if not fondness. Human interaction that is without care for another's needs and well-being is itself a kind of misrecognition. The mutual nature of recognition calls for an intersubjective engagement between individuals of mutual valuing if not mutual affection. Because recognition is a relation of care, recognition comes more easily within personal relationships. According to Stanley Cavell (1976, pp.

238-266), maintaining social relations requires engaging with another individual such that one is existentially involved in the emotional world of the other individual. The involvement Cavell describes need not be intimate but does need to be what he calls a “stance of acknowledgement” of the other individual that is emotional rather than cognitive. This involvement is a recognitional stance of mutual sympathy through which we come to understand that we have a moral responsibility to react to the other individual in specific ways.

I will explore ways in which recognition relations have gone wrong and lead to misrecognition. First I will discuss misrecognitions in which the problem is in vertical recognition, either disengagement from norms or engagement with problematic norms, and then I will discuss misrecognitions in which the problem is in horizontal recognition in which there is insufficient or improper engagement with other individuals.

2. Dimension of disengagement from norms

If an individual or social institution is not engaging with norms, then misrecognition is a likely consequence. If, for example, we ignore the recognition norm that says that productive labor should be rewarded, then we will not respect those who provide productive labor, which is a misrecognition. Individuals are socialized into social norms and learn their value, so broad general disregard for recognition norms would be rare. It is possible that someone can consciously and willfully disregard all norms, taking herself out of mutual recognition relations despite the cost of becoming a pariah in her community. However, given the essentialness of recognition for social functioning, it is far more likely that nonengagement with norms is a specific disregard within a specific situation. At times, we become oblivious to others in our everyday tasks and morally injure others in our inattention, but we can be prodded into awareness at any moment by seeing how we are not treating another appropriately. An example would be being so focused on one’s current activity, such as driving or walking in a crowded area, that one forgets the rules regarding behavior, being either rude or otherwise negligent in our moral obligations to others. This type of misrecognition is

prereflective and is often resolvable by reflecting on the effects of our obliviousness.

There can also be a more deliberate nonengagement with norms when individuals believe that certain norms are not appealing or convenient within a specific social environment. Disengagement from norms is compartmentalized to suit one's own interests without a rejection of moral responsibility in general. Individuals can rationalize away their responsibility to follow specific norms in specific situations, such as in their workplace. Business owners, for example, can neglect the health and safety of their employees to preserve profits while at the same time being honest in their accounting practices. A supervisor could think that recognizing his employees' needs is unimportant because they are his subordinates but still treat friends and family justly. Workers, perhaps in response to their employers' actions, could believe that norms do not apply, or apply differently, in specific situations on the job. A worker could see dishonesty on the job as acceptable if it increases her productivity. Another could rationalize that his theft of company property is acceptable. Another could feel that she does not owe politeness or camaraderie to fellow workers because it is "just a job." By suspending the application of recognition norms in particular aspects of one's life, one is perpetrating misrecognition.

In the above examples of nonengagement, the norms themselves are not causing misrecognition and injustice. It is also possible that engagement with norms leads to misrecognition if the norms themselves are misrecognitions. In the next section, I identify two types of recognition norms that result in misrecognitions, differentiated by whether the norms are tracking putative negatives, which I call "normative discrimination," or putative positives, which I call "pathological recognition." I will discuss each in turn.

2.1. Normative discrimination

What I call "normative discrimination" is the use of recognition-like norms that designate particular social groups as having negative traits that characterize those groups as deficient and inferior, justifying negative consideration and treatment of those people. These norms are a form of negative recognition that

mediates our interactions with certain groups, dictating that the appropriate response is to deny these groups positive recognition and moral consideration. Because the discrimination is guided by recognition-like norms, individuals who perpetrate such discrimination believe they are behaving properly. The normative character of normative discrimination discourages questioning whether the traits it attributes to targeted groups are actually present in individual members of that group. The negative recognition of normative discrimination differs from a negative response to violations of norms such as disapproving of dishonesty or theft. There, the negative response recognizes the rights of those who have been wronged, and a negative response to wrongdoers is an appropriate upholding of norms. When we punish someone who has been convicted of a crime with imprisonment or fines, we are upholding general norms that recognize the rights of victims of the crime, even if we are denying recognition to the convicted individual's freedom and desire to not be punished. In contrast, a normative discrimination is when a trait that should be neutral to moral norms (skin color, ethnicity, or religion) is taken as a negative and all who hold that trait are regarded as less worthy. Normative discrimination is directed predominantly at social groups separated by race, gender, class, and so on. Additional illusory negatives are often attributed to a social group, such as labeling all Muslims as violent, all Jews as dishonest, all gays as promiscuous, all who live in poverty as deserving their poverty because they are lazy, and so on. How the targets of normative discrimination actually are is irrelevant to the negative stereotypes because the perpetrators follow the norms that dictate behavior toward the targeted groups. Veit Bader's (2007) "criteria of ascription," by which he categorizes structural asymmetries of power and practices of discrimination, oppression, and exclusion as being socially defined and ascribed characteristics of targeted groups, are examples of normative discrimination.

Normative discrimination can be taken to the extreme of a group being considered not deserving of any moral consideration. This misrecognition is beyond a lack of awareness of the moral standing of others and is a conscious antagonism toward others. An individual engaged in this comportment considers the appropriate response to other groups of human beings is to deprive the other actively of recognition as a human being. The extreme hostility toward a number of historically marginalized groups such as Gypsies and Jews in Europe, Dalits in

India, and Burakumin in Japan are examples of this extreme normative discrimination. Someone born into one of these groups is condemned for life to misrecognition. The stigma attached to certain diseases or conditions also fits into normative discrimination. Those afflicted with leprosy were cast out from society as unclean. More recently, those afflicted with AIDS have suffered similar pariah status. In both cases, the ostracizing was accompanied by moral rebuke, the victims condemned as immoral simply for having a disease. One could say these people are rendered invisible, but it is more accurate to say they are condemned as unfit to be included and are dispossessed of rights and status.

The institution of slavery in the Americas is a historical example of how the recognition norms of normative discrimination structure human interactions. Slavery is a relationship defined by a malicious use of power, but slavery in the Americas did not result from taking away an existing recognition relation from members of a community and enslaving them. Indentured servitude and debtors' prisons could be construed as a destruction of an existing recognition relation. An impoverished European (already suffering from normative discrimination because of his or her class) was condemned to a debtors' prison because he or she allegedly violated his or her responsibilities as a member of society. The social institution of debtors' prisons operated as a social relation that was used properly or improperly. An individual wrongly accused or condemned to debtors' prison would be suffering an injustice according to society's norms. Slavery operates under a very different set of assumptions because the normative discrimination based on race denied the possibility of recognition relations. Rather than a rupture of a recognition relation of social inclusion, the slave is, as Orlando Patterson (1982) observes, natively alienated. As Frantz Fanon (1967) observed, racism reduces others to a skin to which they are chained and determined. The recognition norms of the dominant culture are imposed onto the oppressed who are represented through normative discrimination as mere animal bodies unable to think, reason, or speak properly. The Native Americans and Africans enslaved by Europeans were always outsiders to the European slavers and had never been afforded recognition other than normative discrimination. The recognition order of European culture negatively recognized non-Europeans as inferior and uncivilized, and this normative assumption framed European encounters with indigenous people throughout the world. Africans and Native Americans had

never been included, so enslaving them was ethically possible in a way that enslaving Europeans was not.

Because slaves were natally alienated by normative discrimination, questions about slaves' integrity, honor, autonomy, or self-respect were nonsensical to anyone who engaged with the norms that specified what the slaves were—property. Slaves were, as Patterson (1982, p. 263) observes, annulled of rights and identity, without ties to past or future, unrecognizable as human beings, at best shadow members of society. The attitude of impossibility of the slaves' social inclusion preceded the enslavement because the normative discrimination framed the recognition relations with the slave whether the slave was captured or born into slavery. Before the violent act of enslavement occurred, the target, reduced to skin as Fanon said, had been deemed to be compatible with enslavement. Whether the assessment was that the slaves were undeserving of freedom or deserving of enslavement, the misrecognition was a normative discrimination against those who possessed the trait of dark skin and, therefore, lacked humanity, dignity, and rationality. It is not so much that the slave was objectified as a tool as much as it was that the slave was tracked as being of no value beyond menial labor.

Our contemporary society does not have slavery per se, though a Marxist theorist could point to low wages as a form of slavery. Our society retains the normative discrimination of what Andrew Sayer (2009) calls “contributive injustice”—the social misrecognition that restricts what members of social classes are allowed to contribute, particularly in terms of occupations. The lottery of birth restricts most individuals to an inheritance of class distinction that limits their economic opportunities, whereas the fortunate inherit wealth either directly or through privileged opportunities for education, jobs, and careers. As Sayer (2011) observes, public attitudes support the idea that greater contributions to society deserve greater compensation, but the public measures the value of contributions on the basis of class and an unequal distribution of labor. The social structure produces unequal opportunities, with jobs with higher social status and compensation going to a privileged class. Most of the problems of distributive injustice stem from this contributive injustice because low-value jobs are given low-value compensation. Sayer (2009, p. 92) correctly observes that what

individuals are allowed to contribute is at least as important as what they receive in terms of resources. This misrecognition is centered on jobs and occupations, but it extends to educational and cultural opportunities, the health hazards and health care one encounters, where one can afford to live, and all of the lifestyle opportunities that go with these. Contributive injustice is a normative discrimination against others who are not allowed to contribute and not allowed to use their talents and explore their possibilities. A wide range of social groups are negatively tracked and restricted as to the occupations they can enter. The normative discrimination against women as weaker and less rational restricts their occupational opportunities and leads to glass ceilings within occupations. Minorities of race, ethnicity, and religion are also negatively discriminated against and restricted to low-value occupations. Mostly, contributive injustice is tied to class, with labor divided between blue-collar and white-collar, and individual workers are subsumed under the norms that designate their social contributions and status.

Despite the fact that contributive injustice damages society, causing it to miss out on the potential contributions of so many, its injustice persists because the normative discrimination is seen as a proper response to how things are. Sayer (2009, p. 87) observes that one of the most common contemporary misrecognitions is underestimating the extent to which structural inequalities give only some individuals preferential access to practices that are socially recognized. Sayer argues that the cause of this unequal distribution of occupations—society's structural inequalities—is likely to be misrecognized as being the deserved product of effort and intelligence. Furthermore, specific individuals' contributions are evaluated according to the unequal distribution of labor, misrecognizing their contributions and qualities. The combination of these two misrecognitions means that regardless of individual traits and efforts, the economically privileged are seen as having earned their wealth through hard work and superior ability, and the economically disadvantaged are seen as deserving of their lack of wealth because they are lazy and incompetent. These misrecognitions hide and reinforce contemporary society's structures that created class inequality, contributive and distributional injustices, and their accompanying normative discrimination norms.

In normative discrimination, a negative preconception, not the other individual, is being seen. The other is being viewed through the negative preconception and treated with hostility on the basis of it. The mistake in normative discrimination is that the perpetrator is guided by his or her own preconceptions (though these preconceptions are usually learned from the culture's recognition norms) rather than the attributes actually possessed by the other individual. The perpetrator assumes, if not insists, that the oppressed others conform to those preconceptions, and the perpetrator is resistant to contrary information. Negative recognition norms are a denial, often with malice, of the positive values and contributions of others who hold particular traits and, thus, are misrecognition. Oppressed individuals are rendered without voice or will, and their experiences, words, and actions are suspected and delegitimized. Today, for example, Muslims are tracked (literally and figuratively) as terrorists; their every word and action is treated as suspect, and their claims for recognition as human beings are delegitimized.

In today's pluralistic society, malice in normative discrimination often reflects social insecurity by dominant groups against minority groups. Racists, sexists, homophobes, jingoists, and antireligious bigots of all stripes imagine themselves harmed by the social inclusion of hated and feared groups. To see those one thinks inferior being treated equally by society is perceived as a moral insult. Normative discrimination also arises in the midst of ethnic and sectarian conflicts. When tensions exist between social groups, all sides can become paranoid and overly sensitive to what the other groups are doing. Actions by the other groups are negatively tracked and perceived as threatening, and the success and well-being of other groups may be perceived as a matter of the others receiving greater and unfair advantages, thus diminishing one's own perceived social position (For example: Cohen, 1972; Vertovec, 2010; Azmanova, 2011; Göle, 2011).

Common targets of negative stereotyping are subcultures and countercultural movements, such as religious sects and youth movements. Stanley Cohen (1972) observed that the behavior of subcultures, such as the violence between mods and rockers in the United Kingdom in the 1960s, is exaggerated by the mainstream culture to hysterical proportions, generating

unwarranted hostility against those subcultures. Members of the subculture are stigmatized as moral outsiders or, as Cohen calls them, “folk devils,” who are defined as a threat to the mainstream social order, values, and interests. The perceived threat becomes a moral panic, rousing normative discrimination against the members of the subculture. Members of a subculture are labeled as deviants, and “once a person is thus type cast, his acts are interpreted in terms of the status to which he has been assigned” (Cohen, 1972, p. 12).

The morality of normative discrimination is easily compartmentalized by perpetrators. Those who deny equal rights to women, minorities, immigrants, or other groups often do not see themselves as being against rights and equality. They would see their exclusion of particular groups not as a double standard but as consistent with and upholding of moral norms. They would justify their disparate treatment with an interpretive narrative of why targeted groups are deserving of exclusion. Superficially rational arguments are used to justify the misrecognition as a case of the victims deserving it and even that there is an ethical demand to misrecognize these individuals because of their traits. The presence of normative discrimination reinforces an environment in which mistreatment of others is defensible. Instances of normative discrimination will be witnessed by other individuals who will become disinclined to offer recognition to socially ostracized individuals, even encouraged actively to misrecognize those individuals. Powerful individuals and institutions can use arguments and persuasion to convince others to engage in normative discrimination against targeted groups or individuals. Similarly, individuals can appeal to interpretive narratives to provide post hoc validity for misrecognition motivated by personal reasons. Often, rational arguments are not needed to tap into fear and hatred of others who are different.

2.2. Pathological recognition

Recognition norms that purport to recognize groups of individuals positively but in practice misrecognize them I label “pathological recognition”—recognition norms that cause injustice. In pathological recognition, social relations are structured by a recognition order that designates social groups, such

as women, as having particular traits that should be positively recognized. Within this recognition order, individuals recognize others by engaging with and applying their culture's seemingly positive recognition norms. This type of misrecognition behavior conforms to a recognition-like structure that socializes individuals into behaviors that emulate recognition, but the norms are pathological in that the cultural recognition norms are concealed misrecognitions, the application of which does not support others' self-realization and autonomy. Pathological recognition norms erode others' autonomy by subsuming them under recognition norms that define them and limit their possibility for recognition—for example, the traditional characterization of the traits and contributions of women. Unlike normative discrimination, the recognition norms of pathological recognition provide an affirmation of the value of targeted individuals. Both forms normatively restrict individuals and their possibilities, but pathological recognition deals with norms that focus on alleged positives that exclude other positives, whereas normative discrimination focuses on alleged negatives.

Pathological recognition encompasses portrayals of negative recognition as domination, as advanced by, for example, Althusser (1971), Markell (2003), and McNay (2008). These theorists address how recognition is used to maintain social domination by motivating subjects to serve the interests of power. Individuals are recognized for adhering to their responsibilities and duties to society, and their recognized compliance gives them a social identity. These theorists tend to reject recognition as irredeemable. Althusser rejects recognition as the central mechanism of ideology, and McNay rejects recognition as a model for emancipatory critique. Markell sees the pursuit of recognition of our identity from social institutions as unobtainable, contributing to injustice rather than emancipating us from it. What these negative conceptions of recognition tend to overlook is that the forces of domination succeed because they are exploiting a positive social mechanism. Recognition can be distorted and used to dominate people because recognition can have a positive influence on individuals but can be difficult for individuals to discern. As Honneth (2007b) says, we need to distinguish the false forms of recognition from its correct morally positive forms, even though identifying “correct and morally required” recognition is even more difficult than Honneth assumes.

What I argue separates false forms of recognition from positive ones is that the former attribute to individuals stereotypical traits and value judgments that subsume individuals under a group definition. These attributed recognition norms hinder those individuals' possibilities for self-realization and receiving recognition for their actual qualities and contributions. Because pathological recognition norms appear to be positive recognition but in practice perpetrate misrecognition, they are false dis-ease-causing forms of recognition. Pathological recognition's positive affirmations are deceptions that mislead individuals into accepting affirmations that limit them. Honneth gives the example of the idea of the heroic soldier, which grants to men who suffer social insignificance and a lack of prospects a type of recognition by becoming part of the military subculture (Honneth, 2007b). This example illustrates how pathological recognition works. In the military subculture, individuals gain a measure of prestige and honor while at the same time being treated as nonautonomous servants of the state, if not used as cannon fodder to achieve aims in which they have little or no involvement or from which they do not benefit. It is, at its core, a pathological recognition that lionizes war and honors "Our Glorious Dead" while downplaying the reality that they are, indeed, now dead. We should not doubt that many served honorably, and whether they acted for their country, families, or comrades in arms, they did their jobs properly and are worthy of our esteem. We also should not doubt that the esteem many individuals give to military veterans is sincere and with cause. Many of those who adopt the pathological recognition norms are not deliberately misrecognizing others but are following social norms, so they believe they are behaving properly. Individual soldiers accepted the pathological recognition that glorifies war because it offers them with a place to belong and a sense of purpose, even if it denied them other options for self-relation and social affirmation. Maybe they had no better options, because of their social position and society's contributive injustice, to achieve social status.

This is why pathological recognition cannot be reduced to ideological machinations of the power structure. We need to discern how pathological recognition relations become part of the social fabric of normative expectations. Pathological recognition works because there is reason for individuals to accept the narrative that distorts individuals and their possibilities, and with time these distortions become cultural traditions, largely unthought and unseen. The

theoretical approach to pathological recognition is best served by understanding it as a largely prereflective form of recognition whose normative conceptions of individuals appears to value their traits and status positively but has lost touch with its tangible effects on individuals. That the putative identities of pathological recognition give individuals a sense of meaning and value is why it is appealing to those being limited by it, despite its imposed limitations. Pathological recognition can exploit two positive aspects of recognition: its importance for individual development and relations-to-self and its normative role as a guide for proper conduct. Pathological recognition can convince people that they are affirming themselves by applying the pathological recognition norms, and in doing so they are affirmed by others as behaving properly.

Pathological recognition remains influential as long as the targeted recipients do not come to realize fully that they are not receiving fair and equal recognition. This is why pathological recognition is accompanied by narratives that justify and maintain its pathological norms. Social norms provide criteria for knowing what is expected of us and for assessing our own actions. We are socialized into the habit of relying on norms to guide our behavior. By distorting recognition norms to give the appearance that following the norms either affirms individuals and/or exemplifies proper conduct, the recognition order can influence people into misrecognition behavior that could advance a conformist ideological agenda. Critique of this ideological structure is necessary to reveal its assumptions and influences in order to open the possibility of individuals' awareness of the structure and motivate them to oppose misrecognition.

A prime example of pathological recognition is the traditional classification of women as caregivers—the normative evaluation of a “good” woman as wife and mother. The traditional social definition of womanhood tracks positive qualities, such as being caring and nurturing. Such a classification of women recognizes women, but in a limiting way. The norms assigning women caregiver roles can place women into restrictive gender roles that, among other effects, define women's care work as part of a woman's natural disposition and thus not real labor warranting compensation. The pathological recognition norms defining women's nature as being caring and nurturing beings assumes that women should pursue caregiving as their life's work—either as a wife and mother or in

caregiving professions such as nursing or teaching. The pathological norms defining women in this way preclude other possibilities for recognition—if one is recognized as being a caregiver, then one cannot easily also be recognized as powerful, creative, or intellectually gifted. Also, women’s caregiver role is recognized as a less valuable contribution compared with the contributions made by men. No matter how good a wife and mother a woman is, she still would not be esteemed highly compared with men. This pathological recognition does not empower women or engender their self-realization. In practice, it limits women’s autonomy and self-image by socializing them into accepting a seemingly positive self-image as a caregiver. What is recognized and honored excludes women as individuals with particular traits or behaviors and is instead an impersonal, stereotypical perception of women, attributed to women as a group. The side effect of this pathological recognition is that other possible roles for women that do not include the recognized traits are discouraged if not outright denied to women, and when women do enter male-dominated occupations, women are paid less.

That gender roles are propagated in terms of recognition helps explain their persistence. By objective criteria, women are being subordinated and harmed by this “recognition,” yet women and men find it difficult to overcome gender roles because those roles’ pathological recognition norms are embedded within society’s other, healthy recognition norms as part of the fabric of cultural attitudes. Men and women are socialized into a society’s set of recognition norms, and the pervasiveness of a culture’s pathological recognition norms gives individuals reason to believe that the norms are true and proper. Because men and women need mutual recognition and social acceptance, it is in their self-interest to adopt recognition norms concerning women in their behaviors. Thus, through recognition, women have a stake in their subordination. What is more, the normative content of recognition normalizes the expectation of the subordination of women. For men, their place as superior to women is affirmed. For women, their subordination feels natural in that it is familiar even when it feels wrong somehow. Pathological recognition norms affirm gender roles while at the same time hiding the damage they inflict.

Pathological recognition is the link between ideological structures and practices of subordination. Ideological forms of subordination cannot prevail without an adequate level of participation from all concerned. Amy Allen (2008, p. 77) says that “regulatory regimes cannot maintain and reproduce themselves; instead, they must be maintained and upheld by the individuals whom they regulate.” Power structures can achieve acceptance and attachment from individuals through recognition norms. Recognition norms provide a moral grammar that individuals can use to gauge their own and others’ behavior. Recognition norms of all types are authoritative and normalizing—those following them see them as the basis for judging themselves and others. Generally, individuals want to do what is considered proper and help maintain social order, and they are given reason to believe that by complying with traditional gender attitudes they are doing good for themselves and others. When socialized individuals adopt pathological recognition norms within their lives, those subordinating recognition norms persist and propagate through the generations.

The basic structures of normative discrimination and pathological recognition can be applied to any group. Both pathological recognition and normative discrimination encourage a lack of engagement with other individuals because those individuals are subsumed under group identities and general norms, which limits recognition relations and the possibilities for individuals. The classification of women as caregivers recognizes women who are quiet, nurturing, and long-suffering, but it does not honor, perhaps does not even see, women who are bold, intelligent, and creative. Pathological recognition and normative discrimination limit individuals’ expression of talents and ideas and often actively silences targeted groups. If, for example, one has a limited view of East Asians as excelling in math and science, one can easily fail to recognize that they have talents and interests in the arts and humanities. If one has a limited view of African-Americans as physically talented in entertainment and athletics, one can easily not recognize their cognitive talents in academia or leadership roles. These stereotypes limit possibilities for recognition relations and opportunities for the individuals being stereotyped, failing to recognize them as human individuals who possess unique talents and personalities. Such

stereotyping is common in human societies across race, ethnicity, gender, religion, and other group traits.

3. Dimension of Horizontal Engagement with Individuals

So far, we have discussed vertical misrecognition behaviors that are shaped by engagements with problematic norms and seen briefly how they discourage engagements with individuals. Now, we move to horizontal misrecognition behaviors that more directly stem from a disengagement from other individuals. The moral aspect of human interaction is that recognition is possible only at a level of engagement in which other individuals are seen as having a moral value. Intersubjective engagement means interaction with and awareness of other individuals as individual human beings with thoughts and desires who deserve moral consideration. Recognition relations need both engagement with norms and engagement with individuals to affirm individual self-realization and freedom.

Depending on how powerful the processes are that diminish a sense of moral responsibility, horizontal misrecognition could be structurally entrenched in interpersonal interactions or could be a by-product of a temporal forgetfulness from which an individual can recover. If it is the former, misrecognition behavior is entrenched and difficult for an individual to perceive, much less overcome. If it is the latter, then perhaps moral responsibility to others is only forgotten or obscured but not entirely abandoned. Not engaging with others intersubjectively precludes the possibility of adequately recognizing others, resulting in misrecognition such that we are oblivious and inconsiderate. These are behaviors in which we are no longer responsive to the other and we no longer recognize the other for who he or she is and how he or she is behaving. Within this type of misrecognition, the perpetrator would not see the other individual, resulting in blindness to the positive contributions and capacities of others, forestalling recognition. However, this is not to say that the individual who is not engaging intersubjectively is deliberately engaging in misrecognition behavior or is even aware of misrecognizing others. A perpetrator could be so engrossed in his or her

activities, even while conducting them in a moral way, that he or she loses awareness of dealing with another human being. We go about our activities aware that others are present but not recognizing them as individuals, and this misrecognition is part of the self-absorption of everyday life. Honneth gives the example of a tennis player who is so focused on winning that she forgets her opponent is her best friend. Her goal has become independent of the context in which it originated, and “any attentiveness for the cooperating partner vanishes completely” (Honneth, 2008, p. 155). The tennis player has not forgotten the rules of the game but has forgotten her opponent’s humanness and is no longer engaging intersubjectively with her friend. Winning the match has become a single purpose independent of her other relations to the world. Such a forgetfulness of others is commonplace in everyday life when we are caught up in everyday tasks and fail to notice and appreciate others as individuals like ourselves. In such forgetful behavior, we do not see other individuals as agents whose contributions and personal well-being should be taken into account. We perhaps remain polite, giving the appropriate gestures of civil behavior, but because we are forgetting the other, our courtesy is cursory and cold. This unintentional disengagement from others is temporary and does not necessarily lead to the elimination of all intersubjective engagements.

This dimension of misrecognition could be a deliberate withdrawal from intersubjective engagements or an involuntary loss of the capacity for intersubjective engagements. Deliberate withdrawal from intersubjective relations is not a forgetting but a denial or defensiveness resulting in not considering the possibility of others’ contributions and personal well-being. A deliberate withdrawal is most likely isolated within certain social interactions rather than across an individual’s entire life—for example, an individual being inconsiderate of others while posting comments online—engaging in thoughtless or aggressive behaviors he or she would not engage in within other interpersonal interactions. Disengagement may also result from general insecurity and anxiety. We are, as Cillian McBride (2013, pp. 136-137) says, recognition-sensitive beings because our well-being depends on receiving recognition, as Honneth says. We need recognition from others, and the potential lack of it is a threat. It is no surprise that social life comes with anxiety about how others will judge us and our actions. Such anxiety limits how much we are willing to risk exposure to negative

judgments from others, and we deliberately hold back from engaging with others. Tension in recognition relations explains a great deal of the social anxiety we all have, and it explains why some individuals protect a positive self-image by telling themselves they do not need others' approval, which is another way intersubjective engagement can be cut off. Within our relationships, changes in circumstances or in other individuals change our recognition relations with them. We are creatures of habit, and we tend to resist change. The threat of changes in relations leads to the temptation to withdraw from intersubjective engagement.

In objectification, one engages with specific other individuals but nonetheless takes them to be the same as others who seem to possess similar traits and capacities. This generalization neglects the distinct traits and needs of the individual, objectifying or commoditizing him or her. In more benign forms, objectification is a general depersonalization, akin to the temporary lack of intersubjective engagement, in which only a general recognition of that individual is possible. Because we are not engaging with that person as a specific individual, we are less open to recognition relations beyond acknowledging him or her as a human being. The other individual is perceived as only a member of a type and is objectified or commoditized. Objectification is often malevolent. Normative discrimination against groups leads to objectification of individuals whose particular traits and capacities are erased by an identity attributed to them that opens up the objectified to abuse. Within patriarchy, women are objectified as beings in service of male desires. Sexual objectification of women has multiple forms and degrees, but the common denominator is that women are the proper objects of male sexual desire, the proper purpose of women being that they are used in order to satisfy that desire. As Timo Jütten (2016) has argued, sexual objectification of women is more than instrumentalisation of them; it is a social meaning imposed on them that undermines their autonomy and equal social standing even at times when they are not being used as an instrument to gratify male desire.

To objectify another individual is to see that person as not worth engaging with intersubjectively. We know that we have more latitude in how we treat objects than in how we treat other human beings, and we can use objects as we see fit. When individuals are objectified, they are objects for our use. In sexual

objectification, the woman is excluded from intersubjective engagement because her normative status is as a sexual object, and not even a particular sexual object, but a fungible one. She is a commodity to be bought, sold, and used. This, I argue, is because objectification, unlike normative discrimination, sees the victim as having use value but not as an individual. The companion to pathological recognition's putative positives of women's gender roles is that women are portrayed as subordinate objects whose value is lessened by intersubjective engagement, plus, intersubjective engagement reveals the woman is not an object. This normative status of women as sexual objects not to be engaged with as individuals enables human trafficking and sexual slavery. Attempts by women to exercise their autonomy in opposition to their imposed social meaning are met with stern and perhaps violent attempts to reassert their objectification.

Self-absorbed disengagement involves behaviors in which the disengagement is driven not by norms but either by a refusal to acknowledge or deal with another or by a moral judgment that the other individual is not worth engaging with as an equal. Perhaps the other is considered to have no qualities of value to the community or not having wants or needs worth considering. What interaction there is with others assumes how other individuals are, and one is resistant to contrary information. Self-absorbed disengagement is misrecognition in its denial of the other individual's value. Various kinds of social ostracizing, shunning, and disregard are examples of self-absorbed disengagement.

Instances of this misrecognition behavior involve restricting the granting of recognition to a select few, while disregarding or denying recognition to others. A line is drawn between those who are "us" and those who are "them," and we engage with "us" but not with "them." Honneth (2007a, p. 227) gives the example of a robber recognizing his companions while misrecognizing his victim. Similar divisions are drawn by almost all individuals who, wittingly or not, engage more with individuals within their group but less with those outside their group. This explains the tendency that all individuals have of according more recognition to those within their social circle than to those outside it. We each have our own lives and our own involvements and relationships, and it is no surprise that we tend to

place more importance on individuals and relationships close to us. The human tendency to view those outside our own group through a lens of our own preconceptions is a common form of misrecognition. This disengagement happens at all levels of society, from ignoring other individuals who are not our friends or family, to governments not seeing members of groups for who they are. Unlike normative discrimination and pathological recognition, this type of misrecognition is not driven by social norms as much as by individuals' decisions of inclusion and exclusion. It does not have the entrenched hostility and dehumanization of normative discrimination, but there is a lack of genuine respect and consideration of those outside one's social circle.

Within self-absorbed disengagement, the engagement is with the preconceived notions of the situation and others and not how the situation and other individuals actually are. This self-absorbed behavior hinders communication and intersubjective recognition. McBride describes an occurrence of self-absorbed disengagement by the British government in 1931: "The Indian delegates had been organized into *religious* groups by the colonial power. Gandhi objected vehemently but colonial officials were immovable in their determination to view Indians primarily through the lens of sectarian division" (McBride, 2013, p. 37). The British self-absorbed disengagement that saw Indians only in terms of British categorization had its most disastrous consequences in the shortsighted partition of the land into Hindu India and Muslim Pakistan. McBride quotes Patricia Williams on the attitudes of white tourists to local African-American churches, saying that for the whites, "no one existed for them who could not be governed by their intensions" (as cited in McBride, 2013, p. 37). In other words, for the white tourists, African-Americans did not exist as individuals, only as objects to be used for their pleasure. These condescending and patronizing attitudes are not seen by perpetrators as disrespectful, and it might not even occur to them that they are misrecognitions. We could count some of this as the manifestation of privilege, but in a broader sense it reveals a common human laziness to engage with others and be open to perceiving them as they are.

So far, I have discussed the dimension of personal engagement only in its positive aspect—that because acknowledging another individual’s positive traits is recognition, the absence of this acknowledgment is misrecognition. There is also personal engagement that is perverse. Intersubjective engagement is perverse when it is perpetrated with intentions contrary to affirming the other positively. In perverse engagement, in regard to the individual engaged with, recognition norms are willfully neglected because the needs and desires of the other are subsumed under the perpetrator’s desires. Unlike normative discrimination that targets a social group, in perverse engagement the perpetrator targets a specific individual, believing either that this specific individual does not deserve to be treated well or that this individual’s deservedness is unimportant in the context of the perpetrator’s larger concerns. Perverse engagement ranges from selfishness to active manipulation of another to sadistic behavior. Examples of perverse engagement would be a bully who targets a specific individual to abuse or a boss who harasses a particular employee. Most bullying and harassment are targeted antagonism that can be understood as perverse engagement. Another example is a con artist engaging with an individual to swindle him or her. The con artist may have selected a target on the basis of the target’s perceived vulnerability or gullibility, and the con artist ignores the norms against theft and dishonesty, specifically in terms of that individual, even if the con artist is honest with others. The con artist is engaging with the other, recognizing and acting in response to the individual’s qualities, only in the service of his or her involvement in the successful swindle, but there is not necessarily a specific antagonism toward the target.

The complexity of the negative aspect of personal engagement can be unpacked further. The misrecognition behavior of perverse individual engagement is characterized by the self-absorbed perpetrator focusing on a personal relation but not on recognition or moral norms. In other words, the character of perverse individual engagement is not “these norms are irrelevant” but “for this specific individual these norms are irrelevant.” A stalker is obsessed with a specific individual to the extent that norms of appropriate conduct are subsumed under the stalker’s desires, which disregards how the other individual actually is. Stalkers often falsely believe the objects of their obsession return their interest and do not take in information to the contrary. If the stalker was engaging

intersubjectively with the other individual, he or she would take into account the expressions of noninterest from the other person. But because the stalker is engaging not with the other but with his or her own attributed identity of the other, there is no intersubjective engagement. An individual seeking revenge on another is engaged with that individual but seeks to harm him or her, so the engagement is not an intersubjective genuine respect. The perpetrator considers that the other deserves to be harmed, not affirmed.

4. Conclusion

The multidimensional view of misrecognition expands the tools available to us to investigate and understand injustice. The misrecognitions perpetrated by and experienced by individuals do not map onto Honneth's typology of misrecognition as being the contrary of recognition—a typology that is too narrowly construed to encompass the diversity of misrecognition behaviors. Rather than seeing misrecognition as simply a violation of recognition norms, the multidimensional view separates positive and negative recognitions. The multidimensional view of misrecognition understands recognition and misrecognition as complex responses by individuals to everyday circumstances that involve varying vertical engagements with norms and horizontal engagements with other individuals. This more complex view of misrecognition reflects the complexity and diversity of human behaviors and helps us to better understand the distinct harms of injustice. The multidimensional view illuminates that, though social institutions set the stage, misrecognition behaviors occur at the microsocial level of individual and small group interactions. The insight that recognition and misrecognition are best viewed in terms of a web of interpersonal relations opens up a microsocial analysis of recognition relations and provides us with the tools to begin to tackle those injustices caused by individuals not being treated *as* individuals. This microsocial analysis adds complexity to our task of understanding misrecognition, because it calls for a deeper analysis of individual relations, attitudes, and actions, but such an analysis will give our social theory greater explanatory power.

References

Allen, A. (2008). *The Politics of Our Selves: Power, Autonomy, and Gender in Contemporary Critical Theory*. New York, NY: Columbia University Press.

Althusser, L. (1971). Ideology and Ideological State Apparatuses. In *Lenin and Philosophy and Other Essays*, (B. Brewster, Trans.). New York, NY: Monthly Press Review.

Azmanova, A. (2011). Against the Politics of Fear: On Deliberation, Inclusion, and the Political Economy of Trust. *Philosophy & Social Criticism*, 27(4), 401–412.

Bader, V. (2007). Misrecognition, Power, and Democracy. In Bert van den Brink & David Owen (Eds.), *Recognition and Power: Axel Honneth and the Tradition of Critical Social Theory* (pp. 244–246). New York, NY: Cambridge University Press.

Cavell, S. (1976). *Must We Mean What We Say?* Cambridge, UK: Cambridge University Press.

Cohen, S. (1972). *Folk Devils and Moral Panics: The Creation of the Mods and Rockers*. London, UK: Grenada.

Darwall, S. (2010). The Second-Person Standpoint. *Philosophy and Phenomenological Research*, Vol. LXXXI Nr. 1.

Fanon, F. (1967). *Black Skin, White Masks*. (C. L. Markmann, Trans.). New York, NY: Grove Press.

Göle, N. (2011). The Public Visibility of Islam and European Politics of Resentment: The Minarets–Mosques Debate. *Philosophy & Social Criticism*, 27(4), 383–392.

Honneth, A. (2003). *The Struggle for Recognition: The Moral Grammar of Social Conflicts*. London, UK: Polity.

Honneth, A. (2007a). *Disrespect*. Cambridge, UK: Polity.

Honneth, A. (2007b). Recognition as Ideology. In Bert van den Brink & David Owen (Eds.), *Recognition and Power: Axel Honneth and the Tradition of*

Critical Social Theory, (pp. 323–347). Cambridge, UK: Cambridge University Press.

Honneth, A. (2008). *Reification: A New Look at an Old Idea*. Oxford, UK: Oxford University Press.

Honneth, A. (2011). Rejoinder. In Danielle Petherbridge (Ed.), *Axel Honneth: Critical Essays*, (pp. 391–421). Leiden, Netherlands: Brill.

Ikäheimo, H. (2015). Conceptualizing Causes for Lack of Recognition: Capacities, Costs and Understanding. *Studies in Social & Political Thought*, 25, <http://journals.sussex.ac.uk/index.php/sspt/article/view/45>.

Jütten, T. (2016). Sexual Objectification. *Ethics*, 127, 27–49.

Markell, P. (2003). *Bound by Recognition*. Princeton, NJ: Princeton University Press.

McBride, C. (2013). *Recognition*. Cambridge, UK: Polity.

McNay, L. (2008). *Against Recognition*. Cambridge, UK: Polity.

Patterson, O. (1982). *Slavery and Social Death: A Comparative Study*. Cambridge, MA: Harvard University Press.

Renault, E. (2011). The Theory of Recognition and Critique of Institutions. In D. Petherbridge (Ed.), *Axel Honneth: Critical Essays*. (pp. 207–232). Leiden, Netherlands: Brill.

Sayer, A. (2009). Contributive Justice and Meaningful Work, *Res Publica*, 15(1), 1–16.

Sayer, A. (2011). Misrecognition: The Unequal Division of Labour and Contributive Justice. In M. Yar & S. Thompson (Eds.), *Politics of Misrecognition* (pp. 87–103). Farnham, UK: Ashgate.

Vetrovec, S. (Ed.). (2010). *The Multiculturalism Backlash*. Abingdon, UK: Routledge.

Zurn, C. (2015). *Axel Honneth: A Critical Theory of the Social*. London, UK: Polity.



COMBINING DEMOCRATIC EQUALITY AND LUCK EGALITARIANISM

ASSOCIANDO A IGUALDADE DEMOCRÁTICA E O IGUALITARISMO DA SORTE

Lars Lindblom*
lars.lindblom@umu.se

Abstract. *The concept of responsibility plays a crucial part in the debate between proponents of democratic equality, like Rawls, and defenders of luck egalitarianism, such as Dworkin. In this paper it is argued that the two theories can be combined, and that they should be combined to achieve a theory of justice that puts personal responsibility in its proper place. The concept of justice requires two different conceptions. The two theories can be combined because they deal with different problems of justice. They ought to be combined because, first, luck egalitarianism needs a theory of background justice, and second, a theory of justice must supply an answer to the question of just individual allocations, something that is not provided by democratic equality. Democratic equality and luck egalitarianism solve each other's problems. The combined theory will lead to allocations of goods that respect both the difference principle and the envy test.*

Keywords: *Allocative Justice, Democratic Equality, Luck Egalitarianism, Responsibility, Social Justice.*

Sumário. *O conceito de responsabilidade desempenha um papel crucial no debate entre os defensores da igualdade democrática, como Rawls, e os defensores do igualitarismo da sorte, como Dworkin. Neste artigo, argumenta-se que as duas teorias podem ser associadas para alcançar uma teoria da justiça que coloca a responsabilidade pessoal no seu devido lugar. O conceito de justiça requer duas concepções diferentes. As duas teorias podem ser associadas por lidarem com diferentes problemas sobre a noção de justiça. Elas devem ser associadas porque, em primeiro lugar, o igualitarismo da sorte precisa de uma teoria de fundo da justiça, e em segundo lugar, a teoria da justiça deve fornecer uma resposta à questão da distribuição individual justa, algo que não é esclarecido pela igualdade democrática. Cada uma das duas teorias, a igualdade democrática e o igualitarismo da sorte, resolve os problemas da outra. A teoria associada levará a uma distribuição de bens que respeita o princípio de diferença e o teste da inveja.*

Palavras-chave: *Igualdade democrática, Igualitarismo da sorte, Justiça alocativa, Justiça Social, Responsabilidade.*

* Lecturer, Department of Historical, Religious and Philosophical Studies, Umeå University, Humanisthuset, 100 44 Umeå, Sweden.

o. Introduction

The notion of responsibility plays a central role in liberal egalitarian political philosophy. The debate between proponents of democratic equality and defenders of luck egalitarianism concerns the appropriate way to hold people responsible for their choices in terms of the theory of justice. The luck egalitarian view is that justice is sensitive to people's choices and that responsibility is therefore essential to the content of the theory of justice, whereas proponents of democratic equality claim that responsibility should not play a role in defining justice. In order to decide which theory best captures the notion of justice, it seems that we need to know where to make the cut between circumstances for which people are and are not responsible.

In this paper, I will argue that we should make not one but two such cuts, and hence that we need two interrelated conceptions of the concept of justice. The theories appear to contradict each other with regards to responsibility, but I will attempt to show that we can and should combine a version of democratic equality and a version luck egalitarianism into a combined view of justice. I will first demonstrate that Rawls's democratic equality view¹ and Dworkin's luck egalitarianism² can be combined, as they deal with two different questions of justice, at least to the extent that the core conceptions, such as their accounts of responsibility and distributive principles, of the two theories are retained. Rawls's theory is a theory of background justice, and I shall argue that we should interpret Dworkin's luck egalitarianism as a theory of justice in terms of individual allocations. Within this interpretation, the theories answer different questions about justice. The two theories should be combined because we want answers to both questions and because neither theory can answer both questions on its own.

¹ The obvious place to start is John Rawls (1971), but see also Rawls (1996) and (2001). Elizabeth S. Anderson defends democratic equality against luck egalitarianism in (1999).

² The term "choice egalitarian" probably better captures Dworkin's intentions, but the "luck egalitarian" label has stuck, so I will use it throughout the article. The canonical statement of Dworkin's theory of equality of resources was published in (1981), which was also reprinted along with other papers on equality in (2000). Other important papers that started this tradition are the classics by G.A. Cohen (1989) and Richard Arneson (1989).

This article does not aim to show that both sides of the debate on justice of the last thirty years have been mistaken. Quite the opposite. It aims to develop a position that allows us to say that both sides have been right.

Rawls's difference principle supplies us with a macro outline of the distribution of income in a just society, but it does not say which person should end up at which point in this distribution. But as Dworkin has pointed out, who gets what is a question we want to have answered when we are thinking about justice. On the combined view, holding a person fully responsible for bad option luck in a society in which background justice is in place is to say that it would be just if he or she ends up among the least advantaged in a society that is designed so that the least advantaged are as well off as possible. This is letting responsibility play a role. A further reason for combining the two theories is that if luck egalitarianism is applied against a Rawlsian background, then it is not susceptible to the counterarguments from callousness that have been put forward by Elizabeth Anderson.

This article consists of seven parts, the first of which provides some background on the concept of justice and the value of responsibility, and a discussion on the major role of the latter value in the quest for an appropriate conception of justice in political philosophy. Here I will also introduce the distinction between social and allocative justice, which will help us in combining the two theories. The second section focuses on Rawls's strict egalitarianism³ and on his views on justice and responsibility. The third section presents Dworkin's criticism of Rawls's theory of justice as fairness and outlines Dworkin's theory of equality of resources. The fourth section is devoted to Anderson's criticism of that theory. The fifth section starts by identifying an impasse. Democratic equality does not include responsibility factors in its criterion of justice, which seems to have the consequence that individual choice plays too little a role, whereas including that value, like luck egalitarianism does, leads to the problems of callousness. I suggest that we could combine Rawls's strict egalitarianism with Dworkin's luck egalitarianism. The sixth section, argues that the two theories should be combined, since they solve each other's problems. The outlines of such

³ I am using the phrase strict egalitarianism in a technical sense that will be explained in the next section.

a combined theory are sketched and some appealing features of such a theory are highlighted. The article ends with some concluding remarks.

1. Social Preferences, the Accountability Principle, and the Concept of Justice

Let us start with H.L.A. Hart's (1994, 156-167) distinction between concepts and conceptions. The concept of justice is our pre-theoretical and everyday notion of how goods ought to be distributed. A conception of justice is a theoretical and normative attempt to spell out a coherent and reasonable specification of that concept.⁴ This section is devoted to laying the groundwork for the investigation of justice from such a perspective. We will do this with the aid of behavioural economics, in order to get a hold on the concept of justice by means of an empirically grounded and tested approach.

When economists started testing game theory empirically, some surprising results emerged, which will help us to identify the outlines of the concept of justice. In the "ultimatum game", one party proposes a split of a resource and the other party either rejects the proposal and neither party gets anything, or accepts and the resource is split according to the proposal. In this game, proposers tend to offer about 40% of the total pie, and respondents commonly reject offers below 20%. In the "dictator game", the respondent does not have the option to reject a proposal. However, the proposers still offer, on average, about 20%. Standard economic rationality would predict that proposers would offer the lowest possible amount above zero in the first game and nothing in the second. This would guarantee that a rational responder would accept the offer in the ultimatum game, and, of course, in the dictator game a rational player would not consider the other player at all when deciding what split to propose (Camerer 2003). There is a fairly clear pattern in the results, which becomes only becomes clearer the less self interest there is at stake in a distributive choice, and this has led

⁴ John Rawls uses this distinction in (1971), on page 5, when he is explaining the role of justice, and Ronald Dworkin makes use of it in (1977, 134-136).

researchers to suspect that there is a structure to our preferences that could explain them. Since the pioneering study of Yaari and Bar-Hillel (1984) in the early eighties, numerous attempts have been made to find this structure in our preferences.⁵

One especially promising option is the accountability interpretation of people's social preference for equitable division. The accountability principle says that an individual's entitlement varies in direct proportion to the value of the individual's relevant discretionary variables, but does not hold the individual accountable for the values of exogenous variables (Konow 1996). A discretionary variable in this model is one that is under the control of the individual; often this has to do with choice. An exogenous variable is one that is not under the control of the individual. The principle implies that individuals should get equal shares if they are equally responsible for the outcomes, and that any fair step away from equality must depend on differences in responsibility exercised by the parties. In empirical tests that include production at the stage prior to the split, this model does exceptionally well.⁶ Social science may have gotten a handle on the concept of justice.

Not only does this research accurately predict people's distributive choices, but it also provides a framework in which the most central views on liberal justice can be systematized. Cappelen *et al.* have proposed such a framework that captures both the central views on justice in political philosophy and the concept of justice as described by the accountability principle. At first glance, the accountability principle might seem like a way of describing luck egalitarianism, but the idea is more general. Strict egalitarians say that there are no relevant discretionary variables, which implies that the distribendum should be distributed equally. Libertarians believe that all variables are to be considered as discretionary, coercion aside, which implies that there is no room for equality and that individual choices should decide the distribution entirely. There are also two intermediate positions. Choice egalitarians—in other words, luck egalitarians—

⁵ Among the more influential papers that have attempted explanations in terms of justice are the following studies by Gary E. Bolton and Axel Ockenfels (2000), Ernest Fehr and Klaus M. Schmidt (1999), James Konow (1996), and Gary Charness and Matthew Rabin (2002).

⁶ See Konow (2000) for an R^2 of 0.983.

believe that people should only be held responsible for their choices. Meritocrats also hold people responsible for their choices, but they also think that we can be held responsible for our personal characteristics, such as our innate talents. These theories agree on the general outline of justice, but they make different responsibility cuts (Cappelen *et al.* 2010). The research concerning people's actual fairness ideals shows that people tend to spread out over the four positions (Cappelen *et al.* 2007). This line of research gives us one way of pinning down the concept of justice, but if we are to find the proper conception of justice, we must engage in normative debate about for what, if anything, we should hold people responsible.

Since the standard views of justice in political philosophy can be categorized in terms of their approach to responsibility, it might seem that if we could locate *the* correct responsibility cut, we would be able to hone in on a solution to the problem of justice. We could then rather easily move from the concept to the (best) conception of justice. However, finding the right conception of justice is more complicated than it might seem. The answer to the question of what is the best conceptualization of a moral concept depends on the specific moral problem with which we are concerned. There are at least two different types of justice problems (Rawls 2001, 50). On the one hand, *allocative justice* concerns the way in which we should divide (or, rather, allocate) a given bundle of goods among a given set of individuals, knowing their preferences, needs, and desires. Here we ask, who should get what? This is the typical problem in economic analysis, and most work in experimental economics focuses on questions of allocative justice. On the other hand, debate in political philosophy often concerns the *structure* of the institutional framework in a just society. This I shall call the question of *social justice*.⁷ David Miller explains it as the moral problem that appears when we are within circumstances that satisfy three conditions: first, there is a group of people connected in a framework that distributes benefits and costs; second, there are institutions that are responsible for this distribution so that we can apply principles of justice to them; third, there are ways of changing this institutional structure in accordance with such

⁷ Rawls (2001, 50) calls this distributive justice, but this terminology seems slightly misleading, since one may say, without misusing ordinary language, that allocative justice also deals with questions of justice in distribution.

principles (Miller 1999, 4-6). Principles of social justice, in this sense, deals with institutional design, and not with the division of bundles of goods. As we shall see below, theories of democratic equality deals with the design of just institution, whereas luck egalitarianism can be understood as concerned with allocative justice.

It makes a world of difference if we are concerned with justice in single economic transaction, as in allocative justice, or with the institutional design of a just society, as in social justice. Even if we could solve the problem of justice in the wider case of social justice, it does not necessarily follow that we would also have a solution appropriate for the narrower case of piecemeal transactions. Each problem must be investigated on its own terms, and might need its own conception of the general concept of justice. Furthermore, there may be interrelationships between the two levels, which should be taken into account when designing the different conceptions. The issue of personal responsibility plays a major part in exemplifying these points in the remainder of this paper. The intuition that I will investigate is if the concept of justice needs two different conceptions for the problems of allocative and social justice.

2. Democratic Equality

John Rawls's theory of democratic equality is an attempt to answer this question:

[h]ow are the institutions of the basic structure to be regulated as one unified scheme of institutions so that a fair, efficient, and productive system of social cooperation can be maintained over time, from one generation to the next? (2001, 50).

The goal is to set up the institutions of the basic structure—the main political and social institutions of society—so that we can say that the distribution of rights, opportunities, and resources is fair, regardless of exactly who gets what, or, to put it another way, regardless of the resulting allocation between actual persons. This is pure procedural background justice, and a clear-cut example of a theory of social justice.

The idea is that a theory of justice should start with institutional background justice, because the basic structure plays such an important role in how people's

lives turn out and in what choices they make. Outcomes for people in real life do not depend solely on the choices that they make. Both cultural backgrounds and innate capabilities play roles in determining the central responsibility variables of choice and talent. What part should these be allowed to play in the conception of social justice? Rawls's answer is clear: none. "No one deserves his greater natural capacity nor merits a more favourable starting place in society" (Rawls 1971, 102). In other words, responsibility for these factors has no place in the design of a just basic structure, and, hence, not in our conception of justice. In Cappelen *et al's* terminology, Rawls is a strict egalitarian, because he holds that there are only exogenous variables at stake in justice. Responsibility comes into play for Rawls *after* we have set up the basic structure justly. Background justice concerns the fair background conditions for responsible choice. The idea is that when background justice is achieved, "[i]t is left to the citizens as free and equal persons, secure in their rights and liberties and able to take charge of their own life, to avail themselves of the opportunities guaranteed to all on a fair basis" (2001, 171).

How, then, do Rawls's two principles of justice deal with the impact of culture and innate talent? The second of his two famous principles consists of two sub-parts: fair equality of opportunity and the difference principle. Fair equality of opportunity says that not only should positions in society be open to all, but all should also have a fair chance of attaining them. "The expectations [of culture and achievement] of those with the same abilities and aspirations should not be affected by their social class" (1971, 73). When this principle is implemented, individuals' choices are no longer determined by the contingencies of the socio-cultural background. This creates the starting points from which individuals can be held responsible, since one very important reason for not holding people responsible is removed. However, we may assume that even if aspirations are functions of culture, we could not reach all the way to equality of aspiration as long as the institution of the family remains (possibly forever), and we probably have at least some abilities that are innate. These facts are part of the reason that we also need the difference principle. If these contingencies of nature cannot or should not be eradicated, there is another solution. We could try to arrange the basic structure so that these contingencies work for the benefit of those least advantaged (Rawls 1971, 63-65). When this is the case, those who gain more,

presumably those with more innate talent, do so in a way that is to the most advantage for the least talented.

The difference principle is generally taken to include the distribution of income and wealth only, but correctly understood, says Rawls, it also implies that there should be a safety net, including medical care for all. The difference principle underwrites equality of opportunity and equal citizenship; in order to function as an equal citizen one needs health care at various stages in life, and therefore the state is obliged to provide such care. This means that the difference principle implies a social minimum (2001, 173-176). When the two principles are in effect, there is a place for the notions of responsibility and therefore desert, because choice is not determined by factors irrelevant to justice. But Rawls points out that “this sense of desert presupposes the existence of the cooperative scheme; it is irrelevant to the question whether in the first place the scheme is to be designed in accordance with the difference principle or some other criterion” (1971, 103). Responsibility ascriptions have their place *within* the basic structure.

3. Luck Egalitarianism

It is sometimes said that Ronald Dworkin revived the egalitarian cause by incorporating the idea of responsibility directly into the theory of equality (Cf. Cohen 1989, 933). In order to see how this was done, we now turn to Dworkin’s theory of equality of resources. This theory is the standard example of luck egalitarianism. A good way to start is to look into the faults Dworkin finds with Rawls’s conception of justice. The first basic problem is that it is not fine-tuned enough. The difference principle works to improve the lot of the least advantaged group in society. Equality, however, says Dworkin “is in principle a matter of individual right rather than group disposition” (1981, 340). Equality of resources is an attempt “to provide a description [of] equality of resources person by person, and the considerations of each person’s history that affect what he should have, in the name of equality, do not include his membership in any economic class” (1981, 340). The second problem concerns responsibility. Rawls fails to take seriously individual history and choice. By invoking the difference principle, Rawls mistakenly “supposes that flat equality in primary goods, without regard

to differences in ambition, taste, and occupation, or to differences in consumption, let alone differences in physical condition or handicap, is basic or true equality” (1981, 343).

In other words, egalitarian justice is achieved when individual entitlements, where appropriate, are based on individual choice. Where Rawls famously starts with a contract, Dworkin starts with an auction. The distribution of goods in a society is just when it mimics the outcome of this auction-based thought experiment. A group of shipwrecked people comes ashore on a deserted island full of resources. They all agree that no one has a prior claim to the resources, *i.e.*, there are no relevant discretionary variables. They decide to hold a Walrasian auction, where each gets an equal amount of markers to use in bidding for the resources he or she individually prefers. They apply the so-called envy test to ensure that equality is preserved. This idea says that a justified division is achieved when no one prefers anybody else’s bundle of goods. The test is satisfied through the auction, because if a person envies someone else’s bundle, he or she is free to bid for it. In this way, every person bears the cost of his or her choice of lifestyle, while equality is preserved. Every immigrant is responsible for the use he or she makes of the markers he or she has been dealt at the start of the auction. The measure of equality of resources is then the allocation that would result from an auction in which all participating parties have equal resources at the start. This is how justice can be individualized.

However, after some time on the island, some people will have worked more than others, some will have fallen sick, others will have been lucky in business, and, of course, accidents will happen. When such things have happened, the envy test will no longer be satisfied. It is here that Dworkin’s famous distinction between brute and option luck comes into play. Option luck has to do with how deliberate gambles turn out, for instance, if a person decides to play the stock market and loses his or her money—or, for that matter, grows very wealthy. Brute luck concerns plain bad unforeseeable luck, *e.g.*, being hit by a meteorite, or in the case of brute good luck, finding a lost treasure. Insurance provides a bridge between brute and option luck. A person who prefers to play it safe can buy insurance, while risk-takers can choose to go without a safety net. If insurance is available, then brute luck is converted into option luck. Thus, the envy test can

still be satisfied. In setting up the auction, we “have already decided that people should pay the price of the life they have decided to lead...[s]o we have no reason to object...to a result in which those who decline to gamble have less than some of those who do not” (1981, 294). We should neither object to the fact that some people gamble and win nor to outcomes where they have gambled and lost. Redistribution from winners to losers is ruled out, because this would mean that the winner would be made to bear the cost of the life choices of the losers.

Dworkin makes one further distinction to clarify for what people are responsible. The goal of equality of resources is that holdings should be ambition-sensitive, but not endowment-sensitive. The allocation of resources must be sensitive to the choices that people make, with regard to, for example, savings, consumption, and production, but it must not be sensitive to exogenous variables, such as talent, handicaps or brute luck. In order to spell out this notion, Dworkin develops a thought experiment where the parties choose insurance packages against such exogenous bad luck. Justice, then, demands that the distribution of resources mimics the outcome of this hypothetical insurance market.⁸ In short, a distribution of goods is just if it could have been achieved from a starting point of equal shares by trade, holding people responsible for ambition and option luck while compensating for bad brute luck and equalizing the result of unequal endowments. In conclusion, Dworkin’s theory of justice answers questions like who should get what; it is a conception of justice that deals with allocative justice. It solves the two problems that Dworkin finds with Rawls’s theory by incorporating the value of responsibility in the conception of justice.

4. Anderson and the Problems of Luck Egalitarianism

We turn now to perhaps the most important critic of luck egalitarianism, Elizabeth Anderson (1999). She finds many faults with this kind of theory, but here we will concentrate on four aspects of it that bring the question of responsibility most clearly into focus. The first is what Anderson calls *the*

⁸ Dworkin discusses this hypothetical insurance scheme to handle the issue of handicaps extensively. However, in what follows it will play a minor role, because John Roemer has conclusively shown that Dworkin’s thought experiment does not produce the conclusions for which he is looking; see Roemer (1996, chapter 7). The parties behind this veil of ignorance would allocate more resources to those who are most able to use them efficiently in the pursuit of utility. In other words, the healthy and vigorous would get the money.

abandonment of negligent victims. If only bad brute luck should be compensated, it follows that the results of bad option luck are of no concern to justice. This implies that it would be just that an ambulance passed by a dying person who has made a deliberate decision not to buy insurance. The second, and similar, problem can be called *the abandonment of the prudent*. A person who is struck by several incidents of bad option luck may end up in a situation in which he or she is unable to afford insurance. It might be more prudent to feed one's children than to pay the insurance bill. Such a person can have made all of the reasonable choices and still end up uninsured. If this person finds himself or herself in the traffic accident above, it is just, according to the luck egalitarian view of justice that the ambulance does not stop to help him or her. Both of these problems are related to the third problem: *the lack of a safety net*. There is something counterintuitive about a theory of justice in which, even if only in principle, the fact that some people have absolutely nothing is considered just. There is in principle no limit to how low one can fall in a society governed by such principles. The final problem we will discuss goes under the heading of *the vulnerability of dependent caretakers*. This has to do with the traditional way to rear children: father works and mother tends the children. The difficulty for luck egalitarianism is that women, being no less talented than men but choosing to stay home to care for their children, will not make much money for themselves. Apparently, they make a lifestyle choice of being dependent on their spouses, and this dependency comes out as entirely just. For these reasons, Anderson presents the following indictment of Dworkin's view of justice:

The fact that these evils are the product of voluntary choices hardly justifies them: free choice within a set of options does not justify the set of options itself. In focussing on correcting the supposed injustices of nature, luck egalitarians have forgotten that the primary subject of justice is the institutional arrangements that generate people's opportunities over time (1999, 308-309).

Luck egalitarianism seems to have some severe drawbacks when conceived as a conception of social justice. How does Rawls's theory of democratic equality handle these four problems?⁹ Let us start with the vulnerability of dependent caretakers. From a Rawlsian perspective, this is first and foremost a question of

⁹ There are clear similarities between the views of Rawls and Anderson, but Anderson prefers Sen's notion of capabilities to primary social goods and also seems to have more republican leanings than Rawls. Therefore, it makes sense to ask this question about Rawls's view, even though both writers position themselves in the democratic-equality camp.

fair equality of opportunity. Life choices should be made against a background where positions are open to all, and everybody should have a fair chance of attaining those positions. That this is not the case becomes obvious when we look at the fact that it is predominantly women who are caretakers in any society we care to investigate. If we are to take fair equality of opportunity seriously, we should strive to create a society in which the choice of becoming a caretaker does not depend on cultural preconceptions of gender. If we could say, which we cannot, under present circumstances, that it is a solely a personal choice to become a homemaker, then we could also consider this problem solved. However, doing so would miss an important part of the complaint, namely that women become economically dependent on men. We can deal with that issue while discussing the abandonment problems and the lack of a safety net.

As we have seen above, the difference principle implies a social minimum, including medical care, which solves the problem of the lack of a safety net by ensuring that there is a limit to how low one can fall in a just society. This kind of safety net guarantees that the ambulance passes by neither the negligent nor the prudent, because it includes an assured level of health care. This also solves both abandonment problems, since people are afforded health care regardless of their responsibility characteristics. Moreover, the difference principle guarantees each person a basic minimum income. If equal basic liberty, fair equality of opportunity, and the difference principle are in effect, we can say that caretakers have all the liberties of citizens, have made the free choice to become caretakers against a full background of options, and are guaranteed an income determined by the difference principle. Men and women are equal citizens, and consequently, are guaranteed individual incomes. Hence, democratic equality can solve all four of luck egalitarianism's problems.

5. How Democratic Equality and Luck Egalitarianism Could Be Combined

We seem to have come full circle. Including responsibility and solving the problem of fine-grainedness with Rawls's conception of justice lead to some troubling difficulties. Not including responsibility leads to the Anderson

problems. How do we proceed from this impasse? First of all, we need to identify where exactly the problems lie. For this purpose, it is instructive to compare democratic equality with the envy test.

The envy test can be applied to any background structure where property rights are guaranteed. Consider a caste system, in which beliefs are widespread that some group in society should do the menial tasks and some other group has a claim to high-ranking positions and where the educational backgrounds of the citizens reflect this. When equal assets have been allocated and trade has taken place, an envy-free distribution of goods will be in place. But this distribution will risk mimicking the caste ideal. If people believe that they are not entitled to positions of power and prestige, they will not pursue such positions. They will not invest in the kind education necessary to move up or in business ventures that would provide advantage. The members of the upper caste will occupy the positions of wealth and power. The lower castes will end up in society's lower positions. The envy test misses this aspect of responsibility, and must be complemented with an account of background justice.

Should luck egalitarianism then be rejected? No, it should be understood as a conception of allocative justice. Since luck egalitarianism is silent on background justice, there will be no contradiction or incoherency in appending a conception of social justice to that theory. As a theory of social justice, we could prefer the Rawlsian conception, but this does not imply that democratic equality is helpful to, or even appropriate for, allocative justice. We must also consider what allocative decisions we should make when background justice has been implemented.

Anderson points out that the difference between democratic equality and luck egalitarianism is that the former envisions justice as a social relationship, while the latter views it as a pattern of distribution. The former aims to secure the social conditions of freedom for all. The goal of the latter is to achieve a just allocation of resources. Democratic justice looks at distributive issues through the lens of considering the effect on people's standings as equal citizens. Often, according to Anderson, "the weight of an interest can be determined by considering the impact on a person's standing as an equal in society" (1999, 332). However, there are cases "where the concepts of equal standing and respect don't

yield a determinate answer to how capabilities should be ranked, the ranking may legitimately be left up to democratic legislation” (1999, 332). As we have seen above, Rawls leaves it up to free and equal citizens to take charge of their lives after background justice is implemented; this also implies that social justice does not pronounce on every question of fairness. Anderson says that a democracy should promote the common good, but the question is then how this good is to be understood, if democratic equality does not yield a determinate answer. It is clearly true that there is room for democratic decisions here, but should we also believe that justice is quiet on the design of the policy that will result? Surely it is not the case that anything goes. Presumably, we should allocate fairly. The problem for democratic justice is that it is not a theory of allocative justice.

To illustrate, I will discuss two areas where the problem of allocation appears: wages and the regulation of traffic. First, let us look at how wages should be set. Anderson says that one thing that is owed to equal citizens is “the right to receive fair value for one’s labor” (1999, 318). Can Rawlsian background justice give an answer to the question of fair wages? It seems not, since the principle that governs the distribution of income in society deals with the distribution between groups. The difference principle points to a specific shape of the curve of distribution of resources in society, but it is silent on the question of whether or not the pay people receive in a society that has implemented that curve is fair. When the difference principle has been applied, it is still possible that within the given range, individuals have wages different from what would be just. We cannot use the two Rawlsian principles of justice to analyze this question, because it is not a problem concerning the background of transactions, but a question of justice *in* transactions. Democratic equality is silent on such problems. Of course, democratic equality will influence what the distributions in such cases will be, but that is because it defines the background for transactions, not because it is a theory which is devised to determine what such outcomes should be. Moreover, we expect wages to take responsibility factors, such as ambition, into account. A person who works more hours, or just works harder, should get more money in return. When background justice is in effect, we would want an allocative principle of justice that takes ambition into account. These are intuitions that are in line with the results we have seen in the experiments in empirical game theory. Democratic equality is neither individualized nor ambition-sensitive, and cannot

be applied to such problems. What was a virtue in social justice becomes a vice in allocative justice. Moreover, this suggests that luck egalitarianism is a plausible candidate for a conception of allocative justice.

It might be thought that this is a problem particular to the economic sphere, but it also affects political issues proper. Assuming that Rawlsian background justice has been implemented, a problem of externalities can serve as an example. Say that a city is experiencing traffic congestion and needs a solution to this problem. The citizens of the city decide to implement a system in which drivers must pay a fixed sum to enter and exit the city. The system would, therefore, hold people responsible for choosing to use their cars. This would presumably cause people with low incomes to decrease their driving the most, while the richer segments of society would make only small adjustments to their driving habits. Would this be just? If we assume the responsibility cut that Rawls made when he developed his theory of justice, it seems that such a policy would come out as unjust. In that case, we are told not to hold people responsible for their choices. This makes strict equality the obvious criterion of justice. However, to make this responsibility cut when we already have assumed background justice would be a strange maneuver. We have implemented a theory of background justice, the purpose of which is to ensure that everyone can develop and exercise responsibility. If we do not hold people responsible for their choices at this point, it would be clear that we are not taking responsibility at all seriously. Therefore, we should look for a theory that says that it is just for those who drive in the city to bear the costs of creating a sustainable traffic situation.¹⁰ Again, this suggests that we should look in the direction of luck egalitarianism for a plausible conception of allocative justice.

Democratic equality is best understood as a conception of social justice, whereas luck egalitarianism is a very plausible candidate for the conception of allocative justice. Democratic equality is silent on issues of allocative justice, and luck egalitarianism does not deal with social justice. This is why democratic

¹⁰ Obviously there might come a point at which background justice is undermined by this policy. If this is the case, social justice ought to take precedence. I will return to the question of precedence below.

equality and luck egalitarianism can be coherently combined. They deal with different problems.

6. Why Democratic Equality and Luck Egalitarianism Should Be Combined

Even if we find that we can combine these two conceptions of justice in one theory of justice, we also need to handle the separate question of why we should go for such a combination. If we look at this question from a Rawlsian perspective, we find that we need to complement the two principles with a conception of allocative justice, that it would be incoherent to argue against the value of responsibility within well-ordered societies, and that we need an allocative principle that is individualized and responsibility-sensitive. The Rawlsian needs to respond to Dworkin's critique. If we instead start from luck egalitarianism, we find ourselves in the position that we need to complement that theory with a conception of background justice, which defines the property rights that the envy test takes for granted and that solves the Anderson problems. The first answer, then, is that the two conceptions solve each other's problem.

We find that that the proponents of democratic equality should be looking for an idea of allocative justice that is individualized and ambition-sensitive, and we are reminded that we had settled for democratic equality as a conception of social justice without having replied to Dworkin's criticisms of Rawls. The first reply should be that democratic equality is silent on the question of allocative justice. It is, in fact, compatible with any number of conceptions of fairness in allocation. It does not say who should get what, but rather talks about the fair background conditions for allocative decisions. The second reply should be to admit that democratic equality is incomplete as a theory of justice if it cannot advise at all on questions of allocative justice. The proponent of democratic equality should, then, welcome the application of luck egalitarianism to problems of allocative justice within societies where social justice is in effect. They ought to welcome a theory that is fine-grained and committed to the view that considerations of responsibility should play a role when fair background conditions are substituted for the natural and social lotteries. By accepting the

combined view, the Rawlsian gains a way of responding to Dworkin's criticism, without having to give up any part of his or her conception of social justice.

It is natural to think of luck egalitarianism as a conception of social justice. However, all the basic intuitions can be retained, even if we, instead, decide to conceive of it as a conception of allocative justice instead. Allocations will still be ambition-sensitive, but not endowment sensitive, even if the allocations are constrained by the difference principle and fair equality of opportunity. Furthermore, allocations can be fine-grained against any backdrop. In a picture, luck egalitarianism says that we should be held individually responsible for how we approach lotteries; it does not say anything about what the structure of the prizes in the lottery should be. This is how the combined view can solve the four Anderson problems, without compromising the core values of luck egalitarianism. The difference principle solves the problems of callousness, by setting up a minimum level of welfare that no one can fall below. It defines the prize structure. Holding people responsible, applying luck egalitarianism, against this background means that the reckless will have to accept to end up among the least advantaged. Now, luck egalitarians could insist on their theory being understood as a theory of social justice, but this would just mean that the least attractive parts of the theory would be retained, without any clear offsetting gains. We have seen that the combined view is ambition sensitive, while avoiding the problems of callousness, whereas luck egalitarianism has the implication that holding people responsible means that, *e.g.*, the reckless should have absolutely nothing. It is hard to see why this should be an essential aspect of the position. This in turn makes it hard to see what could be gained by insisting that luck egalitarianism should be interpreted as a theory of social justice. Furthermore, it is hard to see what a luck egalitarian could have against fair equality of opportunity. Inequalities in opportunities might even seem to fall under the category of endowment inequality. Democratic equality solves Anderson's four problems, and by applying luck egalitarianism within societies that have implemented social justice, we can *also*, and at the same time, solve the two problems that Dworkin finds with Rawls. In order to solve the problems of democratic equality, we need luck egalitarianism, and vice versa. The solution to the impasse is, then, to combine democratic equality and luck egalitarianism.

In order to get a clearer view of how this theory - let's call it the combined view - would work, let us revisit Dworkin's island. The problem faced by its shipwrecked inhabitants is how they should go about dividing a bundle of goods among a group of individuals, knowing their preferences, needs, and desires. If these Robinsons had instead been washed up on an island with a Rawlsian basic structure somehow in place, they could use the same auction mechanism to decide exactly which person would get what resource. They would need a conception of justice that is individualized, that takes people as they are, and that considers responsibility. This is also why they cannot use Rawls's strict egalitarianism to solve their allocative problem. If the background institutions necessary for people to achieve equal power to develop and exercise moral responsibility have been successfully implemented, then it seems unreasonable to again insist on a principle of justice that does not hold people responsible for their choices.

As second way of understanding the question why should democratic equality and luck egalitarianism be combined, would be to take it to mean, why not instead meritocracy or libertarianism? These are the two other conceptions of justice in the Cappelen *et al.* model. Let us start with asking why not meritocracy? One could think that if democratic justice in effect, fair equality has been applied and the value of responsibility is applicable, then it would only be reasonable to hold people responsible for their talents, and not only for their choices. However, we should remember that one of the reasons that we need the difference principle is that inequalities due to innate talent must be alleviated, since the natural lottery cannot be stopped completely. The difference principle assuages the inequalities that result from differences in talents, it does not eradicate them. Any difference in innate talent must be a result of the natural lottery; otherwise fair equality of opportunity would have equalized it. Social justice does not make innate differences in talent fair. Therefore, meritocracy should not take the place of luck egalitarianism.

We should also prefer luck egalitarianism to libertarianism as the conception of allocative justice. The reason for this has to do with the problem of determining to what degree different individuals are responsible for surplus produced in collective ventures, such as firms. Anderson points out that the "the

productivity of a worker in a specific role depends not only on her own efforts, but on other people performing their roles in the division of labor. Michael Jordan could not make so many baskets if no one kept the basketball court swept clean” (1999, 322). The result of production cannot be fully traced back to the responsible actions of any given person. The surplus has been produced collectively. Exactly how many baskets would Michael Jordan have made on littered courts? Possibly quite a few, but surely fewer than he actually managed. This suggests that if the results of exogenous variables should be equalized, then wages should be equal within the firm to the degree that the result of the firm is not traceable to the effort of given individuals.¹¹

Basically, then, these two other conceptions get responsibility wrong. In the case of libertarianism, this conception has no material to work with when there are no identifiable personal responsibility characteristics. Meritocracy does still give the natural lottery too much sway. The second answer, then, to the question of why luck egalitarianism and democratic equality should be combined, is that a combination that includes meritocracy or libertarianism would still have drawbacks that the marriage between democratic equality and luck egalitarianism would not have.

If we opt for the combined view, we would then simultaneously hold two different conceptions of the concept of justice as true. Since they deal with different problems there this would involve no theoretical contradiction. This is why they can solve each other’s problems. However, there would be conflicts in practice, when the accumulative effects of holding people responsible in too many areas lead to violations of the difference principle, as might be the case in the earlier road-toll example.¹² In such cases, social justice should take precedence over allocative justice, since allowing our concerns (based on considerations for justice) about holding people responsible to undermine the preconditions for holding people responsible is getting things backwards. We need to have social justice in effect, in order to be able to apply our conception of allocative justice

¹¹ Or, perhaps, if we want to add some efficiency considerations, the gains should be shared according to the demands of a local difference principle.

¹² Such conflicts show that democratic equality and luck egalitarianism sometimes affect the same thing, but not that they are about the same things. That principles of beneficence and esthetics may come into conflict over whether to set fire to a painting to keep a sick child warm does not show that they concern that same value problem.

properly. Democratic equality fixes the property rights that the envy test takes as its starting point. Giving luck egalitarianism priority would undermine democratic equality, whereas the opposite ordering allows us to maintain the insights of the former theory while avoiding its drawbacks. Therefore, implementing a basic structure governed by the two principles should be given priority over the achievement of allocative justice.

7. Concluding Remarks

This paper has examined the idea that the concept of justice should be conceptualized differently for the problems of allocative and social justice. I hope to have shown that there is a natural way to combine the Rawlsian intuition that we should not be held accountable for the type of society into which we are born and the Dworkinian intuition that justice should be individualized and choice-sensitive. The same underlying concept of justice may be able to explain both intuitions. The question of social justice and that of allocative justice will have very different answers. Democratic equality and luck egalitarianism are two different conceptions of justice and are best interpreted as having different areas of application. We should make two responsibility cuts.

I have suggested that we should be both democratic and luck egalitarians, in (perhaps even lexical) order, because we need to consider both social and allocative justice. To do this, we must interpret luck egalitarianism not as a theory of social justice, but as a theory with the more narrow scope of just allocations. This is contrary to most standard readings of Dworkin's theory, but it is consistent with the justificatory story he presents. Furthermore, such a combination solves difficulties with both positions without compromising their core insights. Rawlsian justice solves the four problems of luck egalitarianism that we have discussed by providing fair equality of opportunity and by applying the difference principle. Luck egalitarianism solves the two major problems that Dworkin found in Rawls's arguments. It gives us a way to achieve justified ambition-sensitive individual allocations against a fair background. It is not sensible to hold people responsible before a social structure that can sustain responsibility is in place. When social justice is implemented, the background structure needed to create

responsible citizens is in place, and the difference principle will have defined what “equal shares” mean in a given society. It would be unreasonable to say that people are responsible and then refuse to hold them responsible. Therefore, we should be luck egalitarians within democratic egalitarian societies. The general liberal egalitarian need not choose between luck egalitarianism and democratic equality.

Acknowledgements

I would like to thank Alexander Cappelen, Bertil Tungodden, James Konow, Mikael Dubois, Kalle Grill, Niklas Möller, Sara Belfrage, Per Wikman Svahn, Isaiah O’Rear, Helen Frowe and Sven Ove Hansson for questions and suggestions. Special thanks to Marcus Widengren and Luciana Widengren for invaluable help. In the context of combining the different yet similar I must also thank Lotta Lindblom.

References

Anderson ES (1999) What Is the Point of Equality, *Ethics* 109:287-337
<http://www.journals.uchicago.edu/doi/abs/10.1086/233897?journalCode=et>

Arneson R (1989) Equality and Equal Opportunity for Welfare. *Philosophical Studies* 56:77-93
<https://link.springer.com/article/10.1007%2FBF00646210?LI=true>

Bolton GE, Ockenfels A (2000) ERC: A Theory of Equity, Reciprocity, and Competition. *American Economic Review* 90:166-93
<https://www.aeaweb.org/articles?id=10.1257/aer.90.1.166>

Camerer C (2003) *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton University Press, Princeton
https://books.google.se/books?id=o7iRQTOeoAoC&printsec=frontcover&hl=sv&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false

Cappelen A, Hole AD, Sørensen EØ, Tungodden B (2007) The Pluralism of Fairness Ideals: an Experimental Approach. *American Economic Review* 97:818-827 <https://www.aeaweb.org/articles?id=10.1257/aer.97.3.818>

Cappelen A, Erik. Sørensen EØ, Tungodden B (2010) Responsibility for what? Fairness and individual responsibility. *European Economic Review* 54: 429-441
<http://www.sciencedirect.com/science/article/pii/S0014292109000890>

Charness G, Rabin M (2002) Understanding Social Preferences with Simple Tests. *Quarterly Journal of Economics* 117:817-869
<https://academic.oup.com/qje/article/117/3/817/1933015/Understanding-Social-Preferences-with-Simple-Tests>

Cohen GA (1989) On the Currency of Egalitarian Justice. *Ethics* 99:906-944
<http://www.journals.uchicago.edu/doi/abs/10.1086/293126?journalCode=et>

Dworkin R (1977) *Taking Rights Seriously*. Duckworth, London
https://books.google.se/books?id=-HuwatdQKhgC&printsec=frontcover&hl=sv&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false

Dworkin R (1981) What is equality? Part 2: equality of resources. *Philosophy and Public Affairs* 10:283-345
<http://www.jstor.org./stable/2265047>

Dworkin R (2000) *Sovereign Virtue – The Theory and Practice of Equality*. Harvard University Press, Cambridge
https://books.google.se/books?id=SIOPnqu3f5kC&printsec=frontcover&hl=sv&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false

Fehr E, Schmidt KM (1999) A Theory of Fairness, Competition, and Cooperation. *Quarterly Journal of Economics* 114:817-868
<https://academic.oup.com/qje/article/114/3/817/1848113/A-Theory-of-Fairness-Competition-and-Cooperation>

Hart HLA (1994) *The Concept of Law*, 2nd edition. Clarendon Press, Oxford

https://books.google.se/books?id=hCoUDAAAQBAJ&printsec=frontcover&hl=sv&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false

Konow J (1996) A Positive Theory of Economic Fairness. *Journal of Economic Behavior and Organization* 31:13-35
<http://www.sciencedirect.com/science/article/pii/S0167268196008621?via%3Dihub>

Konow J (2000) Fair Shares: Accountability and Cognitive Dissonance in Allocation Decisions. *American Economic Review* 90:1072-1091
<https://www.aeaweb.org/articles?id=10.1257/aer.90.4.1072>

Miller D (1999) *Principles of Social Justice*. Harvard University Press, Cambridge
https://books.google.se/books?id=y2wMzJtEZ_8C&printsec=frontcover&hl=sv&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false

Rawls J (1971) *A Theory of Justice*. Oxford University Press, Oxford
https://books.google.se/books?id=kvpby7HtAeOC&printsec=frontcover&hl=sv&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false

Rawls J (1996) *Political Liberalism*, 2nd edition. Columbia University Press, New York
https://books.google.se/books?id=vXGZRYCkaNsC&printsec=frontcover&hl=sv&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false

Rawls J (2001) *Justice as Fairness: A Restatement*. The Belknap Press of Harvard University Press, Cambridge
https://books.google.se/books?id=AjrXZIlbK1cC&printsec=frontcover&hl=sv&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false

Roemer JE (1996) *Theories of Distributive Justice*. Harvard University Press, Cambridge
https://books.google.se/books?id=qQkX24xj9akC&printsec=frontcover&hl=sv&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false

Yaari ME, Bar-Hillel M (1984) On Dividing Justly. *Social Choice and Welfare* 1:1-24
<https://link.springer.com/article/10.1007%2FBF00297056?LI=true>



NORMS FOR THE PUBLIC REMEMBRANCE OF NONHUMAN ANIMALS

NORMAS PARA A CELEBRAÇÃO PÚBLICA DE ANIMAIS NÃO- HUMANOS

Matthew McLennan*
mmclennan@ustpaul.ca

Abstract. *This article builds upon Avishai Margalit's distinction between ethical and moral norms of remembrance. While Margalit is limited by his broadly Kantian framework and restricts his arguments to the remembrance of human beings, the author will argue that the resources exist both in his account and in the particularities of Canadian public life to a) account philosophically for what minimal public ethical norms are in place for the remembrance of nonhuman animals, and b) point towards a more robust, properly moral account of nonhuman animal remembrance. The author will take a recent Canadian case study in the public remembrance of nonhuman animals– the 2012 Animals in War Dedication – to show how existing norms are inherently unstable, pointing beyond themselves to a more species-inclusive, properly moral public perspective.*

Keywords: *morality, ethics, memory, remembrance, animals, Margalit.*

Sumário. *Este artigo baseia-se na distinção feita por Avishai Margalit entre normas éticas e normas morais de celebração. Enquanto Margalit está limitado pelo seu quadro de referência genericamente kantiano e restringe os seus argumentos à celebração dos seres humanos, o autor argumentará que existem recursos, tanto na sua posição quanto nas particularidades da vida pública canadiana, para: a) explicar filosoficamente as normas éticas mínimas que existem para a celebração de animais não-humanos, e b) apontar para uma posição mais sólida e adequadamente moral acerca da celebração de animais não-humanos. O autor usará um estudo de caso canadiano recente sobre a celebração pública de animais não-humanos - a Dedicação de Animais em Guerra de 2012 - para mostrar como as normas existentes são intrinsecamente instáveis, apontando além de si mesmas para uma perspectiva pública mais apropriada e inclusiva em termos de espécies.*

Palavras-chave: *moralidade, ética, lembranças, celebração, animais, Margalit.*

* Saint Paul University.

1. Ethics and Morality of Public Memory

Avishai Margalit has written compellingly of both an “ethics of memory” and a “morality of memory”. His point of departure is a story taken from a Jerusalem local newspaper, in which an army colonel publicly forgot the name of a soldier killed under his command. As Margalit recounts, “There followed a flood of outrage at the officer who did not remember.” (Margalit, 2002, p. 19) This minor episode reveals an aspect of public life that is upon reflection both obvious and under-theorized: remembrance must measure up to certain norms which often are largely implicit until they are violated.

This article is an extension of Margalit’s discussion. What is at issue is the public remembrance of nonhuman animals, a topic which does not factor into his account. The argument is centered upon a recent Canadian case study which illustrates how current practices and norms of the remembrance of nonhuman animals – an “ethics of public memory,” regarding at least some species and classes of nonhuman animal – point beyond themselves to a more robust, species-inclusive “morality of public memory”.

It is important at the outset to define our terms. Margalit distinguishes between *ethics* – our duties based upon “thick” relations to family, community, tribe, etc. – and *morality* – our duties to humanity as such. (Margalit, 2002, p. 7) Note that while a community is ideally both ethical *and* moral, these are conceptually as well as practically separable; it is possible to practice “ethics without moral constraints”, a condition that Margalit calls “tribalism” and which can take extreme, aggressive forms as evinced by the Nazi regime. (Margalit, 2010, p. 122) It is also possible, as in the case of the scientific community, to stand in a moral but not always ethical relation to one’s peers, relating purely through rational discussion and abstract respect. (Margalit, 2002, pp. 145–146)

Within this framework, it is therefore meaningful to speak of distinctly *ethical* and *moral* norms guiding what, whom and how we remember. For instance, I might have an *ethical* duty to visit the grave of my uncle and recount to my children something of his life; kinship is an important value, he played a

role in my upbringing, and so on. Here it would be a question of proper norms of respect for those who sustained me when I was vulnerable, and to whom I owe a duty of rites and narratives of remembrance. As Margalit puts it, “Memory is the cement that holds thick relations together, and communities of memory are the obvious habitat for thick relations and thus for ethics.” (Margalit, 2002, p. 8) On the other hand, I have a *moral* duty to remember, mark and recount such events as the Holocaust, slavery, the genocide of First Nations in the Americas, and the Rwandan genocide. These were attacks on the very idea of a shared humanity, which I will gloss from Margalit’s account as *the presumption of the capacity of each human being to radically change his or her life, i.e. to start over, morally speaking*. (Margalit, 2007, pp. 72–76) In participating in acts of mourning and remembrance of crimes against humanity, I affirm precisely the human ideal that they sought to undermine or destroy.

Margalit is broadly Kantian in his approach to morality, but for him the source of human superiority is immanent rather than transcendent. Thus he cleaves to a modified form of humanism:

I take humanism ... to consist of two claims and not just one: first, that human beings are the only source of justification for ethics and morality; second, that humans are a sufficient source for the justification of ethics and morality. I agree with the first claim but not with the second: I believe that human beings are the only source of justification but that this source is not sufficient. (Margalit, 2002, pp. 183–184)

In this way, Margalit may be read as a “fallibilist” or perhaps “post-metaphysical” Kantian; he maintains that there is no absolute fulcrum for the justification of human ethics and morality other than the process of justification itself. (See also Putnam, 2004)

Like Kant, Margalit makes indirect room for nonhuman animals. But this means that his limitations regarding nonhuman animals echo Kant’s. For Kant, we can speak of indirect but not direct duties to nonhuman animals, since animals are at best only symbols of the human reason which grounds morality. (Kant, 1980, pp. 239–241) Similarly, in Margalit’s case we might speak of ethical but not moral norms regarding nonhuman animals since while many of them participate in our communities, none of them participate in a common humanity. They are perhaps, at best, symbols but not bearers of moral agency. He spells this out explicitly: when we speak of respect for nonhuman animals, really we’re speaking of self-respect. (Margalit, 2007, p. 65)

In describing the cruelty and humiliation which combine to make an inhuman i.e. immoral regime, Margalit relates Hitler's command that the plotters against him in 1944 be "hung like cattle"; "Not treating the dead body of the condemned as human dead body is an expression of humiliation." (Margalit, 2010, p. 65) It is not a question of claiming that nonhuman animals, as opposed to human beings, *can* be treated as mere objects. Rather, the claim is that to treat a human being "like an animal" is inherently immoral. (Margalit, 2007, p. 89) As Margalit puts it: "Morality is the undertaking to regulate human relations between human beings. Human beings should also establish humane relations with animals, but that is a different undertaking." (Margalit, 2010, pp. 140–141) This undertaking would not, by definition, include moral norms of remembrance.

I will repeat that I hold Margalit to take a "broadly", not a strictly, Kantian position. Indeed, he gives an explicit analysis of Kant's norms of human respect that reveal the extent to which he, Margalit, both is and is not a Kantian. (Margalit, 2007, pp. 66–67) The problem here is that, like Kant, Margalit courts risk by grounding the definition of "humanity" in a capacity that is presumably restricted to fully competent adults. Children and those with severe, possibly even moderate, intellectual disabilities are at issue. We might set children aside, to the extent that – all things being equal – they will grow into their moral freedom and therefore already possess it, very loosely speaking, as a "capacity". This move cannot be made for those with irremediable intellectual impairment. At the limit, this would imply that crimes against some human persons who are not moral agents would not technically count as crimes against humanity. Oddly, the Nazi genocide of persons with intellectual disabilities, for example, would be unethical but not immoral on this picture. Subsequently, there would then be an ethical, but not a moral, duty of remembrance of the crime. This is counterintuitive and puts the interests of such persons in a precarious place.

One way to try to escape the problem – assuming that we wish to retain radical moral freedom as the necessary criterion of respect for humanity – would be simply to grant persons with intellectual disability honorary status as moral agents, i.e. an honorary human status. It seems that Margalit is forced to make this move, problematic as it sounds. He claims "the chief premise of morality" to be "the idea that all human beings should be subjected to moral treatment solely

because they are human.” (Margalit, 2010, p. 190) But we have seen that he is precisely *not* talking about the human species, i.e. human DNA as a source of moral standing; he is referring to moral autonomy. The chief premise of morality is therefore either circular and exclusionary (the moral autonomy of some members of the species *homo sapiens* commands the moral treatment of morally autonomous members of the species *homo sapiens*) or it is inclusive but arbitrary (the moral autonomy of some members of the species *homo sapiens* commands moral treatment for all members of the species *homo sapiens*). In this connection, Margalit shows awareness of the need to extend his premise in an honorary if arbitrary way when he critiques common social attitudes towards adults with Down syndrome; specifically, he disagrees that they may be treated as perpetual children or as less than fully human. (Margalit, 2007, p. 110) Agreeing with Margalit that persons with Down syndrome should be treated with respect, I differ from him over the basis of this respect. It is not the high level of moral functioning possessed by autonomous, average adult humans but rather something more basic.

This basis for respect or moral standing – sentience, being a “self” – will be further discussed below. For now, note the problem: Margalit’s picture is vulnerable to the “argument from species overlap.” (Horta, 2014) If human persons who fail to meet the stated criterion for humanity can nonetheless be included under the umbrella of “shared humanity”, then there is no obvious reason why at least some species of nonhuman animals cannot also be included. To shift the basis of inclusion to biological species would be to move the goalposts in an arbitrary way.

2. Case Study: The Animals in War Dedication

The preceding shows how Margalit’s account of ethics and moral norms of remembrance cannot exclude nonhuman animals if it wishes to avoid the charge of arbitrariness. I will now draw upon a Canadian case study to show how the issue is not entirely theoretical. Existing practices in the public remembrance of nonhuman animals are philosophically unstable and, like Margalit’s theory, point beyond themselves to more robust, species-inclusive moral public norms.

On November 3, 2012, the Government of Canada unveiled the Animals in War Dedication in Confederation Park, Ottawa.¹ According to Veterans Affairs, the Government of Canada contributed “more than \$98,000” towards the monument’s creation and unveiling.²

The unveiling ceremony was attended by “Mr. Royal Galipeau, Member of Parliament for Ottawa–Orléans and representative for the Honourable Steven Blaney, Minister of Veterans Affairs, Senator Yonah Martin, Mrs. Laureen Harper, Honorary Patron of the Animals in War Dedication Project, Brigadier-General M.K. Overton, Assistant Chief of Military Personnel of the Canadian Armed Forces, and Mr. Russell Mills, Chair of the National Capital Commission (NCC), along with [World War II and Korean War] Veteran Lloyd Swick, founder of the project” The dedication “consists of three interpretative plaques explaining the roles played by animals during past wars. A bronze statue of a medical service dog stands nearby.” “The footprints of dogs, horses and mules are stamped into the concrete of the Animals in War Dedication, representing the marks they left on the battlefield.”³ The dog stands loyally near the foot of the South African War Memorial. The placement itself is symbolic, since in the Boer War “Canada supplied 50,000 horses for mounted troops.”⁴

In its news release, Veterans Affairs quotes distinguished participants of the unveiling and briefly describes the role of animals in Canadian military campaigns:

‘As a tribute to the efforts of animals who served during crucial battles, we honour their unwavering loyalty, dedicated service, and strong companionship during difficult times’ said Minister Blaney. ‘With the unveiling of this dedication, Canadians now have a place to honour animals who’ve served in war alongside our Veterans,’ said MP Galipeau. ‘Animals have always been a part of our lives and of our culture and should be recognized for their contributions to Canada’s war efforts.’ ‘This dedication in Canada’s Capital Region will help Canadians discover the contributions of animals in war,’ said Mr. Mills, Chair of the NCC (...) A variety of animals were used during war. Mules carried supplies and artillery; horses hauled field guns; carrier pigeons delivered messages to specific destinations; and dogs

¹ See the official website of the monument at <http://aiwdedication.ca/> (accessed Feb.12, 2013).

² <http://www.veterans.gc.ca/eng/department/press/viewrelease/1610> (accessed Feb.12, 2013).

³ <http://canada.pch.gc.ca/eng/1443025436013> (accessed August 16, 2016)

⁴ <http://canada.pch.gc.ca/eng/1443025436013> (accessed August 16, 2016)

worked as messengers, medical assistants, mine detectors and in search and rescue. Dogs are still employed by the Canadian Armed Forces today.⁵

The Animals in War Dedication is not the first of its kind in Canada; the Tympanum of the Peace Tower on Parliament Hill “represents the animals that served during the [First World] war: reindeer, pack mules, carrier pigeons, horses, dogs, canaries and mice”, and bears the inscription: “THE TUNNELLERS' FRIENDS, THE HUMBLE BEASTS THAT SERVED AND DIED.”⁶ Similar monuments may be found abroad, for example Britain’s Animals in War Memorial.⁷

That the monument is of philosophical interest is evident in the language surrounding its unveiling. Unlike their human counterparts, military service animals were “used” for specific purposes; in other words they are considered to be objects, or means to an end. Yet the animals honoured also “worked” at particular jobs; they are thus considered to be labourers, subjects, and, arguably, ends in themselves. The troubled language surrounding the unveiling speaks to a deep philosophical tension.

I do not in the least critique the admirable motives of project founder Lloyd Swick, but rather mount a philosophical criticism based on the clash between what the monument appears to be communicating on one hand, and the policies and practices of the government which erected it on the other. Orienting my argument around Peter Singer’s classic utilitarian defence of animals, as well as Sue Donaldson and Will Kymlicka’s 2011 study *Zoopolis: A Political Theory of Animal Rights*, I will suggest that the Canadian government’s gesture of recognizing the contributions and sacrifices of some species of nonhuman animals is suggestive of a political philosophy that it simply does not believe in, and policies it does not advocate. Were it not for the fact that it continues to support the military use of some nonhuman animals in particular, and existing systems of animal exploitation more generally, the Government of Canada’s erection of the Animals in War Dedication would bear an altogether different and more straightforward meaning.

⁵ <http://aiwdedication.ca/> (accessed Feb.12, 2013).

⁶ http://www.parl.gc.ca/About/House/Collections/heritage_spaces/memorial/stone/3369-e.htm (accessed Sept. 16, 2016)

⁷ <http://www.animalsinwar.org.uk/> (accessed August 16, 2016)

Giving a version of the aforementioned “argument from species overlap” I will show, first, that if we do not want to exclude certain classes of human beings, then we must accord equal moral standing to nonhuman animals. This furnishes grounds for a critique of the government’s unveiling of the new monument, as I will explain. Next, I will question whether and to what extent animals should have differentiated political rights, noting that the new monument treats nonhuman animals as citizens; this too furnishes grounds for a critique of the government’s gesture. I conclude this section by suggesting that the adoption by the Government of Canada of a robust account of animal moral standing and/or political rights would be a necessary condition to removing the contradiction implied by its unveiling the monument. For the Animals in War Dedication to truly honour nonhuman animals in war, the Government of Canada would need to radically alter both its moral and political conception of nonhuman animals, as well as its policy of using some animals for military purposes to this day, as I will describe.

To have “moral standing” is to count in moral deliberations. “Do nonhuman animals count, morally speaking?” is a question that has been debated extensively in academic literature. But it is generally recognized, both within and without the academy, that there should be at least some moral and legal limits to the human exploitation of nonhuman animals. There are major philosophical disagreements as to the specific reasons for which these limits should be drawn, and there are likewise palpable differences over their practical implications. I cannot hope to do justice here to the range of theoretical options pertinent to these questions. For the purposes of the following, however, the discussion will roughly distinguish between consequentialist and rights positions (noting in passing that options such as contractarian, care and capabilities approaches trouble this distinction). Consequentialists make the comparatively modest claim that nonhuman animals have moral standing, i.e. that we should count their interests when we deliberate, and that what happens to them matters, morally speaking. Those arguing from a rights position go further and claim that there are inviolable moral limits to our dealings with nonhuman animals (i.e. they have moral rights in the strict sense).

Recognizing nonhuman animal moral standing is not necessarily tantamount to recognizing species equality. Many people devote considerable yet rather selective efforts to animal welfare. For example, Laureen and Stephen Harper publically supported the Ottawa Humane Society's Foster Program, pleading on the Prime Minister's website for the care and humane treatment of cats and dogs. But Prime Minister Harper's tweet from December 21, 2012 underscores the selectiveness of his commitment to animal welfare. He tweets "@HomerJSimpson Mmm... bacon" and we are treated to a clip from *The Simpsons* in which several characters taunt a vegetarian (singing "you don't win friends with salad").⁸ The political gaff of tweeting about the enjoyment of food during First Nations Chief Theresa Spence's hunger strike should of course be noted, but for our purposes the PM's tweet is interesting in that it communicates very clearly that there will be no special pleading for pigs.

Peter Singer's classic consequentialist account argues convincingly, however, that "all animals are equal" (Singer, 2010); that is to say, that they have equal moral standing. Humans, pigs, dogs, cats, birds and all roughly comparable organisms are to count equally in moral deliberation. The reason for this is simple; all such beings have the capacity to suffer, and therefore have an equal stake in the avoidance of suffering. It may seem that Singer has set the bar for moral standing excessively low, but his reasons are compelling. If the basis for moral standing were something like reason, linguistic ability or, as we saw with Margalit, moral agency, then we would have to exclude a wide swath of humanity from (direct) moral standing. Many human beings are so intellectually disabled that they will never possess these qualities. In some cases – acknowledging both the risk of ableism attendant to this move, as well as the conceptual difficulties of comparing cognition across species (Taylor, 2017; De Waal, 2016) – nonhuman animals could be argued to cognitively outrank them.

The existence of certain atypical (often called "marginal") human cases has in any case long troubled those critics of Singer who would maintain the moral inequality of species. We could by fiat claim that membership in the species *homo sapiens* puts the severely intellectually disabled above animals, morally speaking.

⁸ <https://twitter.com/stephenharper/status/282233267623714816?lang=en> (accessed October 6, 2017)

But then this would be a case of moral preference based on an arbitrarily selected and apparently morally irrelevant factual criterion. We might as well restrict moral standing to sex, or race, or physical strength. Indeed, Singer claims that wherever we are prepared to treat sentient beings differently based on species membership, this is “speciesism”, by analogy with sexism and racism. (Singer, 2010, pp. 569–571) No qualitatively richer, typically human capacity, such as reason or the ability to communicate or to be moral agents, puts us above nonhuman animals – unless we are prepared to be ableist (deny moral standing to humans who lack such intellectual capacities). Getting around this problem by granting special inclusion to the intellectually disabled is, to repeat, speciesist and therefore philosophically and morally indefensible; in short, an intellectually honest account of moral standing sufficiently broad to include severely intellectually disabled humans would also have to include nonhuman animals, irrespective of species. Were Prime Minister Harper consistent in his animal welfarism he would plead for pigs as well as for cats and dogs.

Note however that recognizing the equal moral standing of nonhuman animals is not the same thing as recognizing their rights. Being a utilitarian, Singer does not believe in moral rights at all. Though he typically rejects arguments in defense of nonhuman animal exploitation, he is prepared to grant that there could be times when exploiting nonhuman animals (or humans) serves aggregate utility and is therefore acceptable. He is even prepared to make allowances for eating nonhuman animal proteins. For example, there is no moral presumption in favour of a vegan diet in much of the global South, where people lack readily available alternative sources of vitamin B12. (Singer & Mason, 2006, pp. 226–237) Similarly, it might be argued that there would be no such presumption on the part of Canada’s First Nations. But where protein alternatives and supplements are within reach of urban settler populations, many people of average health would appear to face a utilitarian presumption against nonhuman animal exploitation for dietary reasons.

How then does all this shed light on the Animals in War Dedication? As MP Galipeau puts it, the monument is erected in the spirit that nonhuman animals “should be recognized” for their contributions and, by extension, their sacrifices. Indeed, one of the plaques on the dedication describes how many horses in WWI

“paid the ultimate price” for their contributions. This betrays recognition of their interests, and such recognition, if intellectually honest, automatically grants them moral standing. But note two things. First, recognition of moral standing is not tantamount to recognition of equal moral standing, as we saw in the case of the selective animal welfarism of the Prime Minister. In erecting the monument, the Government of Canada is praising the contributions of some animals while continuing to legally allow and even subsidize the exploitation and deaths of countless others. But speciesism is, to repeat, philosophically indefensible. In this light, the Dedication may be read as a sentimental yet arbitrary gesture.

Second, consequentialist views such as Singer’s are, in principle, perfectly compatible with sending nonhuman animals to war if doing so contributes to aggregate utility. It would appear then that from a rigorously egalitarian consequentialist approach to nonhuman animals, there is nothing amiss with the monument. But bear in mind the following consideration. Canada does not practice conscription of human military personnel. Military service for human beings is, in principle, a free choice. Since nonhuman animals cannot understand the nature of military service and therefore cannot consent to it, their participation is not similarly free. The government thus by definition conscripts animal personnel and as such it makes special allowances for conscription based on species membership. Here we could cite the argument that nonhuman animals lack the capacities of normal adult Canadians and therefore may be conscripted without moral compunction. We could also argue that in some cases, as with explosive-detecting dogs, the deployment of nonhuman animals directly saves human lives.

But note that we do not and would not conscript intellectually disabled humans to military service on these same grounds (indeed, in principle they would be screened out of the application procedure). Imagine a genetic disorder that renders a human being severely cognitively disabled and incapable of managing her own affairs, but imbues her with heightened ability to detect mines and explosives. No self-respecting nation would send her into service on the grounds of her usefulness, and any proposals to that effect would be rightly condemned. We would and do, however, send a dog with comparable capacities to perform such tasks with little hesitation. Considering that a comparable

monument thanking conscripted, severely intellectually disabled human beings for their contributions to military campaigns would rightly be judged an obscenity, the erection of the Animals in War Dedication bespeaks our country's speciesism and is on these grounds open to criticism.

Consider further the military use of children. The practice is widely condemned because children cannot consent to serve in war by virtue of what is recognized, both in law and in moral theory, to be their diminished autonomy and capacity for moral agency. Since in children this lack of autonomy is not considered to give guardians the right to send them to war, by parity we could argue that intellectual disability and the inbuilt cognitive limitations (from a human point of view – See De Waal, 2016) of nonhuman animals do not likewise count in favour of their conscription. Naturally, a “children in war” monument that did not criticize the military exploitation of children could fairly be called an atrocity added to an atrocity. But it might be objected that statistically normal children possess the potential to become fully autonomous adults; perhaps what is so wrong about their deployment in war is the way in which it irrevocably scars them and thwarts their attainment of future goods. Valuing potential in this way however would be no reason to disallow child conscription on the one hand, and practice animal conscription on the other, since to repeat, many intellectually disabled humans lack the potential of children. Such a policy is caught in a dilemma: either we support ableism (potential autonomy is the criterion for moral consideration) or we support speciesism (being human is the criterion for moral consideration).

In sum, the belief in animal welfare as a morally important goal raises the broader question of the basis of moral standing. If we wish to be sufficiently inclusive in our criteria for moral standing that we do not leave out intellectually disabled human beings, while avoiding the philosophically indefensible move of speciesism, then we need to set the bar such that nonhuman animals are put morally on par with humans. This being so, the Animals in War Dedication is problematic because in the same way that an uncritical dedication to conscripted, intellectually disabled humans or children would be. Had the monument been unveiled to memorialize beings mistreated in the past as the result of benighted policies, then things would of course be different. But there is no such indication

on the part of the Government of Canada, which admits without hesitation or regret that dogs are still deployed in military operations.

Animal Rights Theory or ART would point us to similar conclusions while differing importantly from consequentialist accounts like Singer's. As the name would suggest, such a theory envisions full moral rights for nonhuman animals, in the sense of there being inviolable moral boundaries in our dealings with them. On Donaldson and Kymlicka's view these rights are generated by the possession of selfhood, which is a precondition for having interests. Here again, the criterion for rights is fixed widely enough to protect intellectually disabled human beings, and barring a speciesist exception, this entails comparable basic rights for animals.

Bracketing for sake of argument the philosophical defensibility of rights, there is a good deal to recommend ART over Singer's position. Above, I discussed the protection of the severely intellectually disabled as a matter of hypotheticals: philosophically speaking, *if* we want to protect the interests of such humans, *then* we are compelled to protect those of nonhuman animals as well. But from certain consequentialist positions (for example act utilitarianism) it is always, in principle, possible to make exceptions to the protection of such interests. Some consequentialists, such as R.G. Frey, simply "bite the bullet" (Donaldson & Kymlicka, 2011, p. 28) when it comes to atypical or "marginal" human cases, suggesting for example that if utilitarian arguments in favour of animal testing hold water, then we must be open to testing on intellectually disabled humans as well (in fact, there is according to this view an utilitarian presumption in favour of human testing over animal testing, since human models will give us more scientifically accurate and practically useful information) (Frey & Paton, 2010). In response to this kind of reasoning, many philosophers have tried to hammer out a convincing account of inviolable human rights. But in casting their account wide enough to protect intellectually disabled humans, such rights automatically entail animal rights. A representative but clumsy critique of ART which tries to get around this implication comes from Michael Allan Fox (2006), who insists that the very notion of a right implies corresponding duties (i.e. on the part rights-bearers). This of course generates the troubling implication that severely

intellectually disabled humans lack rights, which Fox denies by making a philosophically indefensible speciesist exception.

So much for inviolable moral boundaries, i.e. *negative* rights. The *Animals in War Dedication* is also interesting because it puts the question of nonhuman animal *positive* political rights – that is to say, citizenship – on the agenda. Reading the *Animals in War Dedication* charitably, it is a gesture of gratitude and praise to such animals for defending Canadian sovereignty. If we can thank such beings for their participation in some aspect of political sovereignty, then we can at least meaningfully pose the question of their political status with regard to Canada as a sovereign nation.

This is where Donaldson and Kymlicka's recent intervention in animal ethics proves helpful. They make the admittedly controversial point that because of its focus on inviolable negative rights for nonhuman animals, classical ART generally envisions animal ethics as an ethics of non-interference. (Donaldson and Kymlicka, 2011, p. 9) The *Animals in War Dedication* could in this light be criticized on the grounds that it sentimentalizes a case *par excellence* of interference with nonhuman animals. But Donaldson and Kymlicka point out that its focus on negative inviolable rights renders the classical rights position insufficiently nuanced, to the point of being counterproductive. It does not, for example, make allowances for the fact that humans and nonhuman animals may engage in mutually beneficial and satisfying cooperative relationships. To this extent, animals such as dogs and horses killed in war should indeed be mourned as lost friends and social collaborators. An ethic of non-interference may indeed be the right approach as far as wild and opportunistic/liminal animals are concerned, as Donaldson and Kymlicka argue in their sixth chapter (though this is highly contentious; see Mannino, 2015 and Faria, 2016). Such an ethic would ignore however the rather intuitive claim that as humans we may have positive duties to species we have domesticated on account of our having rendered them, through breeding, incapable of survival in the wild. Notwithstanding what may be plausibly argued to have been the initial injustice of domestication, the existence of domesticated nonhuman animals generates positive duties on our part with respect to them. In other words, we do not remedy the injustice of domestication by abandoning the animals we have already domesticated.

We must bear in mind what, precisely, is entailed by this positive notion of animal rights. Donaldson and Kymlicka do not advocate for animals to have the right to vote, or any such nonsense; rather, in the same way that severely intellectually disabled humans may be considered citizens but minors or wards of state, domesticated animals should be granted citizenship rights entailing no duties on their part, but specific duties on the parts of human institutions and caregivers. As Donaldson and Kymlicka point out, political agency is only one aspect of citizenship, which also includes national identification and popular sovereignty (2011, pp. 55–61). If domesticated nonhuman animals are excluded from citizenship on the basis of their incapacity for political agency, then surely (barring indefensible speciesism) we would have to similarly exclude infants, children, the intellectually disabled, the temporarily unconscious, and those with dementia. Though admittedly counterintuitive, the claim that domesticated nonhuman animals are entitled to citizenship rights holds water for the reason that its denial renders citizenship an unacceptably fragile and exclusionary matter in the case of humans.

Let us ask, then, if the provisional admission of domesticated nonhuman animal citizenship implies the permissibility of such animals being sent to war, or even a duty on their part to serve Canadian military interests. We may dispose of the possibility that such animals have a duty to protect Canadian sovereignty. Since Canada “lags woefully behind” other industrialized nations “regarding even the most minimal reforms” of animal welfare (Donaldson & Kymlicka, 2011, p. 259), it is not evident that domesticated nonhuman animals have a duty to defend (nor any particular stake in) Canadian sovereignty; the claim that they do smacks of the rather repugnant claim that human slaves are duty bound to defend the political territories of their masters. But even if Canada were to become a world leader in progressive animal policies, this would still not generate the duty for nonhuman animals to serve – unless of course we were prepared to admit that progressive disability policies for humans generate a comparable duty on the part of humans who would thereby benefit. As to whether or not it is at least morally permissible to send such (provisional) nonhuman citizens to war, here again parity with the case of intellectually disabled humans rules out an affirmative answer.

The *Animals in War Dedication*, as conceived by Lloyd Swick, is therefore laudable to the extent that it treats nonhuman animals a) as if they have moral standing, however we cash that out, and b) as if they are citizens. But these very laudable recognitions also prove that the gesture of the Government of Canada behind the monument is philosophically confused, and morally problematic. In essence, the monument does not go far enough in its recognition of nonhuman animals.

Summing up the argument of this section, nonhuman animals do indeed have moral standing, by virtue of their capacity to suffer and have interests. The monument recognizes this. But if we adhere to a consistent and intellectually honest animal welfare framework, the speciesism of the monument also immediately becomes apparent. By erecting the monument in light of existing laws and policies the Government of Canada is doing two things: first, betraying a belief that some animals should be arbitrarily singled out for praise, while others unceremoniously exploited, killed, and eaten; second, symbolically thanking and praising creatures with interests of their own for contributions and sacrifices that we would never demand of cognitively comparable members of our own species. The monument is philosophically and politically indefensible by reason of a) its arbitrariness and b) its support of inexcusably exploitative practices.

As to the much more novel question of whether nonhuman domesticated animals should have citizenship rights, the language of the press release surrounding the monument certainly seems to imply it. It errs not in treating domesticated animals as Canadian citizens so much as in treating them as fully autonomous political agents, capable of consenting to go to war. Considering, once again, that we could never justify sending our severely cognitively disabled or our children to war for the reason that they count as citizens to whom we owe special duties of care and protection, the monument makes a grave philosophical and political mistake. It normalizes gratitude to beings that cannot fully understand the nature of the actions for which they are thanked, and that for this reason cannot consent to them in the first place.

Were the Government of Canada's position on the military use of animals and of the industries of animal exploitation radically different, the monument

could serve as a testament to a dark period in the country's history – and a reminder to do better by nonhuman animals in future. Nonhuman animals should be recognized for the sacrifices they have made in the past, but only insofar as we recognize the injustice of requiring such sacrifices from them. A monument like the Animals in War Dedication must be self-critical; we must not allow it to fulfill its ideological function of resolving symbolically, which is to say superficially, a deep contradiction at the heart of Canadian settler society. We must, above all, hold the Government of Canada to the standards it implies in the erection of the Dedication. These imply nothing less than a discontinuation of the use of nonhuman animals in war, and a radical reduction if not abolition of the industries of animal exploitation. To use the language of the first section: the Animals in War Dedication is laudable as an exercise in the ethics of public memory – but it points to a species-inclusive morality of public memory.

3. Conclusion: From Ethics of Public Memory to Morality of Public Memory

Margalit's notion of an ethics of memory – and the corollary notion of a morality of memory – helped to orient our consideration of existing norms surrounding the public remembrance of nonhuman animals. In strict terms, Margalit's framework proved to be limited. Since nonhuman animals do not fall into the orbit of human nature – what Kant qualifies as the freedom of human reason, and what Margalit qualifies as the radically open moral agency of each human life – it is at best possible on Margalit's account to speak of an ethics, rather than a morality, of public memory concerning nonhuman animals. As such, Canadian settler norms regarding nonhuman animal remembrance are “thick”, “tribal” and imply no universally binding duties; indeed, as we saw they are through and through speciesist – both because they hold nonhuman animals *in toto* at a lower status than human beings, and because only certain species of domesticated nonhuman animal are included in the “tribe”. Thus it is conceptually possible, according to existing norms, to pay tribute to certain species of domesticated animal – dogs, horses, even carrier pigeons who participated in Canadian military campaigns – while systematically abusing, exploiting and killing other species for food, fur and other products on an

industrial scale in a country that boasts some of the least progressive animal welfare regulations among industrialized nations.

I argued however that this conceptual possibility – the speciesism inherent in existing norms of remembrance – is unstable. The ethics of public memory for nonhuman animals as practiced by the Government of Canada, evinced by the notable public gesture of the Animals in War Dedication, gestures beyond itself, in spite of itself, since it a) implicitly recognizes nonhuman animal moral standing and social citizenship and b) is undermined by its own speciesism. Since there is no question of retracting our existing commitment to nonhuman animal commemoration, it is therefore a question of going forward – remembering, for example, the labour and deaths of animals in agriculture and animal husbandry in addition to the sacrifices of military animals. But once we make this commitment, as I have shown, we are on our way to taking seriously the troubling question: what does it mean to *honour* those whom we use and kill without their consent?

Such public gestures are therefore troubling, but they are to be supported in the interim and seized upon as an opening towards more progressive and consistent public values.

References

De Waal, F. (2016). *Are We Smart Enough to Know How Smart Animals Are?* New York and London: W.W. Norton & Company.

Donaldson, S., & Kymlicka, W. (2011). *Zoopolis: A political Theory of Animal Rights*. New York: Oxford University Press.

Faria, C. (2016). *Animal ethics goes wild: the problem of wild animal suffering and intervention in nature*.

<http://www.tdx.cat/handle/10803/385919>

Fox, M. A. (2006). The Moral Community. In H. LaFollette (Ed.), *Ethics in Practice: An Anthology* (3rd ed.). Malden, Oxford and Carlton: Blackwell Publishing.

Frey, R. G., & Paton, S. W. (2010). *Vivisection, Morals and Medicine: An Exchange*. In H. Kuhse & P. Singer (Eds.), *Bioethics: An Anthology*. Malden, Oxford and Carlton: Blackwell Publishing.

Horta, O. (2014) "The Scope of the Argument from Species Overlap." In *Journal of Applied Philosophy*, 31, pp.142-54.

Kant, I. (1980). *Lectures on Ethics*. Indianapolis and Cambridge: Hackett Publishing Company.

Mannino, A. (2015). "Humanitarian Intervention in Nature: Crucial Questions and Probable Answers." *Relations: Beyond Anthropocentrism* Vol. 3, No 1.

Margalit, A. (2002). *The Ethics of Memory*. Cambridge and London: Harvard University Press.

Margalit, A. (2007). *La Société décente*. Paris: Flammarion.

Margalit, A. (2010). *On Compromise and Rotten Compromises*. Princeton and Oxford: Princeton University Press.

Putnam, H. (2004). *Ethics without Ontology*. Cambridge and London: Harvard University Press.

Singer, P. (2010). All Animals are Equal. In H. Kuhse & P. Singer (Eds.), *Bioethics: An Anthology*. Malden, Oxford and Carlton: Blackwell Publishing.

Singer, P., & Mason, J. (2006). *The Ethics of What We Eat: Why Our Food Choices Matter*. Emmaus, PA: Rodale Books.

Taylor, S. (2017). *Beasts of Burden: Animal and Disability Liberation*. New York and London: The New Press

**8TH BRAGA MEETINGS ON ETHICS
AND POLITICAL PHILOSOPHY,
*WITH GUSTAF ARRHENIUS***



INTRODUCTION

Alexandra Abranches*
alexandra@ilch.uminho.pt

Eze Paez**
ezepaez@ilch.uminho.pt

Nine years ago the members of the Political Theory group at the University of Minho (now the Centre for Ethics, Politics and Society) decided to hold the first edition of the Meetings on Ethics and Political Philosophy. It was conceived as a yearly encounter among researchers of the various disciplines in practical philosophy, fostering an amicable discussion among peers which would attract prestigious scholars and provide young researchers with an opportunity to present their work. So far this initiative has been met with success. Its last edition—the eight, in June 2017—gathered more than 68 participants, selected from the more than 162 abstracts that were received.

This very first issue of the *Ethics, Politics and Society* journal includes a dossier with papers based on some of the presentations delivered during the 8th Meetings. One of the defining features of our conference is the keynote address by at least one philosopher of international renown. Professor Gustaf Arrhenius (Institute for Futures Studies, Stockholm) was one of the keynotes in last year's edition, and he has been so kind so as to contribute to this issue with a paper entitled 'The Democratic Boundary Problem Reconsidered'. In it Professor

* Integrated Researcher at CEPS (Center for Ethics, Politics and Society), University of Minho.

** Integrated Researcher at CEPS (Center for Ethics, Politics and Society), University of Minho.

Arrhenius addresses the problem of determining the criteria by which to identify the members of the people entitled to participate in collective decisions. He defends the All Affected Principle, as well as and democracy as a normative ideal, from important objections that have been levelled against them.

A selection of the speakers, chosen because of the quality of their contributions, were also approached to submit an article based on their presentations. Herein the six articles which received the highest scores are collected. They reflect the diversity of research interests which come together at the Meetings, including metaethics, normative ethics, political philosophy and philosophy of law.

Regarding metaethics, Ashley Lane (University of London) discusses Jacksonian functionalism in “Are moral functionalism’s moral *a priori* commitments really *a priori*?” Moral functionalism was developed by Frank Jackson and Philip Pettit. One of its central claims is that it is *a priori* that a particular descriptive property playing a particular moral role is identical to a particular moral property, even if we can only know *a posteriori* what the actual descriptive property is. Thus, we know *a priori* that the property of moral rightness is whatever plays the moral-rightness role, though whether the actual descriptive property is being an act that maximises utility, one that is universally willable, one not reasonably rejectable or some other altogether different property is something to be discovered *a posteriori*. Adapting an objection made by D. H. Mellor against similar claims in Jacksonian metaphysics, Lane argues that these allegedly *a priori* claims of moral functionalism are actually *a posteriori*, since their truth can only be ascertained through *a posteriori* investigation.

For his part Josh T. U. Cohen (University of Cambridge) delves into feminist normative ethics in “Gender Identities and Feminism”. Cohen is concerned with the rift in feminist philosophy between those accepting non-binary and trans identities, on the one hand, and radical feminism, which denies such identities. The former accept the principle of first person authority (FPA) about gender, allowing for gender self-categorization. The latter resist such principle. They claim that female subjugation is rooted in biology, so that the FPA is conceptually flawed and the political recognition which non-binary and trans people seek may

hinder women's liberation. In order to mend this rift Cohen proposes a conceptual framework (which he calls 'radical FPA feminism') that tries to preserve the FPA and to be compatible with understanding women's oppression as stemming from biology.

Nevertheless, in line with Prof. Arrhenius' paper, the bulk of the contributions deal with problems in political philosophy and philosophy of law. In "Are we post-justification? Stout's case for self-knowledge, political justification and public philosophy", Deven Burks (University of Luxembourg) raises the question of whether self-knowledge is necessary for having justified political beliefs. Burks claims that thick self-knowledge (an agent's knowledge of her own beliefs as well as further beliefs, attitudes, reasons and history at work in their formation) is indeed necessary and identifies this position with Jeffrey Stouts' views on public discourse and public philosophy. Concerned with Brian Leiter's naturalistic critique of political philosophy, he goes on to argue that Stout's position can work within the limits set by Leiter and even complement it.

Daniel Guillery (University College London) makes a contribution to the history of political philosophy with "Hobbes: A Voluntarist About the Possibility of State Enforcement?". In his article he argues that it would be a mistake to read Thomas Hobbes as a voluntarist regarding state enforcement, even if he is a voluntarist about political obligation. Rather, such possibility follows from there being no condition that can render state enforcement impermissible. Guillery contends, however, that this can only be one part of Hobbes's argument for state legitimacy. For the argument to be complete, his scepticism about state-independent morality is also required.

The last piece on political philosophy is penned by Stephen McLeod (University of Liverpool). In "Basic Liberties, the Moral Powers and Workplace Democracy", McLeod elaborates on one of the three Rawlsian arguments (the Fundamental Liberties Argument) for an entitlement to an element of workplace democracy, as discussed by Martin O'Neill. McLeod agrees with O'Neill that this argument should be rejected because, even if it can be modified to withstand O'Neills critique, it is invalid. As an alternative, though inspired by it, McLeod presents in this paper the Argument from Risk to the Moral Powers. According to this view, the exercise of the moral powers (a capacity for a sense of justice and a

capacity for a conception of the good) is severely jeopardised in the absence of economic-democratic entitlements, including the entitlement to a degree of workplace democracy.

Finally, Damiano Simoncelli (University of Genoa-FINO Consortium) completes our dossier with an article on philosophy of law—“From Natural Law to the Golden Rule: Aquinas Revisited”. Simoncelli sets out to reinterpret the Thomistic account of natural law as a form of the golden rule. The author thereby attempts to avoid the traditional misunderstandings associated with the grounding of natural law on human nature and a shared human good. Simoncelli believes his reinterpretation can be fruitful in the development of an intercultural ethics that manages to eschew moral relativism.

This ensemble of contributions from various disciplines makes the first issue of *Ethics, Politics and Society* a fine representative of what the journal has been conceived to be. It is also representative of the sort of rigorous discussion we aspire to in our research Centre.



THE DEMOCRATIC BOUNDARY PROBLEM RECONSIDERED

O PROBLEMA DA DELIMITAÇÃO DEMOCRÁTICA RECONSIDERADO

Gustaf Arrhenius*
gustaf.arrhenius@iffs.se

Abstract. *Who should have a right to take part in which decisions in democratic decision making? This “boundary problem” is a central issue for democracy and is of both practical and theoretical import. If nothing else, all different notions of democracy have one thing in common: a reference to a community of individuals, “a people”, who takes decision in a democratic fashion. However, that a decision is made with a democratic decision method by a certain group of people doesn’t suffice for making the decision democratic or satisfactory from a democratic perspective. The group also has to be the right one. But what makes a group the right one? The criteria by which to identify the members of the people entitled to participate in collective decisions have been surprisingly difficult to pin down. In this paper, I shall revisit some of the problems discussed in my 2005 paper in light of some recent criticism and discussion of my position in the literature, and address a number of new issues.*

Resumo. *Quem deve ter direito a participar em que decisões no processo democrático? Este “problema da delimitação” é uma questão central para a democracia, e tem importância tanto teórica como prática. Todas as diferentes noções de democracia têm, pelo menos, uma coisa em comum: uma referência a uma comunidade de indivíduos, “um povo”, que toma decisões de forma democrática. No entanto, que uma decisão seja tomada de acordo com um método democrático por um determinado grupo de pessoas não é suficiente para que a decisão seja democrática ou satisfatória numa perspectiva democrática. O grupo tem também de ser o grupo certo. Mas o que é que faz com que um grupo seja o grupo certo? Tem sido surpreendentemente difícil determinar os critérios com os quais devemos identificar os membros do povo que têm o direito de participar em decisões colectivas. Neste artigo, irei visitar alguns dos problemas discutidos no meu artigo de 2005 à luz de críticas recentes e da discussão da minha posição na literatura, e abordarei algumas questões novas.*

* Institute for Futures Studies; <http://www.iffs.se/forskning/forskare/gustaf-arrhenius>.

o. Introduction

Who should have a right to take part in which decisions in democratic decision making? This “boundary problem” is a central issue for democracy and is of both practical and theoretical import.¹ If nothing else, all different notions of democracy have one thing in common: a reference to a community of individuals, “a people”, who takes decision in a democratic fashion. However, that a decision is made with a democratic decision method by a certain group of people doesn’t suffice for making the decision democratic or satisfactory from a democratic perspective. The group also has to be the right one. But what makes a group the right one? The criteria by which to identify the members of the people entitled to participate in collective decisions have been surprisingly difficult to pin down. Resolving the boundary problem raises a number of theoretical problems, some, it seems, quite intractable.

Although the boundary problem is a fundamental issue in democratic theory, surprisingly little attention has been given to it in the classical canonical treatises on democracy. As Robert Dahl put it in the seventies, “how to decide who legitimately make up ‘the people’ (...) and hence are entitled to govern themselves (...) is a problem almost totally neglected by all the great political philosophers who write about democracy” (Dahl 1970, 60). Cf. (Dahl 1989, 119ff), which is rather surprising. Actually, very little had been written about this topic until Dahl’s own work on it and Frederick G. Whelan pioneering paper in the eighties. In the last ten years or so, however, there has been a significant and welcome improvement and there is now a burgeoning literature in the area. In this paper, I shall revisit some of the problems discussed in my (Arrhenius 2005) paper in light of some criticism and discussion of my position in the new literature on the boundary problem and address a number of new issues.

¹ (Robert Dahl 1989) refers to this problem as “the problem of the unit” (p. 193), “the problem of inclusion” (p. 119), and sometimes as the “boundary problem” (pp. 146-7). (Robert Goodin 2007) calls it “the problem of ‘constituting the demos’” and others have referred to it as the “demos problem”. Frederick G. Whelan calls it “the boundary problem” in his (1983) pioneering article on the subject, and so shall I. (Goodin 2007 fn. 1), thinks that “calling it ‘the boundary problem’ makes the issue seem more a matter of geography than it necessarily is”. As should be clear below, I don’t conceive of the boundary problem as a matter of geography and I use the term “boundary” in its general sense, like, for example (Frege 1970, 159), when he talks about concepts having “a sharp boundary” or as it is defined in Oxford English Dictionary: “That which serves to indicate the bounds or limits of anything whether material or immaterial; also the limit itself”.

1. Practical and Theoretical Boundary Problems

In difference from many philosophical problems, the boundary problem is clearly also a pressing practical political problem. For example, what is the relevant constituency for a democratic solution to the Northern Ireland conflict?² Should a treaty be approved by the citizens (or their representatives) of Northern Ireland alone or should it involve those of the United Kingdom and the Irish Republic as well? The latest treaty — “the Good Friday Agreement” — was subject to a referendum in Northern Ireland and the Republic of Ireland whereas the citizens of Great Britain were represented by their government. It is hardly a solution acceptable for an old-style Unionist, since she would prefer a referendum in United Kingdom of Great Britain and Northern Ireland or perhaps only in Northern Ireland.³ Yet, such a referendum would not impress an Irish nationalist who would consider these boundaries arbitrary and illegitimate, nothing more than a kind of international gerrymandering. Still, both the Unionist and the Irish nationalist could be dedicated democrats in the sense that they think that a fair solution should be based on a democratic referendum.⁴

The proposed secession of Quebec from Canada, Scotland from UK, UK from EU, Catalonia from Spain, Kosovo from Serbia-Montenegro, and the like, raise similar problems.⁵ The boundary problem also arises in the context of migration. For instance, consider the dilemmas confronted by host states. What voting rights, if any, should non-citizens have? Should they enjoy these rights only at certain levels—say, local rather than national elections—or only over certain issues (Beckman 2006)? Analogous questions arise from the perspective of migrants’ countries of origin (Grace 2003; Bauböck 2006, 2007; Rubio-Marín 2006; López-Guerra 2005, 2014). Do long-term emigrant diasporas in Europe,

² (Whelan 1983, 23) discusses this example at length.

³ It is telling that in the referendum about the treaty, an estimated 96% of the Catholics supported it whereas only 52% of the Protestants gave it its blessing. See (Encyclopædia Britannica 2014)

⁴ And what of the millions of Irish who left the island, in part because of the conflict?

⁵ It is interesting albeit worrisome to note how rapidly the number of practical instances of the boundary problem have increased in the short time span since my (Arrhenius 2005) paper.

for example Turkish and Kurdish communities in Germany, have a right to participate in democratic decision-making in their homelands?

The boundary problems also appear in more humdrum small-scale cases. Consider for instance the local referendum about congestion charges in Stockholm. Was it right that only people living in the city of Stockholm had a vote? One might claim, as many did, that the inhabitants in the surrounding suburbs who on a regular basis commute to central Stockholm also should have had the vote.

It is perhaps not equally clear that the boundary problem is an issue in the justification and legitimacy of democracy but consider how questions like the above should be decided democratically. One suggestion could be to have a referendum about who should have say in these questions. But who should be allowed to take part in such a referendum? And so on without any end, we seem to end up in an infinite regress. This chain of reasoning has led some to draw quite gloomy conclusions regarding both the ability of democratic theory to solve the boundary problem in a satisfactory manner and the scope and the legitimacy of democratic decision-making. In his pioneering paper on the boundary problem, Frederick G. Whelan, concluded that:

(...) democratic theory *cannot itself* provide any solution to disputes that may – and historically do – arise concerning boundaries. (...) It may not be surprising that democracy, which is a method for group decision-making or self-governance, cannot be brought to bear on the logically prior matter of the constitution of the group, the existence of which it presupposes. Nevertheless, strong claims are frequently made for democracy, both by its philosophical advocates and by ideologues and activists of the modern world; democracy is commonly put forward as the sole foundation of legitimate government, and as the sole legitimate method to make binding public decisions of all sorts. (...) The boundary problem does, however, *reveal one of the limits of the applicability of democracy*, and acknowledgement of this may have the beneficial effect of moderating the sometimes excessive claims that are made in its name.⁶

Likewise, Dahl stresses that “we cannot solve the problem of the proper scope and domain of democratic units from within democratic theory” and that “in solving this particular problem [the problem of constituting the people] democratic theory cannot take us very far. Democratic ideas, as I have said, do

⁶ (Whelan 1983, 40,42 my emphasis). (Näsström 2003, 2004) repeats the same claim.

not yield a definitive answer” and it cannot be solved even by “reasoned inferences from democratic principles and practices”.⁷

Although I don’t agree with Whelan’s and Dahl’s gloomy conclusions, as I shall discuss below, I indeed agree that the boundary problem reveals a problem at the heart of the very idea of democracy. That a decision is made with a democratic decision method by a certain group of people (or by an elected assembly that represents the group) doesn’t suffice for making the decision democratic or satisfactory from a democratic perspective. The group also has to be the right one. But what makes a group the right one?

Actually, the problem I have referred to above as “the” boundary problem is just one among a number of such analogous problems. The perhaps most discussed boundary problem concerns people’s capabilities as political agents, their political competence. In order to effectively further one’s interests through democratic processes one must, arguably, possess a certain degree of knowledge and rationality. The question then becomes how we should decide the relevant political competence for membership in the political community. In practice, this minimal competence is often approximated by age, but one could ask which age is the appropriate one, and whether the same age limit is appropriate for all democratic processes, e.g., parliamentary elections versus local elections. Moreover, we could not only ask who should have a right to vote but also who should have a right to run for office, and whether these rights go together (historically, this has often not been the case). Another boundary problem concerns beings that lack the capacity to take part in the democratic process but who are going to be affected by policies adopted and that could be represented by proxies, for example minors, future generations, and animals.⁸ Members of the two first groups are likely to have the same political competence as present people, but they cannot take part in the democratic process since they are simply not around.

⁷ (Dahl 1989, 207, 209) In Dahl’s terminology, the “scope” of a democratic unit is the set of matters that are to be decided by it, and the “domain” is the set of persons who comprise it. Cf. (Barry 1991).

⁸ For a discussion of future generations in connection with the boundary problem, see (Tännsjö 2005), (Bergström 2005), and (Arrhenius 2015).

I shall not discuss these problems here, but my suggested solution to Whelan's boundary problem --- Which people, given that they are politically competent and can take part, ought to have a right to take part in which decision-making processes? ---- has clear implications for how we should approach these boundary problems too.⁹

2. Schumpeter and Ross: A Historical Detour

Before turning to the discussion of a general answer to the boundary problem, it will be useful to look at two classical definitions of democracy. As mentioned by Dahl in the quote above, the boundary problem has been almost ignored by the great theorist of democracy in the past. A case in point is Joseph Schumpeter's revisionist definition of democracy:

The democratic method is that institutional arrangement for arriving at political decisions in which individuals acquire the power to decide by means of a competitive struggle for the people's vote. (Schumpeter 1976, 269).

Schumpeter is here discussing democratic governance of states and developed his definition in analogy with firms in a capitalist market: Firms compete with others to sell products at a profitable prize and only the most competitive firms survive. As he writes elsewhere "(...) we have restricted the kind of competition for leadership which is to define democracy, to free competition for a free vote. --- Free, that is, in the same sense in which everyone is free to start another textile mill" (Schumpeter 1976, 271 and 272, fn. 6). He developed his definition by looking at states that people called "democratic" and extracted what he thought they had in common. Hence, his definition is based on a denotation (extension) analysis of the term "democracy" as it was used in his time and environment and as such, I surmise, quite accurate.¹⁰

⁹ For reasons that will become clearer below, I also think that the answer to problem of what kind of majority --- simple, qualified, or unanimity --- that should be required to pass, abolish, or amend certain laws, and of what kind of voting system that should be used -- direct democracy, representational democracy, plurality voting, Borda count etc --- is analogous to the answer I'm going to give below to the boundary problem.

¹⁰ As Julia Mosquera pointed out to me, there is a kind of epistemic boundary problem lurking here since one can wonder whose language use is taken into account in Schumpeter's denotation analysis.

Notice, however, that nothing is said about the boundary problem, that is, who has a vote or for whose vote one can compete to acquire power. Schumpeter's denotation analysis and definition are remarkably incomplete in this respect but also, I'm afraid, representative of the definitions that have been proposed in the literature.¹¹

One might think that Alf Ross' definition is an exception. He defines an "ideal type" of democracy with three dimensions that can be fulfilled to varying degrees:

1. [I]ntensity, that is with respect to the *size of the population of people* who are allowed to take part in referenda and elections. ---

2. [E]fficiency, that is, with respect to the effectiveness of the popular will in deciding issues. ---

3. [E]xtensity, that is, with respect to the scope of popular influence and control over the *different branches of government*.¹²

Here, one might think that the clause regarding Intensity answers the boundary problem: The more people that are allowed to take part in a vote, the greater the degree of democracy. However, this is not exactly what Ross had in mind since he has presupposed a constitution of a "people" and Intensity is the percentage of the "people" who are (legally) allowed to take part in a vote. As he writes just before the passage quoted above: "(...) the people's influence on the exercise of public authority, can vary with respect to(...)".¹³

So at least with respect to Schumpeter and Ross, Dahl is right that the boundary problem has been ignored in the canons of democratic theory. Let's

¹¹ The same holds for (Schumpeter 1976, 250) statement of the "classical" definition of democracy: "The democratic method is that institutional arrangement for arriving at political decisions which realizes the common good by making the people itself decide issues through the election of individuals who are to assemble in order to carry out its [the people's] will". Again, nothing is said about what constitutes the people.

¹² See (Ross 1968, 101–2), my translation, italics in original.

¹³ (Ross 1968, 101), my translation. Still, Ross definition seems to have something to say about the boundary problem inside a demos and thus it might have interesting implications regarding, for example, economic democracy. For a discussion, see (Arrhenius 2012, 2017).

now turn a possible answer to the boundary problem and its implication for how we should conceptualise democracy.

3. Democracy as a Normative Ideal or Decision Method

It will be useful to consider an important distinction between two ways of understanding democracy which unfortunately hasn't been observed sufficiently in the discussion. In general, we should distinguish between normative ideals, on the one hand, and practical decision methods or rules for regulating social interactions (e.g., social norms, laws, institutions), on the other hand.¹⁴ Roughly, a normative ideal states the ultimate goal that that we strive towards, such as the just or good society (i.e., the considerations that ultimately make actions, policies, institutions etc, right, just, or fair), whereas a decision method is a strategy for decision-making which we use to achieve the goal specified by the ideal. We use the normative ideal, in conjunction with empirical considerations (e.g., economical and psychological facts), to evaluate and rank alternative decision methods, social norms, laws, institutions, etc., for different situations and contexts, in respect to how well they would promote the ideal. In that sense, the application of a certain decision method is justified by our normative ideal whereas the ideal is justified by being in accordance with our considered normative judgments and by satisfying other relevant epistemological and methodological criteria.¹⁵

Take utilitarianism as an example. According to utilitarian ideal, we should maximise people's well-being, or expected well-being. Now, a common complaint against utilitarianism contends that it is self-defeating since in many instances it is practically impossible to calculate the value of the outcomes of the alternative actions available to a person. Consequently, if we try to apply the utilitarian principle in every single case, we are likely to choose the wrong action since our calculations are bound to be wrong. This is, however, no argument against

¹⁴ For a related distinction between criteria of rightness and decision methods, see (Bales 1971) and (Danielsson 1974, 28–29). Danielsson and (Tännsjö 1992) make the distinction in connection with democratic theory. See also (Brink 1986, 421–27); (Kymlicka 1990, 29).

¹⁵ For a discussion, see e.g., (Rawls 1971) and (Tersman 1993).

utilitarianism as a normative ideal but an argument against utilitarianism as a practical decision method. One can still accept utilitarianism as a normative ideal but hold that in practical deliberation, we have to rely on “rules-of-thumb” and approximations: help those that are in distress, be honest, do not break promises, obey the law, etc., on the individual level, and rule of law, democracy, independent mass media, and so forth, on the institutional level. Whether we should accept these practical decision methods, on the other hand, have to be judged against the utilitarian ideal in a “cool hour” when we have enough time and resources to evaluate the consequences of the general application and implementation of these decision methods relative to the goal specified by the ideal. Likewise for an advocate of equality of resources, well-being, power, etc, and other normative ideals.

There are two important lessons to draw from this example. Firstly, one can reject a theory as a decision method but still accept it as a normative ideal and *vice versa*. Secondly, even if we don't find a particular decision method satisfactory in regard to some case, it doesn't follow that it isn't useful in other cases. The utilitarian decision method, for example, might be a good one for some governmental bodies. Again, this has to be decided by evaluating the decision method against the ideal. A normative ideal, on the other hand, we expect to be applicable to any actions, rules, or institutions, that falls under its domain without exceptions.

What does this distinction have to do with democracy? A theory of democracy can also be taken either as a normative ideal or as a practical decision method. As R. J. Pennock puts it succinctly in a discussion of Wollheim's paradox: “One must distinguish at the outset between democracy as an ideal and democracy as a practical device for approximating the ideal” (Pennock 1974, 88).

If one wants to defend democracy as a normative ideal, for example as an idea about fair distribution of power, then one needs to show that it is in a reflective equilibrium with our considered judgments about democracy and power. Such a normative ideal need not be directly applicable to choice-situations in the real world. Rather, we use the normative ideal, in conjunction with empirical considerations, to evaluate and rank alternative practical decision methods for different situations and contexts.

For those who study how democracy works in practice, it is probably more common to view democracy as a kind of decision method, as a matter of institutional arrangements. Schumpeter is a case in point. Implicit in much reasoning about democracy, however, is also the idea that democracy is a kind of normative ideal. For example, it is presumed in many lofty political declarations, often expressed in terms of justice and equality.¹⁶ Although I shall not dwell much on the details of such a theory in this paper, I think that the most promising approach is to take democracy as a normative ideal concerning fair distribution of power.¹⁷

Let me here take the opportunity to point out that there is an ambiguity in the discussion of the boundary problem.¹⁸ As we have formulated the problem, it concerns who should have a right to take part in different decisions. This can be interpreted in at least three ways. On one interpretation, it is about who should have a right to take part, *all things considered*, that is, when we have taken into account all relevant moral and political aspects (efficiency, prosperity, freedom, equality, etc.). Complete normative ideals such as utilitarianism and Rawlsian liberalism answer this question. On a second interpretation, it is about who should have a right to take part given that other important normative ideals are not compromised too much. Here, the answer to the boundary problem yields a normative pro tanto reason for including and excluding people in a democratic decision process.

Both of the above interpretations take the boundary problem as a normative problem. On a third interpretation, the boundary problem concerns the conceptual question regarding who should have a right to take part for a decision procedure to be *democratic* or *more democratic* than another procedure.¹⁹ Ross' theory discussed above is an example of a partial answer (inside a given demos) to this interpretation of the boundary problem. Here we are discussing

¹⁶ See (Næss, Christophersen, and Kvalø 1956) for a list of such slogans.

¹⁷ The best developed version of this kind of ideal in the literature is (Brighthouse and Fleurbaey 2006). Two other examples are Danielsson's (1974) suggestion to take problems of preference aggregation, such as Arrow's impossibility theorem, as problems of just distribution of influence, and Christiano's (1996, 2002) theory of democracy as an ideal of equal chances to affect the outcome.

¹⁸ Including my (Arrhenius 2005).

¹⁹ For similar distinctions, see (Beckman 2009).

democracy not as a normative ideal but in terms of a non-normative ideal type of democracy (just as we can give ideal type definition of a circle). In itself, an answer to this problem has no normative implications since it says nothing about who ought to be given a say, all things considered.

A connection between the two interpretations is often presumed, however, by an implicit normative premise according to which a decision ought to be taken as democratic as possible given that other important values would not be too compromised. Democracy is thus understood as a partial normative ideal that must be weighed against other partial normative ideals to yield an answer to the problem of who should have a right to take part in a decision, all things considered.

Unless otherwise indicated, I shall take the boundary problem in the latter way below. The answers to this problem will thus specify who should have a right to take part in a certain decision in order to make it more democratic, but also who ought to have a say given that other important normative ideals are not compromised too much. For reason of space, I have to leave the interesting question of how to weigh the democratic ideal against other ideals to another time.

4. Democracy as Decision Method and the Boundary Problem

Does Whelan discuss democracy as a decision method or as a normative ideal? Whelan's position is unclear on this issue. He sometimes writes like he has a normative ideal in mind, for example, when he talks about democracy as "the sole foundation of legitimate government" (Whelan 1983, 40). Someone who takes democracy as a practical decision method justified by a normative ideal doesn't hold that democracy is the foundation of legitimate government but that its legitimacy derives from the normative ideal. For utilitarians, for example, democracy (of some kind) is justified if and only if it maximises people's well-being as compared to alternative decision methods. For Rawlsian liberals, to take another example, democracy is justified (roughly) if it is the best decision procedure for the safeguarding of basic civil liberties, equal opportunity and the well-being of the worst-off. For Nozickian libertarians, democracy is justified

insofar it respects people's property-rights, and so forth for other normative ideals.

On the other hand, much of Whelan's writing seems to concern democracy as a decision method. For instance, he spends quite a lot of space on discussing boundary principles based on territory, nationality, culture, or geography. These principles are quite obviously poor candidates for a boundary principle seen as part of a democratic ideal.

The territorial state principle is an illustrative example. According to this principle, we should just take existing territorial states as a given and include every person residing or born in a particular state in the democratic process governing that state. As Whelan himself points out, the territorial state principle cannot give any guidance in cases where borders of territorial states are in question. It falsely assumes that we can take boundaries between territorial states as something already fixed and undisputed. Needless to say, territorial disputes are frequent in human history and some of the most tragic and bloody conflicts in the present revolve around the issue of establishing boundaries where no entrenched territorial boundaries exist.

Secondly, the territorial state principle has a very limited scope. Arguably, any social union, from the world community to the family, is part of the domain of democracy, i.e., are candidates for being democratically organised.²⁰ The territorial state principle only addresses one particular boundary problem and leaves open the question of how to delimit participation in decision making procedures in other social unions.

Thirdly, even if we lived in the best of possible worlds where all territorial boundaries were settled, these boundaries would still be irrelevant and ad-hoc from a normative point of view. Suppose the U.S. Government decides to resume atmospheric nuclear tests and predicts that fallout would cause several deaths and injuries. The test would either be performed above the Nevada desert, where the fallout will only affect U.S. citizens, or next to the Mexican border where, because of wind conditions, it would only affect Mexican citizens. As good democrats, the U.S. Government arranges a referendum in the U.S. where, not

²⁰ See (Cunningham 1987, 51), for the same view.

surprisingly, the vast majority of U.S. citizens votes for the Mexican border alternative. According to the territorial state principle, this would be democratically impeccable decision relative to the boundary problem.

The above example is of course just fictional²¹ but it is easy to find analogous real cases. One example is the nuclear plant Barsebäck on the south coast of Sweden just across from Copenhagen. The choice of location for Barsebäck was under the democratic control of the Swedish people through their elected representatives in the parliament whereas the Danes didn't have hardly any democratic influence over the choice.²² The same holds for the regulations regarding the maintenance of the plant and the eventual decision to close it down. Yet, the Danes would, arguably, have had to carry most of the burden had a serious accident happened.²³ Given the territorial state principle, this is no problem from the perspective of democracy, a view few Danes would agree with, I surmise.

As these examples show, the main flaw of the territorial state criterion is that it is completely insensitive to who is affected by a decision. This will not suffice in a world where pollution, goods, and capital move more or less freely over state borders.

The three other boundary principles mentioned above which Whelan discusses --- nationality, culture and geography --- share the same flaws as the territorial state principle. Since many of these problems are rather obvious, it is hard to believe that anyone has seriously suggested them as a boundary principle for a democracy as a normative ideal and Whelan himself doesn't state clearly what he has in mind. However, Whelan's discussion of this topic becomes more interesting if we see it as a criticism of certain rules of thumbs for who should take part in which decision, that is, as partial boundary principles for democracy as a decision method. We can then see his criticism as an effort to point out when

²¹ When this paper was written. Given the recent political development in US foreign politics toward Mexico, one might fear that the example will not stay fictional forever, or at least that some similar cases will appear.

²² Admittedly, one can argue that they had some influence since the Danes had the opportunity to lobby the Swedes by taking advantage of Swedish freedom of speech laws and the like. This highlights the important question for democracy of how to regulate and facilitate lobbying. I'm grateful to Shlomi Segall for pressing this issue.

²³ Within a zone of 40 kilometres from Barsebäck, all of Copenhagen is covered (and also the Swedish cities Helsingborg, Landskrona, Lund, and Malmö).

these rules are useful and when they are not. Moreover, I don't doubt Whelan's claim that these principles are commonly discussed among political scientists, geographers, and others, but it is quite likely that what these theorists had in mind was a useful rule of thumb for certain cases, not a general principle applicable in all cases.

There is, however, a problem for Whelan if his discussion is supposed to be about democracy as a decision method: his main charge against democracy loses its force. If we take democracy as a practical decision method and again raise the question of who should take part in which decision procedures, it seems clear that this has to be answered by the normative ideal that motivated the choice of democracy in the first place, in conjunction with empirical considerations.²⁴ Again, if our normative ideal is utilitarianism, then the allotment of voting rights and the scope of democratic decision-making should be devised such that the total welfare is maximised, and likewise, *mutatis mutandis*, for other possible ideals. In other words, it is a misplaced criticism of democracy as a decision method that it “cannot itself provide any solution to disputes that may – and historically do – arise concerning boundaries” since it never was supposed to do it by itself. (Whelan 1983, 40). Consequently, for Whelan's challenge to have any bite, we have to take it to be about democracy as a normative ideal.

5. Democracy as a Normative Ideal and the All Affected Principle

An intuitively attractive boundary criterion for a democracy as a normative ideal is the *All Affected Principle*: The people that are relevantly affected by a decision ought to have, in some sense and to varying degrees depending on how much they are affected by it, influence over the decision.²⁵ I think it is fair to say that it is implicit in much reasoning in the democratic tradition and many (most?) the contemporary democratic theorists who explicitly discuss the boundary problem endorse some version of this principle: “Everyone who is affected by the

²⁴ Cf. (Dahl 1989), (Barry 1991).

²⁵ I'm here assuming that the All Affected Principle will answer both the question of whom should have influence and how much influence they should have. One could take the principle in a less expansive way to just answer the former question and leave it to another principle to answer the latter one. As we shall see below, however, much criticism of the All Affected Principle assumes the more expansive version. Thanks to Krister Bykvist for pressing this issue.

decisions of a government should have the right to participate in that government” (Dahl); “In a perfect democracy all who are thus affected [by a decision] play some part” (Cohen); “[A]ll affected interests should have a say” (Goodin); “Power in any decision-making process should be proportional to individual stakes” (Brighouse & Fleurbaey).²⁶

It is easy to garner intuitive support for All Affected Principle. We don’t think that the curriculum imposed by the School board of Waco, Texas, is any business of Icelanders since they are not relevantly affected by this decision. Likewise, people in Luleå (far up north in Sweden) should not, in most cases, have much of a say on how the public transportation is organized in Stockholm, e.g., whether to increase the number of buses to a certain suburb. However, what kind of hair spray the teachers use in Waco might be the business of Icelanders too, i.e., if the hair spray used destroys the ozone-layer. Similarly, whether state tax revenue should be used to subsidise the public transportation system in Stockholm is arguably an issue that the people in Luleå, qua taxpayers, should have some form of influence over.

According to the All Affected Principle, how much power you ought to have over an issue depends on how much your interests are at stake. In actual democratic practices we approximate this standard by having different issues handled on different levels: councils, provinces, regions, states, European, and so forth. The subsidiarity principle, frequently invoked in the discussion of decision making in the European Union, is in one of its popular interpretations --- “decisions should be taken as closely as possible to the citizen” --- very much along the lines of All Affected Principle. The general prescription of this principle is that an issue should be handled by the democratically run body that represents

²⁶ (Dahl 1970, 64); Cohen (1971 p. 8); (Goodin 2007, 50); (Brighouse and Fleurbaey 2010, 2). To the best of my knowledge, the first formulations of the All Affected Principle are by Robert Dahl and Carl Cohen. It has been formulated in different ways, both by its advocates and by those opposing it. Frederick G. Whelan, in his influential paper on the boundary problem, defines the All Affected Principle as “all those people who are affected by a particular law, policy, or decision *ought to have a voice* in making it” (Whelan 1983, 16). Ian Shapiro suggests that “[e]veryone affected by the operation of a particular domain of civil society should be *presumed to have a say* in its governance”; Lars Bergström claims that “the all-affected principle ... says that every individual who is affected by a given decision *should have a vote*”; and Torbjörn Tännsjö renders the All Affected Principle as “[e]veryone who is affected by a decision *should be allowed to take part* in it” (Shapiro 1996, 232); (Bergström 2009, 1); (Tännsjö 2007, 5). See also (Cunningham 1994, 147) and (Cunningham 1987, 25–26). My emphasis in all the quotes.

the social union that best approximates the set of relevantly affected people relative to the type of issue.

One reason why many people agree with the All Affected Principle is, of course, that it is quite imprecisely formulated and thus open to many interpretations and precisifications. As others and I have stated it, it doesn't say anything about what amounts to being relevantly affected, or what it means to have influence over a decision, and to what degree one should have influence given that one is affected in a certain way. To forestall some possible misunderstandings, let me just hint at what I think an analysis of these concepts would and would not look like.

Just as in the discussion of welfare and equality, we need to develop a measure or index of what should count as being relevantly affected by a decision by consulting our considered judgements about which effects on people's lives are of such significance that they should have a say in a decision, and how much influence they should be assigned. The idea is that one is relevantly affected by a decision if and only if one is made better or worse off relative to the measure by the different possible outcomes of the choice situation, that is, the difference between the values of the outcomes for the individual given the index.

Such a theory would in many respects be similar to the theories of welfare that have been suggested in the discussion of utilitarianism and to the theories regarding "the currency of egalitarian justice" suggested in the discussion of how a just society would look like.²⁷ One might also think that one could just import a currency from these areas, such as Rawls' "primary goods" or Sen's "capabilities", as an explication of "relevantly affected". This is suggested by Brighouse and Fleurbaey and an advantage with this approach is that it might bring democratic decision making more in line with what is good from the perspective of justice and morality.²⁸ However, our judgment about when people are affected by a decision in such a way that they should have some influence over it may be different in many respects from our judgment about when people's well-being is

²⁷ See e.g., (Rawls 1971); (J. Cohen 1989), (Dworkin 1981b, 1981c, 2002); (Sen 1985, 1992).

²⁸ (Brighouse and Fleurbaey 2010, 15). Roughly, if people vote in accordance with what is good for them from the perspective of the metric of social justice, then the winning alternative will also be the one that maximises social justice.

affected, or about the relevant goods for the state to distribute in an egalitarian fashion. The example of “nosy preferences” is a case in point: Even if I am so disgusted by the lewd literature that you read, or by your choice of bedroom activities, that my well-being is seriously at stake, it still seems that I shouldn’t have any power over you in regards to such activities. Rather, you (and your partner if one is needed) should have all the power to decide such issues.²⁹ Similar considerations are likely to appear for other currencies developed for a different context, I surmise.

A quite popular suggestion, usually presented as an alternative to the All Affected Principle, is that those who are legally bound by the laws should have the right to take part in making the laws.³⁰ This might very well be a better exegesis of the common phrase “government by the governed” or, as Lincoln once expressed it, “A government of the people by the same people”.³¹ It might also be more in line with how we historically have thought about democratic governance.

The scope of the “Legally Bound” or “All Subjected Principle” is quite unclear, however. A person who spends a fortnight in South Africa every year is arguably legally bound by the laws of South Africa, at least during the time she is in the country. Does that mean that she should have some kind of influence on the South African elections according to the All Subjected Principle?

It is sometimes suggested that the All Subjected Principle will keep voting rights and other democratic influence roughly along the lines of current democratic practices, or at least extend it less widely and counterintuitively as compared to the All Affected Principle.³² As the above example indicates, this is not clear but depends on how we spell out “legally bound”. Actually, on a natural reading of what it means to be bound by or subjected to a law, the All Subjected Principle entails that we should include everyone. On this reading, you’re bound by a law if you are liable to prosecution were you to violate the law. For example, I’m bound by the law in Sweden to wear seat belts whenever traveling in a car

²⁹ See e.g., (Sen 1970); (Dworkin 1981a, 2000).

³⁰ See (Miller 2009); (Beckman 2006, 2009, 2014); (Tännsjö 1992); (Owen 2012); (López-Guerra 2005); and (Dahl 1989). For an extensive discussion of the All Subjected Principle, see (Goodin 2016).

³¹ Lincoln in Message to Congress, 1861, quoted from (Næss, Christophersen, and Kvalø 1956, 285).

³² See e.g., (Miller 2009, 224).

even if I actually never go by car since if I were to go by car and not wear a seat belt, I would be liable to prosecution. Hence, all of us, irrespective of where we live, are bound by the laws in South Africa since were we to violate them, by going to South Africa and doing something against the law, we are indeed liable to prosecution.

Moreover, the circle of people subject to legal duties doesn't always correspond to the territorial jurisdiction of the state, as illustrated where the law includes provisions of "universal jurisdiction".³³ In addition, there is a distinction to be drawn between being subject to legal duties and being subject to coercive institutions enforcing the law. These don't always coincide as is illustrated by cases where people are beyond the reach of public authorities and yet subject to the law.³⁴

The point here is that the All Subjected Principle also needs an explication of relevantly affected although in terms of *relevantly legally affected*.³⁵ Rather than taking the All Subjected Principle as an alternative to the All Affected Principle, I suggest that it is more fruitful to see it as a version of the latter but with a specific currency, namely being relevantly legally affected (to different degrees).

This is analogous to the different version of other normative principles. Take, for example, Utilitarianism. The same formal principle ("An action is right if and only if it maximises welfare") can be combined with different conceptions of welfare to yield different versions of Utilitarianism: Hedonistic Utilitarianism, Preference Utilitarianism, etc. And just as there are different versions of hedonism, which yield even more versions of Utilitarianism, there will be many different versions of legally affected, yielding different versions of the All Subjected Principle.

Most importantly, the right currency of the All Affected Principle might in the final analysis turn out to be a quite complicated combination of different

³³ See (Goodin 2016) for an extensive discussion of this issue.

³⁴ A recent suggestion is that the All Subjected Principle should be interpreted as including both requirements; i.e., a person is subjected in the relevant sense if and only if the person is both subject to legal duties and coercive institutions (Beckman 2014).

³⁵ Such an explication would also consider whether people can be more or less legally affected by different laws.

aspects including being affected in certain legal *and* non-legal ways (e.g., affected wellbeing). This important possibility is obscured if they take the All Subjected Principle as an alternative competitor to the All Affected Principle.³⁶

Another fundamental question for the All Affected Principle is when can we say that a person has had correct influence over a decision? How to analyse “influence over a decision” is a tricky question which needs its own essay. A starting point, however, could be to analyse it in terms of whether an individual’s preferences ordering could determine the collective ordering in some possible situations, what I have called “potential influence” and which the standard measures of voting power measures. You have potential influence on a decision if there is a possible situation (i.e., a possible set of individual preference orderings or voting patterns of the involved people) where you are decisive, that is, where your preference or vote will determine the outcome.³⁷ In addition, we should also consider an individual’s probable and actual influence. An individual’s probable influence in a situation is the probability of her being decisive whereas an individual’s actual influence given a number of issues is the number of times she is decisive divided with the number of decided issues (Arrhenius 2008a, 2018). Such an analysis has to be supplemented with an analysis of an individual’s influence on other peoples’ preferences and beliefs, and her influence on the agenda. For example, a person may have great influence on a decision by just being the kind of person that many people trust (e.g., an expert or a charismatic leader), or by having influence over what issues that are discussed in the mass media, or by having control over which alternatives that are on the voting agenda.³⁸

We would then have to consider what kind and degree of influence that should be given to an individual depending on how she is relevantly affected. This

³⁶ There might be interpretations of “legally bound” which would be somewhat counterintuitive to subsume under “relevantly affected”, for example being bound by an unenforced or even unenforceable law (I’m grateful to Bob Goodin for suggesting this possibility). However, my guess is that any reasonable explication of “relevantly legally affected” would rule out such “effects” as irrelevant for giving people a say in a decision.

³⁷ Measures of potential influence were first proposed by (Penrose 1946), (Shapley and Shubik 1954), and (Banzhaf 1965, 1966, 1968) (see (Felsenthal and Machover 1998) for an overview). See also (Danielsson 1974), (Goldman 1974), (Morris 1987). I discuss these kinds of measures further in my (2008, 2018) papers.

³⁸ See (Arrhenius 2008a, 2018). Research on the increasing influence of lobbyists and so-called “policy professionals” in Western democracies are highly relevant here. See e.g., (Svallfors 2016; Svallfors and Tyllström 2018).

can vary, a point that is often overlooked in the discussion of the All Affected Principle (more on this below). Sometimes it could be a vote (perhaps with differential weights), sometimes a veto, sometimes only a right to participate in the deliberation or the right to put forward proposals, sometimes a combination of these and all the other possible ways of having influence over a decision.

Again, the All Affected Principle will have very different implications depending on what explication of “relevantly affected” and “influence” with which it is coupled. This is a rather obvious but important point since many arguments against the All Affected Principle is actually arguments about how “relevantly affected” and “influence” should be understood (the discussion of the All Subjected Principle above is a case in point). Hence, much of the criticism of the All Affected Principle misses in this sense its target. Let us now turn to a couple of such examples .

6. Nozick’s suitors and Bergström’s Complaint

Robert Nozick argues against the principle that “people have a right to a say in the decision that importantly affect their lives” with a number of examples where the principle purportedly gives the wrong answer (Nozick 1974, 268–69). Here is one:

If four men propose marriage to a woman, her decision about whom ... to marry importantly affects each of the lives of those four persons, her own life, and the lives of any other person wishing to marry one of these four men, and so on. Would anyone propose, even limiting the group to include only the primary parties, that all five persons vote to decide whom she shall marry?³⁹

Nozick answer is a resounding “no” and I think most people would agree. This will not, however, worry a proponent of the All Affected Principle. She can happily agree with Nozick and argue that when it comes to such vital interests as to whom to marry, the individual should have a veto right. On the other hand, one might argue that normally the four suitors have a right to try to influence the

³⁹ (Nozick 1974, 269). Nozick (p. 269) also gives the following example: “Does Thidwick, the Big-Hearted Moose, have to abide by the vote of all the animals living in his antlers that he not go across the lake to an area in which food is more plentiful?”. I don’t find this example very counterintuitive, at least not if we suppose that the animals living in Thidwick’s antlers are conscious mentally competent beings and that it is a life-and-death question for them but just a matter of greener grass for Thidwick. In that case, it seems reasonable that the animals should have not only a vote but perhaps also a veto right against Thidwick’s proposed course of action.

decision in the sense that they may present their case, send flowers and poems, etc., as a corollary of some form of freedom of speech. And the woman in Nozick's case is of course also free to just ignore such courting and can demand and demand that the suitors stop with their efforts, and so forth.

In other words, to think that Nozick's example is a decisive argument against the All Affected Principle is to make a triple mistake. Firstly, the All Affected Principle is flexible since it can be coupled with different notions of "relevantly affected", and, secondly, "having a say" ("having influence") need not to be equated with voting rights, and, thirdly, one can give people different degrees and kinds of influence relative to how they are affected by a decision.

I would say the same thing about the alleged counterexamples presented by Lars Bergström that evidently is meant to show that the All Affected Principle "is not very plausible" (Bergström 2006, 7). For example, he writes that

this can be seen if one tries to apply it to individual agents. Everyone makes a lot of decisions that affect the interests of other people. But hardly anyone would maintain that all these decisions should instead be taken collectively by all the people that are affected by them. This is not only impossible in practice. It is not even a desirable ideal that one should try to realize as far as possible. It does not seem to leave much room for personal decisions – except perhaps for very trivial ones. It would be a threat to individual freedom (Bergström 2006, 7–8).

Again, since individual freedom is quite a basic interest, a proponent of the All Affected Principle can happily agree that the individual should have most of the influence in many decision and even be a dictator about certain decisions, for example such decisions that concerns her human rights.

7. Whelan's Objection

What is then the Whelan's original problem with the All Affected Principle? He worries that it "would require a different constituency of voters or participants for every decision" (Whelan 1983, 19). Similarly, Dahl writes that the "logic of the [all affected principle] (...) is that for every different set of persons affected there be a different association or decision-making unit" (Dahl 1970, 64). In other words, the All Affected Principle demands what is practically impossible.

This is surely true about the All Affected Principle taken as part of a practical decision method but misses the target if we take it as part of a normative ideal. As

with utilitarianism, the All Affected Principle might not be possible to use as an everyday decision method but it might still be correct as part of a democratic ideal. As such, it is part of an ideal that we use to evaluate the practical procedures that we implement in the real world in respect of how well they approximate the ideal. We will never be able to create a perfect democratic system but that is not an argument against trying to approximate it. To take an analogue example: It would be silly to criticise a criterion of “tall person” on the grounds that we cannot in practice measure length exactly. Although there is always going to be borderline cases, there are clear examples of procedures that are better and worse according to the All Affected Principle, as the examples used in this paper illustrates.

Whelan raises another problem that at first sight looks more damning:

The deeper problem is that before a democratic decision could be made on a particular issue (by those affected), a prior decision would have to be made, in each case, as to *who* is affected and therefore entitled to vote on the substantive issue... And how is this decision, which will be determinative of the ensuing substantive decision, to be made? It too should presumably be made democratically --- that is, by those affected --- but now we encounter a regress from which no procedural escape is possible. (...) Thus to say that those who will be affected by a given decision are the ones who should participate in making it is to attempt to bypass the crucial question, and to propose what is *logical as well as a procedural impossibility* (Whelan 1983, 19 last emphasis mine).

These are harsh words but Whelan’s reasoning begs the question. Why should we determine who is relevantly affected by certain decision by a prior democratic decision?⁴⁰ Why shouldn’t it, as I suggested above, be determined by a theory of the currency of relevantly affected and an analysis of the consequences of different courses of action, policies, and institutional structures on people’s interests?

When we try to figure out which theory of democracy is the best one, and which conception of relevant effects is correct, we have to weigh the evidence for and against different theories – it is an epistemic question, not a practical political

⁴⁰ One might think that there is an obvious end to Whelan’s regress: when everybody is included. This is, however, not true, since there is nothing that guarantees an expansion of the number of involved people for each step of the regress (I’m grateful to Wlodek Rabinowicz for pressing this issue). However, one might consider having a vote on what counts as being “relevantly affected” since this is bound to be an issue that reasonable people will disagree over. We could then use this notion of relevantly affected for any future decision. Intuitively, this is a vote that should involve everyone. Of course, this is impossible in practice but can still be an ideal that we should try to approximate, quite similar to the hypothetical consensus used in the contractarian tradition in political philosophy (e.g., Hobbes, Locke, Kant, Rousseau, and Rawls).

one. I don't think any democratic theorist has thought that whether or not their theory is correct depends on people's opinions about it. Compare again with utilitarianism according to which an act or policy is right if it maximises people's well-being. No utilitarian, however, has ever made the absurd suggestion that what constitutes well-being should be determined by the utilitarian principle.⁴¹

8. Another Worry

Here's another but related objection to the All Affected Principle that I at first found more troubling. Assume that you're the pivotal voter for the choice between two alternatives A and B, and that the choice will affect my relevant interests. It follows from the All Affected Principle that I should have some kind of influence over the choice, which is fine. However, your choice how to vote will also affect my interest, since your vote will determine whether A or B is chosen. Hence, it looks like I should have an influence on how you vote, according to the All Affected Principle, which might seem counterintuitive. Moreover, should we then vote on how you should vote in the first vote? Again, the outcome of that vote might also affect me. Thus, should we also vote on how you should vote in the second vote? We seem to get into an infinite regress.

A first answer to this objection is that it can be avoided by a reasonable theory of "relevantly affected". Arguably, it is one of my fundamental interests, like freedom of speech, to be decisive when it comes to which alternative I vote for. Thus, I should be decisive on such issues. Moreover, it seems reasonable in a democratic society that people should have the right to influence other people's voting behaviour by arguments and discussion, and that they in that sense should have some influence over how other people vote. Whether there will be a regress or not depends on which theory of "relevantly affected" the All Affected Principle is combined with, and there seems to be promising candidates that stops the regress in the first step. Of course, the devil is in the detail and to properly show this we need to work out our theory of relevantly affected.

⁴¹ It might be that Whelan himself believes in a normative theory according to which a principle can only be justifiably implemented in a constitution via some kind of democratic decision. This is an untenable idea since it would imply that all decisions to implement democracy in undemocratic countries would be unjustified.

More importantly, I think there is only a spurious infinite regress here. Consider a case with two outcome A and B in which three people, p_1 , p_2 , and p_3 are involved:

A: $(p_1, 1), (p_2, 1), (p_3, 0)$

B: $(p_1, 0), (p_2, 0), (p_3, 1)$

For each person, there is equally as much as stake in the choice between A and B, namely one unit (we are assuming that the numbers in the case is measuring relevantly affectedness on at least an interpersonally comparable ordinal scale). It follows, according to the All Affected Principle, that each person should have the same influence over the decision which could be secured by giving every person one vote and using the majority principle. Assume that, unsurprisingly, p_2 will vote for A and p_3 for B. Then p_1 is the pivotal voter for the choice between A and B and the choice will affect p_2 's and p_3 's relevant interests since they get either 0 or 1 unit depending on p_1 's choice.

Assume that it then follows from the All Affected Principle that p_2 and p_3 should have influence over p_1 's vote in proportion to how they are affected and that there is a vote on how p_1 should vote in the first choice. What is then at stake in this second vote? Well, either it will result in p_1 voting for A and then we get outcome $(p_1, 1), (p_2, 1), (p_3, 0)$, that is, A. Alternatively, it will result in p_1 voting for B and then we get outcome $(p_1, 0), (p_2, 0), (p_3, 1)$, that is, B. Likewise for the other possible voting patterns in this case. Thus, if we give people proportional influence over other people's votes, then we just reproduce the original choice situation and the distribution of influence will be the same. So no infinite regress.

9. Miller's Objection

David Miller argues that the All Affected Principle has a problem of circularity since to know who to include in a decision according to principle, we already need to know which outcome the decision will produce:

[I]f we take "affected interests" in its simplest sense to mean being made significantly better or worse off by the policies a *demos* adopts, then for any question that has to

be decided the domain is likely to vary with the outcome, creating a problem of circularity. That is, if group X has to make a decision between two policies, one of which is detrimental to group Y, but the other of which has no impact on that group, then to know whether the members of group Y should be included in the demos alongside the members of group X, we have already to know which outcome will occur. (Miller 2009, 215)

This argument, however, is based on an uncharitable analysis of what proponents of the All Affected Principle mean with being affected by a decision. You are affected, I and other have suggested, only if something is at stake for you, that is, if the two outcomes are of different value for you, given the currency of the All Affected Principle. Consider the following representation of Miller's case above:

Policy 1: (X, 10), (Y, -10)

Policy 2: (X, 5), (Y, 0)

One can say that the choice of policy 2 have no impact on group Y if one, like Miller, "assumes that a person's interests are affected by a decision when it makes him or her better or worse off by comparison to the status quo ante". But that is not what the proponents of the All Affected Principle have had in mind (at least not me). As we discussed above, the idea is that one is relevantly affected by a decision if and only if one is made better or worse off depending on the possible outcomes of the choice situation, that is, the difference between the values of the outcomes for you given the currency of the All Affected Principle. Group Y is affected in the above case since 10 units are at stake for them in the choice between policy 1 and 2. So the people affected doesn't vary with the outcome that actually comes about but only with the choice situation.

Miller seem to suggest, however, that with this approach, we face another devastating problem:

If to avoid that problem we say that everyone who may *possibly* be affected by a decision should be included in the constituency that makes it, then the *demos* will expand in all directions, depending on which possibilities are contemplated in the deliberation leading up to the decision. The upshot is that if we adopt the affected interests principle, we can only be sure of avoiding the circularity problem by making democracy's domain universal, that is to say by including every (...) human being in the demos. (Miller 2009, 215). Cf. (Goodin 2007, 52–53).

This is, however, a false dichotomy. There is a third alternative, namely that we include those people whose interests are *actually* at stake in a certain decision, like the Y- and X-people in the case above. We should not include those people whose interests might possibly be at stake, or would have been at stake if we were in another choice situation. Alternatively, we can take probabilities into account but then we should of course adjust people influence to match the probability that their interests are at stake. This means that for most decisions, the vast majority of people will have no or close to no influence at all. In both cases, it doesn't follow that everybody should be included.

10. Summary

I hope to have showed that the seemingly intractable theoretical problems raised by the democratic boundary problem are solvable and that the All Affected Principle is still very much a live candidate as a boundary principle for democracy as a partial normative ideal. Moreover, I hope to have contributed to a fruitful way of approaching the boundary problem so that future work can move ahead on the important question of developing a currency for the All Affected Principle and what notion of influence with which the principle should be coupled.

To forestall some possible objections, however, let me point out that I haven't given a full-fledged defence of the All Affected Principle and of democracy as a normative ideal but have focused on clarifying how to understand these two ideas and the boundary problem, and a rebuttal of the criticism that has been directed towards the principle. We might, in the end, reach the conclusion that democracy is not an attractive normative ideal as compared to other such ideals – further inquiry has to determine this. However, the reason will not be that the All Affected principle “propose[s] what is logical as well as a procedural impossibility” or is “a threat to individual freedom” or that it makes “democracy's domain universal”, etc., but that another normative ideal is more in reflective equilibrium with our considered judgements. Recall also that democracy need not be our only ideal, we might have other ideals that the democratic ideal has to be weighed against in reaching a final theory of how a just or good society should look like.

Secondly, the All Affected Principle fits well with democratic ideals that concern distribution of influence or power. It might not fit other conceptions very well, however. The epistemic democrat is a case in point. According to the epistemic conception of democracy, what makes democracy valuable is that it generates better decisions, that is, it is a better “truth-tracker” than alternative decision-making procedures.⁴² The All Affected Principle doesn’t fit well with this conception of democracy. Rather, one could argue that inclusions and exclusions should be based on whether it is likely that the voter in question is better than chance at getting the right answer.⁴³ If there is good reason to believe that the average voter will be better than chance at getting the correct answer, then we have a *prima facie* argument for including as many people as possible in a decision making process. If not, the argument works the other way around, that is, as a reason for excluding people from the decision. In other words, the criterion of inclusion and exclusions for the epistemic democrat has to do with what will generate the most accurate decisions, not with who is relevantly affected by the decision.

I find it more plausible to take epistemic democracy as practical decision method that is justified by an appeal to some normative ideal, such as the utilitarian value of reaching the right decisions in many contexts. Some might not agree, however. If so, then we have a normative democratic ideal for which the All Affected Principle is not a suitable boundary criterion but for which there is another promising candidate: competence at getting the answers right.⁴⁴

⁴² See, for example (Estlund 1990, 1993, 1997, 1998), (List and Goodin 2001), (J. Cohen 1986), (Pettit 2001), (Bovens and Rabinowicz 2006; Condorcet 1972).

⁴³ See (Arrhenius 2005).

⁴⁴ I would like to thank Ludvig Beckman, Krister Bykvist, Tim Campbell, Jerry Cohen, Bob Goodin, Marc Fleurbaey, David Miller, Julia Mosquera, Lennart Nordenfeldt, Sarah Philipson, Shlomi Segall, Orri Stefansson, Wlodek Rabinowicz, Folke Tersman, Torbjörn Tännsjö, Andrew Williams, Ashwini Vasanthakumar, Björn Wittrock, and Annette Zimmermann for stimulating discussions and criticism. Earlier versions of this paper were presented at the Jerusalem Political Philosophy Forum, The Hebrew University, 5 March, 2008, the *Nordic Network for Political Ethics*, Vejle, June 1, 2011, *Nuffield Political Theory Workshop*, May 15, 2017, and *VIII Braga Meetings on Ethics and Political Philosophy*, 9 June, 2017. I would like to thank the audience at these occasions for their comments. Thanks also to IEA-Paris and the Swedish Collegium for Advanced Study for being such generous hosts during some of the time when this paper were written. Financial support from the Swedish Research Council is gratefully acknowledged.

References

Arrhenius, Gustaf. 2005. 'The Boundary Problem in Democratic Theory'. In *Democracy Unbound: Basic Explorations*, edited by Folke Tersman, 14–29. Stockholm Studies in Democratic Theory. Stockholm: Stockholm University. Filosofiska institutionen.

———. 2008a. 'Democracy and the Measurement of Influence'. Mimeo. Stockholm University, Department of Philosophy.

———. 2008b. 'Measuring and Distributing Potential Influence'. Mimeo. Stockholm University, Department of Philosophy.

———. 2012. 'Politisk Och Ekonomisk Demokrati'. In *Tillsammans : En Fungerande Ekonomisk Demokrati*. Stockholm: SNS förlag.

———. 2015. 'The All Affected Principle and Future Generations'. Mimeo. Stockholm University, Department of Philosophy.

———. 2017. 'Political and Economic Democracy'. Mimeo, Institute for Futures Studies.

———. 2018. 'Democracy for the 21th Century Research Challenges'. Forthcoming in Wittrock et al *Sociology looks at the twenty-first century*.

Bales, R. Eugene. 1971. 'Act-Utilitarianism: Account of Right-Making Characteristics or Decision-Making Procedure?' *American Philosophical Quarterly* 8 (3): 257–265.

Banzhaf, John F. 1965. 'Weighted Voting Doesn't Work: A Mathematical Analysis'. *Rutgers Law Review* 19: 317–43.

———. 1966. 'Multi-Member Electoral Districts - Do They Violate the "One Man, One Vote" Principle'. *Yale Law Journal* 75 (7): 1309–38.

———. 1968. 'One Man, 3,312 Votes: A Mathematical Analysis of the Electoral College'. *Vill. L. Rev.* 13: 304.

Barry, Brian. 1991. *Democracy and Power: Essays in Political Theory 1*. Oxford: New York: Oxford University Press.

Bauböck, Rainer. 2006. 'The Rights of Others and the Boundaries of Democracy'. *European Journal of Political Theory* 6 (4): 398–405.

———. 2007. *Migration and Citizenship. Legal Status, Rights and Political Participation*. Amsterdam University Press.

Beckman, Ludvig. 2006. 'Citizenship and Voting Rights: Should Resident Aliens Vote?' *Citizenship Studies* 10 (2): 153–65.

———. 2009. *The Frontiers of Democracy: The Right to Vote and Its Limits*. 1 edition. Basingstoke: New York: Palgrave Macmillan.

———. 2014. 'The Subjects of Collectively Binding Decisions: Democratic Inclusion and Extraterritorial Law'. *Ratio Juris* 27 (2): 252–270.

Bergström, Lars. 2005. 'Democracy and Future Generations'. In *Democracy Unbound: Basic Explorations I*, edited by Folke Tersman, 190–92. Stockholm: Filosofiska institutionen, Stockholms universitet.

———. 2006. 'Democracy and Political Boundaries'. Mimeo, Department of Philosophy, Stockholm University.

———. 2009. 'Principles of Voting Power'. Mimeo. Stockholm University, Department of Philosophy.

Bovens, Luc, and Wlodek Rabinowicz. 2006. 'Democratic Answers to Complex Questions: An Epistemic Perspective'. *Synthese* 150 (1): 131–53.

Brighouse, Harry, and Marc Fleurbaey. 2006. 'On the Fair Allocation of Power'. Mimeo. CERSES, CNRS.

———. 2010. 'Democracy and Proportionality'. *Journal of Political Philosophy* 18 (2): 137–55.

Brink, David O. 1986. 'Utilitarian Morality and the Personal Point of View'. *The Journal of Philosophy* 83 (8): 417–38.

Christiano, Thomas. 1996. *The Rule of the Many: Fundamental Issues in Democratic Theory*. Boulder, Colo: Westview Press.

———. 2002. 'Democracy as Equality'. In *Democracy*, edited by David Estlund, 30–55. Malden, MA: John Wiley & Sons.

Cohen, Carl. 1971. *Democracy*. Athens: University of Georgia Press.

Cohen, Joshua. 1986. 'An Epistemic Conception of Democracy'. *Ethics* 97 (1): 26–38.

———. 1989. 'Deliberation and Democratic Legitimacy'. In *The Good Polity : Normative Analysis of the State*, edited by Alan P. Hamlin and Philip Pettit, 17–34. Oxford: Basil Blackwell.

Condorcet, Jean Antoine Nicolas de Caritat. 1972. *Essai Sur L'application de L'analyse à La Probabilité Des Décisions Rendus à La Pluralité Des Voix*. New York: Chelsea Publ.co.

Cunningham, Frank. 1987. *Democratic Theory and Socialism*. Cambridge ; New York: Cambridge University Press.

———. 1994. *The Real World of Democracy Revisited, and Other Essays on Democracy and Socialism*. Atlantic Highlands, N.J: Humanities Press.

Dahl, Robert A. 1970. *After the Revolution? Authority in a Good Society*. New Haven: Yale University Press.

———. 1989. *Democracy and Its Critics*. New Haven: Yale University Press.

Danielsson, Sven. 1974. 'Two Papers on Rationality and Group Preferences'. *Filosofiska Studier*, Uppsala: Filosofiska föreningen och filosofiska institutionen, 21.

Dworkin, Ronald. 1981a. 'Is There a Right to Pornography?' *Oxford Journal of Legal Studies* 1 (2): 177–212.

———. 1981b. 'What Is Equality? Part 1: Equality of Welfare'. *Philosophy & Public Affairs* 10 (3): 185–246.

———. 1981c. 'What Is Equality? Part 2: Equality of Resources'. *Philosophy & Public Affairs* 10 (4): 283–345.

———. 2000. *Sovereign Virtue: The Theory and Practice of Equality*. Cambridge, Mass: Harvard University Press.

———. 2002. 'Sovereign Virtue Revisited'. *Ethics* 113 (1): 106–143.

Encyclopedia Britannica. 2014. 'Good Friday Agreement'. *Encyclopedia Britannica*.

Estlund, David. 1990. 'Democracy Without Preference'. *The Philosophical Review* 99 (3): 397–423.

———. 1993. 'Making Truth Safe for Democracy'. In *The Idea of Democracy*, edited by David Copp, Jean Hampton, and John E. Roemer, 71–100. New York: Cambridge University Press.

———. 1997. 'Beyond Fairness and Deliberation: The Epistemic Dimension of Democratic Authority'. In *Deliberative Democracy: Essays on Reason and Politics*, edited by James Bohman and William Rehg. MIT Press.

———. 1998. 'The Insularity of the Reasonable: Why Political Liberalism Must Admit the Truth'. *Ethics* 108 (2): 252–75.

Felsenthal, Dan S., and Moshé Machover. 1998. *The Measurement of Voting Power: Theory and Practice, Problems and Paradoxes*. Cheltenham: Edward Elgar.

Frege, Gottlob. 1970. 'The Fundamental Laws of Arithmetic II'. In *Translations from the Philosophical Writings of Gottlob Frege*, edited by Peter Geach and Max Black, translated by Peter Geach, 2nd edition. Oxford: Blackwell.

Goldman, Alvin I. 1974. 'On the Measurement of Power'. *The Journal of Philosophy* 71 (8): 231–52.

Goodin, Robert E. 2007. 'Enfranchising All Affected Interests, and Its Alternatives'. *Philosophy and Public Affairs* 35 (1): 40–68.

———. 2016. 'Enfranchising All Subjected, Worldwide'. *International Theory* 8 (03): 365–89.

Grace, Jeremy. 2003. 'The Electoral Rights of Conflict Forced Migrants: A Review of Relevant Legal Norms and Instruments'. IOM International Organization for Migration Discussion Paper No.1. Participatory Elections Project (PEP).

Kymlicka, Will. 1990. *Contemporary Political Philosophy: An Introduction*. Oxford [England] : New York: Clarendon Press ; Oxford University Press.

List, Christin, and Robert E. Goodin. 2001. 'Epistemic Democracy: Generalizing the Condorcet Jury Theorem'. *Journal of Political Philosophy* 9 (3): 277–306.

López-Guerra, Claudio. 2005. 'Should Expatriates Vote?' *Journal of Political Philosophy* 13 (2): 216–234.

———. 2014. *Democracy and Disenfranchisement: The Morality of Electoral Exclusions*. Oxford, New York: Oxford University Press.

Miller, David. 2009. 'Democracy's Domain'. *Philosophy & Public Affairs* 37 (3): 201–228.

Morriss, Peter. 1987. *Power: A Philosophical Analysis*. Manchester: Manchester Univ. Pr.

Næss, Arne, Jens A. Christophersen, and Kjell Kvalø. 1956. *Democracy, Ideology and Objectivity : Studies in the Semantics and Cognitive Analysis of Ideological Controversy*. Oslo: Oslo University Press.

Näsström, Sofia. 2003. 'What Globalization Overshadows'. *Political Theory* 31 (6): 808–34.

———. 2004. 'The An-Archical State : Logics of Legitimacy in the Social Contract Tradition'. Doctoral Thesis, Stockholm: Stockholm University, Faculty of Social Sciences, Department of Political Science.

Nozick, Robert. 1974. *Anarchy, State, and Utopia*. Malden, MA: Blackwell.

Owen, David. 2012. 'Constituting the Polity, Constituting the Demos: On the Place of the All Affected Interests Principle in Democratic Theory and In Resolving the Democratic Boundary Problem'. *Ethics & Global Politics* 5 (3): 129–52.

Pennock, J. Roland. 1974. 'Democracy Is Not Paradoxical'. *Political Theory* 2 (1): 88–93.

Penrose, Lionel S. 1946. 'The Elementary Statistics of Majority Voting'. *Journal of the Royal Statistical Society* 109 (1): 53–57.

Pettit, Philip. 2001. 'Deliberative Democracy and the Discursive Dilemma'. *Philosophical Issues* 11 (1): 268–299.

Rawls, John. 1971. *A Theory of Justice*. Cambridge, Mass.: Belknap Press.

Ross, Alf. 1968. *Varför Demokrati?* Stockholm: Tiden.

Rubio-Marín, Ruth. 2006. 'Transnational Politics and the Democratic Nation-State: Normative Challenges of Expatriate Voting and Nationality Retention of Emigrants'. *New York University Law Review* 81 (117): 117–47.

Schumpeter, Joseph A. 1976. *Capitalism, Socialism, and Democracy*. 5th ed. London: Allen and Unwin.

Sen, Amartya. 1970. *Collective Choice and Social Welfare*. Mathematical Economics Texts 5. San Francisco: Holden-Day.

———. 1985. *Commodities and Capabilities*. Amsterdam: North-Holland.

———. 1992. 'Justice and Capability'. In *Inequality Reexamined*, 73–101. New York: Cambridge, Mass: Russell Sage Foundation ; Harvard University Press.

Shapiro, Ian. 1996. *Democracy's Place*. 1 edition. Ithaca: Cornell University Press.

Shapley, L. S., and Martin Shubik. 1954. 'A Method for Evaluating the Distribution of Power in a Committee System'. *American Political Science Review* 48 (3): 787–92.

Svallfors, Stefan. 2016. 'Politics as Organised Combat – New Players and New Rules of the Game in Sweden'. *New Political Economy* 21 (6): 505–19. <https://doi.org/10.1080/13563467.2016.1156662>.

Svallfors, Stefan, and Anna Tyllström. 2018. 'Resilient Privatization: The Puzzling Case of for-Profit Welfare Providers in Sweden'. *Socio-Economic Review*, January. <https://doi.org/10.1093/ser/mwy005>.

Tännsjö, Torbjörn. 1992. *Populist Democracy: A Defence*. London ; New York: Routledge.

———. 2005. 'Future People and the All Affected Principle'. In *Democracy Unbound: Basic Explorations*, edited by Folke Tersman, 186–89. Stockholm Studies in Democratic Theory 1. Stockholm: Filosofiska institutionen, Stockholms universitet.

———. 2007. 'Future People, the All Affected Principle, and the Limits of the Aggregation Model of Democracy'. In *Hommage à Wlodek. Philosophical Papers Dedicated to Wlodek Rabinowicz*, edited by Toni Rønnow-Rasmussen, Björn Petersson, Jonas Josefsson, and Dan Egonsson. Stockholm University.

Tersman, Folke. 1993. *Reflective Equilibrium: An Essay in Moral Epistemology*. Stockholm Studies in Philosophy 14. Stockholm: Almqvist & Wiksell International.

Whelan, Frederick G. 1983. 'Democratic Theory and the Boundary Problem'. In *Liberal Democracy*, edited by J. Roland Pennock and John W. Chapman. New York: New York University Press.



ARE MORAL FUNCTIONALISM'S MORAL A PRIORI COMMITMENTS REALLY A PRIORI?

SERÁ QUE OS COMPROMISSOS A PRIORI DO FUNCIONALISMO MORAL SÃO, DE FACTO, A PRIORI?

Ashley Lane*
ashleyleane1999@yahoo.co.uk

Abstract. *Moral functionalism, a metaethical theory developed by Frank Jackson and Philip Pettit, claims that we can attain moral knowledge by ascertaining the commonplaces about morality that are typically accepted by actual agents. It has important a priori commitments; whilst we may discover a posteriori that a particular descriptive property is identical to a particular moral property, it is a priori that the thing that is identical to the moral property, whatever that thing actually is, plays a particular role. Jackson holds a particular metaphysical position, and moral functionalism is a development of that position as it applies to ethics. In this paper I adapt an objection made by D.H. Mellor against Jackson's metaphysics to show that moral functionalism's a priori commitments are actually a posteriori. We can only discover if moral functionalism's purportedly a priori claims are true through a posteriori investigation.*

Keywords: *Moral functionalism, metaethics, metaphysics.*

Sumário. *O funcionalismo moral, uma teoria desenvolvida por Frank Jackson e Philip Petti, afirma que podemos obter conhecimento moral determinando os lugares comuns acerca da moralidade que tipicamente são aceites por agentes concretos. É uma posição que tem compromissos a priori importantes; podemos descobrir a posteriori que uma propriedade descritiva particular é idêntica a uma propriedade moral particular, mas é a priori que a coisa que é idêntica à propriedade moral desempenha um papel moral particular. Jackson defende uma posição metafísica particular e o funcionalismo moral é um desenvolvimento desta posição aplicada à ética. Neste artigo adapto uma objecção feita por D. H. Mellor contra a metafísica de Jackson para mostrar que os compromissos a priori do funcionalismo são, de facto, a posteriori. Só podemos descobrir se as afirmações alegadamente a priori do funcionalismo moral são verdadeiras através de uma investigação a posteriori.*

Palavras-chave: *funcionalismo moral, meta-ética, metafísica.*

* PhD student, Birkbeck, University of London.

Moral functionalism, a metaethical theory developed by Frank Jackson and Philip Pettit, claims that we are able to attain moral knowledge by ascertaining the commonplaces about morality that are typically accepted by actual agents. However, whilst we may be able to ascertain these commonplaces through *a posteriori* investigation, moral functionalism has certain purportedly *a priori* commitments regarding the extension of moral terms, which cannot be verified *a posteriori*. Moral functionalism represents the application of Jackson's general metaphysics to ethics, and moral functionalism's *a priori* commitments are a type of commitment that is also found elsewhere in Jackson's general metaphysics.

Even if these commitments are true, though, and contrary to what Jackson claims, they are not *a priori*. We can use an objection made by D.H. Mellor against Jackson's general metaphysics to generate a problem for moral functionalism. Jackson believes that we identify things such as water with what actually plays a particular role; in the case of water, H₂O plays a particular role which makes it identical with water. Mellor denies that this is true. As Jackson claims that moral properties are identical with whatever plays particular "moral property" roles, we can adapt Mellor's objection to criticise moral functionalism. The problem does not show that the commitments are false, but it does show that they can only be ascertained through *a posteriori* investigation. They therefore cannot be *a priori*, and it may be possible to show *a posteriori* that they are false. In the first section of this paper, I will set out moral functionalism's commitments. In the second section, I will outline Mellor's objection and adapt it to show that the commitments are not *a priori*.

1. Moral functionalism

Jackson and Pettit (1995) claim that agents use moral terms "in a way that presupposes a large network of connections with other terms, both evaluative and descriptive" (p. 22). We accept commonplaces about the terms we use, and this allows us to identify particular descriptive properties that ordinary moral thinking tells us are moral properties. We are thus able to grasp moral concepts. For example, when we use the term "fairness", it is a commonplace for us that if

an action is fair, that is usually a reason to do it. Agents who use moral terms competently know that the commonplace is true.

[T]he meaning of relevant moral terms will be fixed by roles which certain commonplaces give them, and so moral thinking is bound to involve the attempt to use commonplaces as a base, and holding on to as much of that base as possible, or at least to the parts considered most secure, to fix opinions on particular questions. (p. 26)

Grasping the commonplaces correctly means that we grasp the concepts relating to the commonplaces correctly, which then means that we can correctly use the terms that denote moral properties. Once we know about the roles that moral properties play and the moral commonplaces that we accept, we can find out which descriptive properties we can identify with the moral properties.

Moral functionalism has a certain type of *a priori* commitment (It may have others, but I will only focus on one type here). Jackson (1998) claims that

[w]hat is *a priori* according to moral functionalism is not that rightness is such-and-such a descriptive property, but rather that *A* is right if and only if *A* has whatever property it is that plays the rightness role in common folk morality, and it is an *a posteriori* matter what the property is. (pp. 150-151)

“Common folk morality” is the moral theory that endorses the moral commonplaces accepted by most people in a particular society (pp. 117-118). Moral functionalism proposes that we look at the commonplaces typically endorsed by people, and that we use that as the basis of a suitable moral theory. We hold on to as many of these commonplaces as possible, and they allow us to identify moral properties with particular descriptive properties. However, to know what moral properties are, we must know *a priori* that moral properties play a particular role. Only after that can we use *a posteriori* evidence to ascertain which descriptive properties are moral properties, since they play that role. To take a non-moral example, we discover *a posteriori* that water is H₂O. However, we have *a priori* knowledge that the substance that is water is the substance that plays a certain “watery” role of being transparent, drinkable, and so on. The *a posteriori* discovery is that H₂O plays the role, and so is identical to water.

Since the moral commonplaces we accept may conflict irreconcilably with each other, we can refine common folk morality into a mature folk morality. As people continue debating moral issues and refining their moral views, we work out which commonplaces to keep and which to drop. Jackson claims that this gets

us closer to a mature folk morality, a “folk morality that has been exposed to debate and rational reflection”. It makes “good sense of the raft of sometimes conflicting intuitions about particular cases and general principles that make up current morality” (p. 113).

In his metaethics, as in his general metaphysics, Jackson separates the concepts that actual agents use from what he calls the “essence” of what these concepts relate to. For example, what is essential about water is that it is H₂O, but we do not need to know that to grasp the concept “water” (p. 50). People grasped the concept and could refer to water long before they knew its chemical composition. It is this distinction that makes clear the *a priori* commitments of moral functionalism.

Jackson accepts that the investigation of which commonplaces we endorse is *a posteriori*. The *a priori* element becomes apparent when we talk about the application of terms in the actual world or in counterfactual situations. Let T be a particular term. T may apply to various entities, events, relations, etc., and this defines its extension. It may have different extensions in different possible worlds. The *A-extension* of T in *w* is the actual extension of T in world *w*, where *w* may be our world or another possible world. In our world, the A-extension of “water” is all and only the occurrences of water.

We may ask, given the assumption that we are talking about our actual world, what T would apply to under various counterfactual situations. The answer would give us the *C-extension* of T, which would tell us what T would apply to across possible worlds. For example, the term “water” applies to all the watery occurrences in a world, so in the actual world the A-extension of “water” would be all and only the occurrences of what we call “water”, which are occurrences of H₂O. In another world, where substances of chemical composition XYZ are called “water”, the A-extension of “water” in that world would be all and only the occurrences of XYZ. But since in our actual world water is H₂O, the C-extension of “water” in the actual world is the occurrences of H₂O in every possible world (p. 49). The C-extension would not extend to any occurrences of XYZ, even though people on other possible worlds may call XYZ “water”.

For “water”, the A-extension and the C-extension in our actual world are the same. “Water” applies to all and only occurrences of H₂O under both extensions.

But, says Jackson, there is an epistemological difference. Before we found out that water was H₂O, we did not know its essence. So we could not determine its C-extension.

[I]n order to pick out water in a counterfactual world, we need to know something about *relationships* between the counterfactual world and the actual world that we could only know after discovering that in the actual world H₂O plays the watery role. (p. 50)

Before we found out that water was H₂O – an *a posteriori* discovery – we could not say whether the stuff that plays the watery role in a counterfactual world was part of the C-extension of “water”. For consider a counterfactual world that is identical with the actual world except that the stuff that plays the watery role in the actual world is H₂O and the stuff that plays the watery role in the counterfactual world is XYZ. Until we find out that water in the actual world is H₂O, how do we work out that the stuff playing the watery role in the counterfactual world does not fall under the C-extension of “water”?

A-extensions (in our actual world) do not need *a posteriori* knowledge of the actual world. “Water” is whatever plays the watery role in the actual world. We do not need to know that water is H₂O in order to know that. And this applies to any counterfactual world as well, because when we work out A-extensions of “water” in a counterfactual world, we do not compare the watery substance in that world to the watery substance in any other world.

Since we can know A-extensions in the actual world without knowing what the actual world is like, Jackson concludes that such knowledge is *a priori* (p. 51).¹ In the actual world, it is *a priori* that the A-extension of “water” is just the watery stuff of our acquaintance. Water is whatever happens to fulfil the watery role. What precisely that stuff is (H₂O) is an *a posteriori* question, but we do not need to know what precisely the watery stuff is in order to know that water is the watery stuff of our acquaintance.

What about the term “fairness”? The A-extension of “fairness” (at our actual world) will be the thing that we are acquainted with because it fulfils a certain “fairness” role, and we know this *a priori*. What we must discover *a posteriori* is exactly what descriptive property fulfils the role. So even though we can

¹ We cannot analyse all terms like this (such as names and demonstratives), and I will discuss this briefly later.

examine many moral commonplaces *a posteriori*, it is *a priori* that moral properties are simply the things that play the various “moral property” roles, that play certain roles in the eyes of moral agents. We can roughly say that moral properties are those that fulfil certain functions according to our common folk morality, which we can refine to become a mature folk morality. Metaethical theories that analyse moral concepts in ways that are completely alien to our ordinary conception of them are most likely wrong (p. 31). The way to find out what our ordinary conception is is to go out and look at what moral commonplaces most people endorse. What is *a priori* is our knowledge of the A-extensions of moral terms. What is *a posteriori* is our knowledge of the “essence” of moral properties. We investigate *a posteriori* what descriptive properties play the “moral property” roles, and so discover which descriptive properties are identical with which moral properties. This identification reveals the “essence” of moral properties, just as we discover that the “essence” of water is that it is H₂O.

Both A-extension propositions (propositions that express A-extensions) and C-extension propositions can be commonplaces. The proposition “The property of moral rightness is whatever plays the moral-rightness role” is a commonplace, and it expresses an A-extension. However, Jackson and Pettit (1995) also believe that for many people it is a commonplace that saving lives is morally more important than being fair (p. 23). This seems to be an *a posteriori* commonplace, since humans might have come to believe it a commonplace that being fair is morally more important than saving lives instead. We find out such commonplaces *a posteriori*, and in fact we come to identify particular descriptive properties as moral properties *a posteriori*. Such identifications are discoveries of moral C-extensions. Commonplaces are able to express A-extensions or C-extensions, and thus can be either ascertainable *a priori* or *a posteriori*.

Despite the significant *a posteriori* elements in moral functionalism, Jackson tells us that it ultimately rests on an *a priori* foundation of A-extensions. However, this foundation is not actually *a priori*. In the next section, I will examine an objection from D.H. Mellor that Jackson’s endorsement of A-extensions in his metaphysics fails. I will use this objection to argue not that moral functionalism’s A-extensions fail, but that they are *a posteriori*. They are thus vulnerable to *a posteriori* objections.

2. Mellor's objection

Return to Jackson's metaphysics. Consider the following argument:

- a) Sixty percent of the earth is covered by H₂O.
- b) Water is the stuff that plays the watery role.
- c) H₂O is the stuff that plays the watery role.

Therefore

- d) Sixty percent of the earth is covered by water.

(a) and (c) are *a posteriori*. Before we could know (c), we had to check that H₂O actually plays the watery role. According to Jackson (2003), (b) is *a priori* (p. 87). It is something we can know *a priori* about water.

Consider now:

- e) Ben performed an act that maximised utility.
- f) The property of moral rightness is whatever plays the moral-rightness role.
- g) The property of maximising utility plays the moral-rightness role.

Therefore

- h) Ben performed a morally right act.

(e) and (g) are *a posteriori*, as (g) acts as a commonplace that we must discover *a posteriori*. (f) is *a priori*, which is what we should expect, since it is an A-extension proposition. Jackson's justification for (b) and (f) being *a priori* is that they are about reference fixers. As far as (b) goes, "water" refers to what plays the watery role, the stuff that plays the role of a colourless drinkable liquid that makes up the ocean and so on. It does not indicate what actually plays the watery role or what the watery role consists of.

The same happens with (f). (f) may be *a priori*, but note how little information it contains. It does not tell us a great deal about moral properties or what their characteristics are. But without it, Jackson claims, we could not

discover which descriptive property is identical with the moral-rightness property. The *a posteriori* discovery rests on *a priori* assumptions about moral properties.

Currently, then, we have *a priori* A-extension propositions that the moral functionalist uses to justify moral C-extension propositions. The critic's aim is to show that this type of justification always fails, and that *a priori* A-extension propositions cannot justify any C-extension propositions. There are at least three strategies what we can use.

a) Moral A-extension propositions are false, and so cannot help justify any C-extension propositions.

b) Moral A-extension propositions are really *a posteriori*, contrary to what moral functionalism claims.

c) Moral A-extension propositions cannot be used to justify any moral C-extension propositions, even if the A-extension propositions are true.

If option (c) is right, then moral functionalism has *a priori* commitments, but these commitments are effectively redundant in moral arguments. We cannot use them to justify any moral C-extension propositions, and so their role in a folk morality is very small. Moral C-extension propositions must be justified on other grounds which have nothing to do with A-extensions. However, as my focus is on options (a) and (b), I will not develop option (c) here.

Option (a) looks attractive. The proposition "X is whatever plays the X-role" (whether X is a moral property or something else) may be wrong in several ways. Most obviously, the X-role may not exist, although this is unlikely to be a truly damning objection to Jackson's metaphysics. For example, "X" must not be a name; London exists, but it is not the case that the thing called "London" must play a "Londony" role. People just applied the name "London" to a particular entity. However, all Jackson has to do is exclude names from his metaphysics, and he is back in business. It seems hard to claim that there is no watery role for water to play, because the watery role is just a collection of properties that the thing that is water has to have. "Water" is not a name. The same point holds for demonstratives such as "this" and "these". The proposition "This is whatever

plays the thisy role” is false, because there is no thisy role, but Jackson has no reason to claim otherwise.

More interestingly, it may be that we are mistaken about the nature of X, and hence of what the X-role should be. Suppose a scientist finds a new chemical that appears to dissolve plastic. For something to be this chemical, he believes, it has to play a role that includes the ability to dissolve plastic. It comes to be generally accepted that if something cannot dissolve plastic, it cannot be the chemical. But then it is discovered that the scientist made a mistake, and the chemical does not dissolve plastic. The chemical role includes the ability to dissolve plastic, the chemical itself cannot dissolve plastic, and so the chemical is not identical with whatever plays the chemical role.

This is not the only way in which we might be mistaken about the relation of X to the X-role. Take the example of water. We say that water is transparent, drinkable, found in oceans, and so on, and to Jackson these properties constitute the watery role. It was not necessarily the case that anything has all these properties, though, or that there was only one thing that did. We might have examined oceans and found out that what we call “water” is actually a mixture of H₂O and XYZ.

Another possibility is that the thing that actually plays the X-role also has other properties that X itself does not have, and so the thing that plays the X-role cannot be X. D.H. Mellor (2003) argues that H₂O is not identical to water, despite playing the watery role. Even if we allow that ice and steam are water, “no single H₂O molecule can be water, since it instantiates hardly any of water’s laws, having no solvent powers, density, freezing or boiling points, or latent heats” (p. 224). Of course, Mellor does not deny that water exists or that it is made up of H₂O. What he denies is that H₂O is identical to water, despite it playing the watery role.

It is tempting to think that we can make the same sort of objection about moral A-extensions, and thus show that option (a) is correct. Actually, it is difficult to defend option (a), and it is much better to use option (b) to criticise moral functionalism. I shall outline a couple of ways in which one might try to use option (a) before moving on to Mellor’s objection. Initially, it appears that Mellor’s objection can be used to show that option (a) is right. However, it really ends up showing that option (b) is correct.

If option (a) is correct, then the following moral A-extension proposition is false:

MRP: The property of moral rightness is whatever plays the moral-rightness role

MRP can be questioned in several ways. First, it may be the case that it is impossible for anything to play the moral-rightness role. This could happen, for example, if the role is logically inconsistent. If the property of moral rightness exists, MRP cannot be true. Not only that, but the moral functionalist is left trying to find another way in which moral agents can recognise moral properties. MRP offered a straightforward way to do so – check what descriptive properties play a certain role, and those properties are moral properties. Now another method must be found.

So the objection throws up two problems for the moral functionalist. The first problem is metaphysical – if moral properties are not the properties that play a particular moral role, what are they? The second is epistemological – how can agents recognise moral properties if the properties do not play a particular role? However, the moral functionalist need not worry yet, because it first has to be shown that it is impossible for anything to play moral roles. The moral functionalist will of course deny that it is impossible, because we seem to identify moral roles all the time with the aid of the moral commonplaces we accept. The “fairness” role is the role of being even-handed, impartial, unaffected by biases, and so on. Where does the logical inconsistency lie in this? And if we can identify an inconsistency, why can we not just alter our conception of the role to exclude it? Jackson explicitly states that we can refine our moral commonplaces to exclude inconsistencies and so develop a mature folk morality.

Another way that MRP may be false is if there is no connection between the property of moral rightness and the moral-rightness role. Suppose we discover a new non-moral property, Y, and we ascribe a number of characteristics to it that come to define the Y role. We then discover that Y has near to none of those characteristics. The property of Y is not what plays the Y role. Could this not be the case for the property of moral rightness? The trouble with this objection is that it once again does not allow for correction and refinement. As we find out more about Y, we refine our understanding of what the Y role is, so once we have

realised our mistake we change our conception of the Y role to reflect characteristics that Y actually has. Our conception of the Y role changes so that whatever plays the Y role is indeed identical to Y. This does not seem impossible, so the moral functionalist can say the same thing about moral properties. We may be currently mistaken about what the “moral rightness” role is, but by examining the property of moral rightness we can correct our beliefs about what the “moral rightness” role is, so that whatever plays that role is identical to the property of moral rightness.

The third objection is a variation of Mellor’s metaphysical objection. Something plays the moral-rightness role, and there may be such a thing as moral rightness, but the property that plays the moral-rightness role is not identical to the property of moral rightness. Take MRP. Moral-Mellor may say about it:

Assume that there is a property that plays the moral-rightness role. Let us say that the property of maximising welfare plays it. That certainly does not mean that such a property is the property of moral rightness, because it may not play that role at all times and under all conditions (H₂O plays the watery role, but it does not always do so.). And we certainly cannot assume ahead of investigation that there is any single descriptive property that is identical to the property of moral rightness, or that there is a collection of descriptive properties that constitute the property of moral rightness.

Call this the Moral-Mellor objection. This objection concedes a certain amount to the moral functionalist. It allows that agents can recognise moral properties by examining what plays moral roles (Even if water is not identical to what plays the watery role, examining what plays the watery role will help us identify examples of water.). It is consistent with the claims that there are moral properties, that there are moral roles that descriptive properties play, and that we can know about moral properties and roles. There may also still be a metaphysical connection between a moral property and a moral role. If something does not play the watery role, then it cannot ever be an occurrence of water. Similarly, if something does not play the role of a moral property, then it can never be identical to the moral property. The only claim that is rejected is the moral functionalist’s claim that moral A-extension propositions are correct. It may be possible to know the proposition “Maximising the welfare of the homeless is morally right”, even though the property of maximising welfare is not identified with any particular moral property.

At this point, the moral functionalist can reasonably object that we cannot just assume that moral A-extension propositions are false. Perhaps when we investigate further, we will discover that the propositions are true. This is a fair comment, but we can now show that even if they are true, they are not *a priori*. H₂O is identical to water, says Jackson. Not so, says Mellor. Mellor's objection appears to be an *a posteriori* objection. H₂O does not always physically act like water; for example, occurrences of H₂O need not have any boiling point. By observation we attribute various features to occurrences of water and of H₂O, and it is a matter of *a posteriori* fact that these features are not always the same.

So it appears that we discover whether the proposition "Water is the stuff that plays the watery role" is true or false by *a posteriori* observation. H₂O plays the watery role, but that does not necessarily mean that it is identical to water. We cannot discover *a priori* that it is water, and so we cannot discover *a priori* that water is identical to the stuff that plays the watery role. By observation we attribute various features to occurrences of water and of H₂O, and it is an *a posteriori* matter whether these features are always the same. The proposition is an *a posteriori* proposition, and we can argue analogously about moral A-extension propositions. We can only discover if, say, MRP is true by investigating *a posteriori* what people take the moral-rightness role to be, seeing what plays that role, and then checking whether it conforms exactly with the content of "the property of moral rightness". If the proposition is true, we can only discover that it is true by *a posteriori* investigation. If MRP can only be shown to be true *a posteriori*, though, that means that it is not *a priori* and therefore option (b) is correct. Moral A-extension propositions, if they are true, can only be verified *a posteriori*, not *a priori*.

A possible disanalogy between Mellor's original objection and the Moral-Mellor objection is that whilst we discover *a posteriori* that water is not identical to what plays the watery role, this is not the case for moral properties. With water, we make the identification of water with whatever plays the watery role *a priori*, and we make the identification of water with H₂O *a posteriori*. However, the objection claims, we make the identification of a moral property with whatever plays the "moral property" role *a priori*, and also we make the identification of

the moral property with the descriptive property that plays that role *a priori*. Moral C-extension propositions are thus *a priori* themselves, not *a posteriori*.

The moral functionalist cannot make this claim. The less important reason for this is that in Jackson's general metaphysics, he states that C-extensions are knowable *a posteriori*. The moral functionalist may choose to drop this requirement for moral C-extensions, and so detach moral functionalism from the rest of Jackson's metaphysics, but this brings us to the main reason. We ascertain moral C-extensions by examining certain moral commonplaces, and those commonplaces are discovered *a posteriori*, by observing how actual agents behave and what moral judgements they make. Moral functionalism is committed to "holding on to as [many of these commonplaces] as possible", so the moral functionalist is committed to saying that C-extension propositions are *a posteriori*. It follows that if our discovery of moral C-extensions is based on a *a posteriori* discovery of moral commonplaces, and if we can only ascertain moral A-extensions by ascertaining moral C-extensions, then the discovery of moral A-extensions is based on a *a posteriori* discovery as well.

Even if there are true moral A-extension propositions, this does not mean that option (b) is wrong. The moral functionalist hopes to identify moral properties with particular non-moral descriptive properties by moving from the *a priori* to the *a posteriori*. He starts with *a priori* moral A-extensions, and uses them to justify identifying moral C-extensions, which is done *a posteriori*. But there are other metaethicists, such as Jesse Prinz (2007), who in effect work from the other direction. They attempt first to work out *a posteriori* that particular non-moral descriptive properties are moral properties. In Prinz's case, he identifies moral properties with properties that dispose moral agents to have particular emotions.² After doing so, such metaethicists may if they want say that moral properties are whatever plays certain roles. They thus move from C-extensions ("Moral property X is identical to non-moral property Y.") to A-extension propositions ("Moral property X is whatever plays the X-role.")

² Prinz and the moral functionalist actually agree on a lot. They are both reductionist moral realists who believe that we must investigate human behaviour and attitudes *a posteriori* to discover which non-moral properties are identical with moral properties. However, Prinz gives much more importance to emotions such as guilt in his metaethics, and he certainly would not accept *a priori* moral A-extension propositions.

However, the A-extension propositions are justified by the C-extension propositions. The X-role is defined by the content of the non-moral property Y. In this case, the metaethicist can say that there are true moral A-extension propositions. However, they are not *a priori*, because they are justified by *a posteriori* C-extension propositions.

3. Conclusion

Moral functionalism claims that there are moral A-extension propositions that can be ascertained *a priori*. These propositions provide justification for *a posteriori* discovery of moral C-extensions. Two types of objection to moral functionalism's claims here are that a) moral A-extension propositions are untrue (and hence cannot justify anything), and b) moral A-extension propositions are not *a priori*.

D.H. Mellor objects that A-extension propositions are untrue – it is not the case, for example, that water is identical with whatever plays the watery role. An analogous argument can be made that moral A-extension propositions are untrue for the same reason, but this goes too fast. We cannot just declare automatically that moral A-extension propositions are false. However, we can use the objection to point out that we can only ascertain *a posteriori* that moral A-extension propositions are true. So moral functionalism's commitment to *a priori* A-extension propositions becomes a real problem. If moral A-extension propositions are true, they cannot be *a priori*.

I must point out here that I have been concentrating only on the A-extension propositions that moral functionalism relies on. Moral functionalism may have other *a priori* commitments, and I say nothing about them. However, A-extension propositions are particularly important for moral functionalism, since to Jackson they are prior to C-extension propositions. The moral functionalist thinks that even if we discover *a posteriori* that descriptive property X is moral property Y, we do so because X plays the Y role, and the identification of Y with what plays the Y role is discovered *a priori*. But the moral functionalist is wrong

on that point. We have to go out and look, just as we have to go out and look at whether water is whatever plays the watery role.³

References

Jackson, F. (1998). *From metaphysics to ethics*. New York, NY: Oxford University Press.

Jackson, F. (2003). From H₂O to water: The relevance to a priori passage. In H. Lillehammer & G. Rodriguez-Pereyra (Eds.), *Real Metaphysics: Essays in Honour of D.H. Mellor* (84-97). London: Routledge.

Jackson, F., & Pettit, P. (1995). Moral functionalism and moral motivation. *Philosophical Quarterly*, 45(178), 20-40.

Mellor, D.H. (2003). *Real metaphysics: Replies*. In H. Lillehammer & G. Rodriguez-Pereyra (Eds.), *Real metaphysics: Essays in honour of D.H. Mellor* (217-238). London: Routledge.

Prinz, J. (2007). *The emotional construction of morals*. New York, NY: Oxford University Press.

³ An earlier draft of this paper was presented at the VIII Meetings on Ethics and Political Philosophy at the University of Minho. I would also like to thank Hallvard Lillehammer and Cristian Constantinescu for their helpful comments.



GENDER IDENTITIES AND FEMINISM

IDENTIDADES DE GÉNERO E FEMINISMO

Josh T. U. Cohen*
joshcohen93@gmail.com

Abstract. *Many feminists (e.g. T. Bettcher and B.R. George) argue for a principle of first person authority (FPA) about gender, i.e. that we should (at least) not disavow people's gender self-categorisations. However, there is a feminist tradition resistant to FPA about gender, which I call "radical feminism". Feminists in this tradition define gender-categories via biological sex, thus denying non-binary and trans self-identifications. Using a taxonomy by B. R. George, I begin to demystify the concept of gender. We are also able to use the taxonomy to model various feminist approaches. It becomes easier to see how conceptualisations of gender which allow for FPA often do not allow for understanding female subjugation as being rooted in reproductive biology. I put forward a conceptual scheme: radical FPA feminism. If we accept FPA, but also radical feminist concerns, radical FPA feminism is an attractive way of conceptualising gender.*

Keywords: *feminism, gender, identity, sex, transgender, non-binary.*

Sumário. *Muitas feministas (por exemplo, T. Bettcher e B.R. George) defendem um princípio de autoridade na primeira pessoa (FPA) sobre o género, ou seja, argumentam que não devemos (pelo menos) desautorizar as auto-categorizações de género das pessoas. No entanto, há uma tradição feminista resistente à FPA sobre género, a que eu chamo "feminismo radical". As feministas desta tradição definem categorias de género através do sexo biológico, negando então as auto-identificações não binárias e trans. Usando a taxonomia de B. R. George, começo por desmistificar o conceito de género. Também podemos usar a taxonomia para modelar várias abordagens feministas. Torna-se mais fácil ver como as conceptualizações de género que permitem o FPA muitas vezes não permitem compreender a subjugação feminina como estando enraizada na reprodução biológica. Proponho um esquema conceptual: a FPA do feminismo radical. Se aceitarmos a FPA, mas também as preocupações feministas radicais, a FPA do feminismo radical é uma forma apelativa de conceptualizar o género.*

Palavras-chave: *feminismo, género, identidade, sexo, transgénero, não-binário.*

* Philosophy student at the University of Cambridge.

o Introduction

Many people identify their gender as being non-binary, so consider themselves to be neither (wholly) men nor women (Titman, 2014). Mainstream western society generally does not recognise non-binary identities,¹ and non-binary people are often subject to *misgendering*, i.e. they are categorised as a gender they do not identify as. Talia Mae Bettcher (2009) argues that misgendering is a form of transphobia, involving a refusal to recognise trans people (including non-binary people) as who they claim to be. Generally, trans people suffer greatly in western society, often facing harassment, discrimination and violence, in addition to social exclusion, high levels of stress and mental health problems (Diamond, Pardo, & Butterworth, 2011). It is therefore imperative that we listen to, think with, and be empathetic towards trans people's diverse lived experiences.

However, some radical feminists have problematised aspects of non-binary and trans identities (Cox, 2016; Reilly-Cooper, 2016). Many radical feminist critiques are relatively easy to answer. For example, they have problematised the notion that non-trans women are more privileged than trans people, and, as a result, have problematised assigned male at birth (AMAB) trans women and non-binary people having access to feminist/women-only spaces. Trans people are also thought to transition for dubious (either mistaken or menacing) reasons. For example, trans people have been thought to misunderstand their own desire to transition, or to desire transition in order to hurt individual women and/or the feminist cause. Such arguments have been addressed comprehensively (primarily on the subject of trans women) by Sandy Stone (1991), Emi Koyama (2003), and Bettcher (2014).

Other radical feminist arguments are have not been satisfactorily addressed. Such arguments hold that, if one accepts the radical feminist claim that women's

¹ There are some notable legislative exceptions and efforts to change this, some of which I will note here. In Australia, there is a legally sanctioned "X" category for indeterminate sex category, which some are trying to get re-defined as meaning non-binary gender (Ansara, Webeck, Carpenter, Hyndal, & Goldner, 2015). As of 2017, Canada allows for non-binary gender identification on legal documents (Busby, 2017). Germany's Federal Constitutional Court ruled in 2017 in favour of recognising non-binary gender identity (Bundesverfassungsgericht, 2017). In the UK, the House of Commons Women and Equalities Committee's 2016 report on Transgender Equality concluded that non-binary gender identities should be legally recognised (House of Commons Women and Equalities Committee, 2016).

oppression stems from female sexual biology, then gender self-identifications are not authoritative. They hold that non-binary identities are based on a flawed understanding of gender, and that the social and political recognition of non-binary identities undermines women's liberation.

I will engage with a number of radical feminist criticisms in order to develop an account of gender that listens to both radical feminist concerns and accepts trans identities as legitimate. I call this account radical first person authority (FPA) feminism.

1. Trans Terminology

Terminology used when discussing trans issues is often contested. §I of (Bettcher, *Feminist Perspectives on Trans Issues*, 2014) is a useful explainer. Here I will attempt to give a *relatively* neutral illustration of some key terms, many of which will be re-examined in §3.

Trans people are those who do not identify with the sex/gender they were assigned at birth,² e.g. transsexual, transgender, and non-binary people (Vincent, 2016). Someone's being trans does not indicate anything about their sexual orientation, biology, or whether they have had (or want) gender-affirming medical intervention. Non-trans people are sometimes referred to as being cisgender or cis. For example, a cis woman is someone assigned female at birth (AFAB) who identifies as a woman.

Non-binary people do not identify themselves as being men or women,³ and thus do not identify with the binary sex/gender they were assigned at birth (Vincent, 2016). "Non-binary" functions as an umbrella term. It encompasses those who identify as genderfluid (switch between genders), agender (no/neutral gender), bigender (both man and woman), as a demi-boy/girl/man/women (predominantly boy/girl/man/woman), or as some other sub-category (Vincent, 2016). Some identify with a sub-category yet not as "non-binary", or may prefer to be called genderqueer. Many non-binary people have a preference for gender neutral pronouns, most commonly they/them/their.

² I will discuss the sex/gender distinction in §3.

³ Non-binary people may identify as being men and/or women to some degree, but do not identify with either exclusively.

Stone (1991) argues that all transsexual people are beyond the gender binary. However, non-binary people (in the sense I am discussing here) are non-binary due to their non-binary gender-identities, not on the basis of any medical intervention or condition. I will discuss the nature of gender-identity in §3.

2. Feminist Principles

There is division within feminism between those who accept trans women's self-identification as women in certain politically important ways, and those who do not. A similar division occurs on the issue of accepting the legitimacy of non-binary self-identifications.

My explication of both "sides" oversimplifies diverse traditions, but for my purposes these sketches will suffice.

2.1 First Person Authority Feminism

I will refer to feminists who accept trans self-identifications as first person authority (FPA) feminists, as they believe that people have a kind of authority over their own gender categorisation.⁴ At a minimum, they hold that other people should not disavow trans people's self-identifications. FPA feminists do not necessarily agree on what someone's gender is constituted by, and most do not give explicit outlines of the metaphysics of a subject's gender or how this relates to their self-identification. What unites FPA feminists is agreement that we should generally defer to people's first personal statements about their own gender.

Bettcher (2009) argues that people should have ethically justified FPA over their own gender categorisation. She notes that consensual gendering already happens in trans-friendly subcultures. In such contexts, when someone is asked their gender they are expected to respond with a statement about who they believe they are, in some deep existential sense. One does not "have FPA over being a woman but only over one's *believing* one is a woman" (p.111). As a result of this practice, it is not ethically permissible to question someone's stated gender, as one does not have the ethical authority to deny other people's avowals

⁴ For example, Bettcher (2009), B. R. George & R. A. Briggs (2016), Katherine Jenkins (2016), and Emi Koyama (2003).

about their own mental life. To do so would be controlling, invasive and potentially harmful.

Bettcher further argues that this subcultural practice is superior to mainstream practices, in which questions about gender status are answered on account of genitalia. Bettcher holds that the culturally enforced mandatory admission of genitalia is tantamount to sexual abuse. Rather than sharing facts about our sexual biology, gender labels should communicate our self-understanding. Dominant gender practices may also be deemed immoral for avowedly feminist reasons: dominant gender practices involve subjugating women, and so should be rejected wholesale (Bettcher, 2012).

Burkay Ozturk (Forthcoming) has argued that Bettcher's notion of ethical FPA is too strong, and that there are circumstances in which one can reject someone's self-identification. He argues that if someone rejects a person's self-identified gender without harming them, without violating their privacy, and without denying their "negotiative dignity," such a denial can be permissible.

Nevertheless, consistent with Ozturk, there is good reason for *generally* deferring to gender self-identification. For example, feminism often aims to take people's lived experience seriously. Many trans people have testified that misgendering is a harm to their mental health, and that this is a harm that cannot be otherwise avoided. These testimonies should be taken seriously, especially as they are corroborated by clinicians (Wiseman & Davidson, 2011). Avoidably harming other people's mental health is not morally permissible, all things being equal, so we should avoid such harm by accepting people's gender self-identifications.

According to Iris Young's (2009) account of oppression, misgendering is oppressive. Misgendering is a form of cultural imperialism: society's "dominant meanings" make the trans community's perspective "invisible", while simultaneous stereotyping its members, and making them "the Other" (p.66). Mainstream society systematically and coercively subjects non-binary and trans people to its norms. The dominant culture assigns gender-category on the basis of genitalia, which systemically erases non-binary and trans self-understanding, as well as positioning trans people as freakish outsiders. Additionally, Young considers violence to include "incidents of harassment, intimidation, or ridicule

simply for the purpose of degrading, humiliating, or stigmatizing group members” (pp.67-68). Non-binary and trans people frequently experience such abuse, which may come in the form of purposeful misgendering (Diamond, Pardo, & Butterworth, 2011).

Furthermore, there is uptake of FPA about gender in progressive circles (George & Briggs, 2016). Many people hold that a person’s self-identification is the most ethically sound, or socially viable, indicator of their gender. At the very least, it is worthwhile to see how good a theory consistent with gender self-identification can be. Following Sally Haslanger’s (2012) methodology of ameliorative analysis, I hold that analyses should be done on the basis of political usefulness: a theory of gender should support feminist aims.

2.2 Radical Feminism

Feminists who do not accept self-identifications as authoritative often self-identify as radical feminists, which is how I will refer to them.⁵ Radical feminists argue that trans people’s self-identifications should not be authoritative, regardless of social or medical transition. Their approach to trans issues is predicated on their theory of gender, which I will summarise in relatively neutral language.

Most babies are assigned male at birth (AMAB) or assigned female at birth (AFAB) in accordance with their genitals.⁶ As children grow, they learn gender roles from their parents and wider society. Divergent norms are instilled in them: AMAB are trained to be masculine men, i.e. domineering, individualistic, and aggressive, while AFAB are trained to be feminine women, i.e. submissive, cooperative, and kind. Radical feminists understand “men” and “women” to be the two sex-classes/castes/roles in the patriarchy, i.e. the dominant gender system. These categories may be referred to as “gender- classes/castes/roles”, or “genders”.

⁵ E.g. (Raymond, 1994), (Jeffreys, 2014) and (Reilly-Cooper, *Sex and Gender: A Beginner's Guide*, 2015). Of course, many FPA and/or trans feminists may be radical feminists in the broader sense, or self-identify as radical feminists too. Many canonical radical feminists, e.g. Catherine McKinnon, accept trans identities (Williams, 2015).

⁶ Radical feminists rarely use terms such as AFAB or AMAB, as many believe categorisation as male or female is non-problematic and assignation independent.

The sex-classes are hierarchical: men are privileged and women are oppressed. Because of their sexual biology, women are expected to be available for male sexual satisfaction, gestate children and raise them. Heterosexuality is compulsory. Women are expected to perform domestic labour and raise children without formal remuneration. This inequality in domestic labour division continues despite women joining the formal workforce (Jones, 2016). Women's oppression is intrinsically linked to being AFAB: it is the recognition (or assumption) of their sex organs that lead to membership in the subordinate sex-classes/genders.

Gender membership is seen as imposed and inherently oppressive. Radical feminists aim to decouple assigned sex from gender roles. Slogans such as "biology is not destiny" communicate that female people need not occupy traditional roles, and should not be forced or coerced to do so. Distinguishing being AFAB/classed as a woman from performing a women's role/acting in a feminine way is of central importance to radical feminists.

Although I do not have space to investigate the claims made by radical feminists, I will take understanding female subjugation as stemming from traditional reproductive roles to be (at least) useful for feminism: This analysis seems to have advanced women's liberation, and it seems likely that it will continue to do so.

3. Unpacking Gender

Central to the issue of trans/non-binary identities in feminism is the notion of gender. The category "woman" is often seen to be some social category, i.e. a gender, and not a biological category, i.e. a sex. The sex/gender distinction has been problematised by many feminists (notably Judith Butler (1990)), but here I will assume that it is, at least, useful. Even if there are many people with a variety of intersex conditions, the vast majority of people are unproblematically assigned a sex at birth. Even if the biology of sex assignment is post-discursive, socially constructed and culturally informed, it is dependent on different sets of facts from gender (e.g. genitalia and/or chromosomes vs. pronouns and/or gendered clothing), and it is thus distinguishable from whatever "gender" is.

Radical feminists define genders as a class/caste related to biological sex, but trans people hold that their “gender” is not determined by their biological sex. Trans people make another element of gender salient: gender-identity. Some FPA feminists claim that it is gender-identity that *makes* one a woman, or non-binary, or a man (Cohen, 2017). Additionally, the notions of gender expression, performance or presentation are important. These notions capture what an individual *does* that is gendered, i.e. how they dress or to act in gendered ways.

B. R. George (2016) clarifies many ambiguities within gender talk with a taxonomy of gender concepts. It notates the loose clusters of characteristics, behaviours and labels important to our notion of “gender”, and the relations between them. There are three nodes: sexed-biology (**s**), gender-practice (**p**), and gender-category (**c**). **s** unites the biological characteristics that a society takes to indicate reproductive roles, for example beards, ovaries and XY chromosomes. **p** is constituted by behaviours which a society takes to be gendered, e.g. raising children, wearing a suit, or being assertive. **c** contains the various sex/gender categories a society may place people in, e.g. female/male and woman/man, and also cisgender, non-binary, transgender, intersex, etc.⁷ The distinction between these nodes can be messy: some theorists (such as Butler (1990) and Jennifer McKittrick (2015)) hold that gender-classes are in some way constituted by practice. However, it is possible to distinguish between category and practice, even if one is constituted by the other, and most of the theories explored within this paper do not hold that this is the case anyhow.

⁷ The notion of gender-category is distinct from the radical feminist’s sex-role/class/caste. It includes categories that they leave do not, and is not *necesarrily* heirarchical. Thus, I call them “gender-categories”, not “gender-classes”.

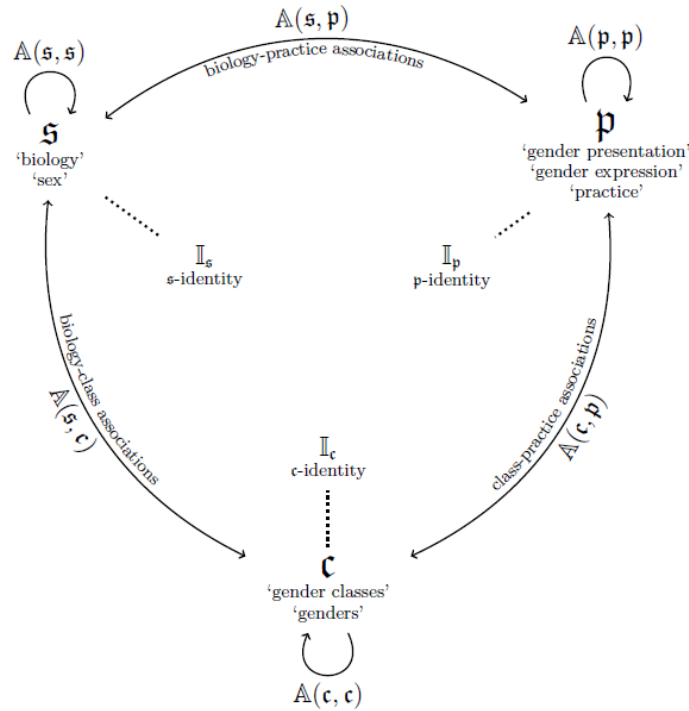


Figure 1. George's Taxonomy

There are six families of associations, notated $A(x,y)$, which are stereotypes, norms or rules. Inter-node associations connect the three nodes. $A(s,p)$ are associations connecting sexed-biology, s , to gender-practice, p . For example, “bearded people should be assertive,” or “people with ovaries should raise children.” $A(s,c)$ are associations connecting sexed-biology, s , to gender-category, c . For example, “women have ovaries,” or “people with XY chromosomes are male.” $A(p,c)$ are associations connecting gender-practice, p , to gender-category, c . For example, “people who are male should have short hair,” or “people who are sensitive are female.” There are also intra-node associations. $A(s,s)$ are associations between different aspects of sexed-biology, s . For example, “people with testicles don’t have ovaries.” $A(p,p)$ are associations between different gender-practices, e.g. “if you care for children you should not wear a suit.” $A(c,c)$ are associations between gender-categories, c , e.g. “men are male” or “females are not men.” Associations are reflexive, i.e. $A(x,y) = A(y,x)$.

Additionally, there are gender-identities/inclinations, notated I_x , for each of the respective nodes (s , p , and c). Gender-identities are not about the recognition of other people’s perceptions or categorisations, but are instead about how one would, or would like to, perceive or categorise oneself, and be perceived and

categorised by others. Sex-identity, I_s , involves attitudes towards or identifications with aspects of sexed-biology. For example, I_s includes wanting to have large breasts, or disliking one's penis. Practice-identity, I_p , contains attitudes/identifications with gender-practice. Wanting to be a nun or believing oneself to be assertive are examples of I_p . Category-identity, I_c , involves attitudes/identifications with gender-categories. Examples of I_c include identifying as male, wishing one were a woman, or feeling oneself to be agender. Category-identity is what I have previously been referring to as "gender-identity".

George's taxonomy omits associations between gender-identities, and omits associations between gender-identities and nodes s , p , and c .⁸ I include $A(c, I_c)$, $A(p, I_p)$ and $A(s, I_s)$, i.e. associations between category and category-identity, practice and practice-identity, and sex and sex-identity. Examples of these include: all those who identify as women are women, those who want to wear dresses should wear dresses, or that people should accept their sexed biology. I also include associations connecting gender-identities, such as $A(I_c, I_p)$, e.g. those who want to act effeminately should identify as women, or $A(I_c, I_s)$, e.g. those who want to have a penis should identify as men.⁹ I include these additions in my revision of George's taxonomy, which I take to model mainstream gender practices:

⁸ §4.4 of (George, 2016) allows for such additions to the taxonomy.

⁹ There are plausibly associations connecting gender-identities to other nodes, such as $A(I_c, s)$, but including them is unnecessary for my purposes.

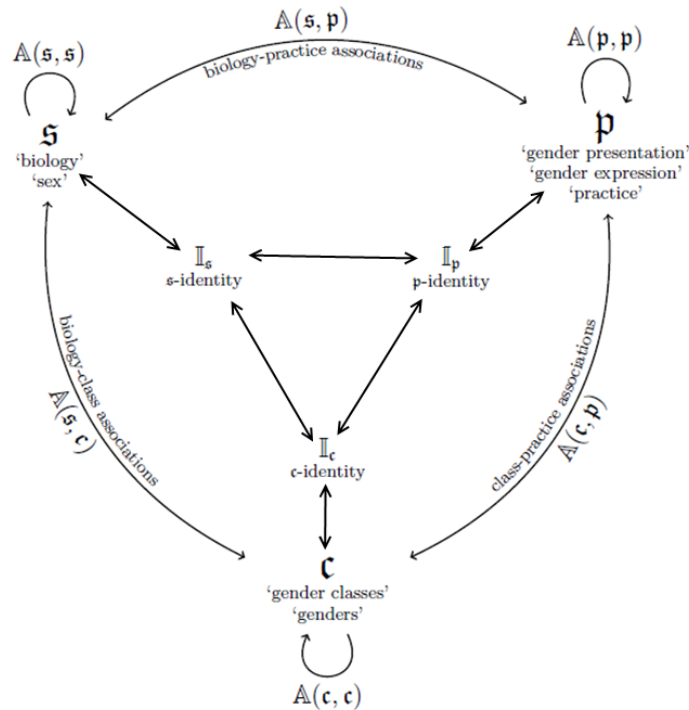


Figure 2. My additions to George's Taxonomy

4 Applying the Taxonomy

The notation of the taxonomy does not enable us to capture many dimensions of our experiences of gender, e.g. the difference between recognising the existence of an association and advocating an association, but it can nonetheless be used to clarify several things. When speaking about gender, people may be taking about any one (or combination) of the 3 nodes, 3 identities, and 18 associations.

I will use the taxonomy to illustrate the multiple possible dimensions of trans-ness, as I understand it, and to illustrate different models of feminist thought.

4.1 Types of Trans

Trans activists differentiate between “body dysphoria” and “social dysphoria” to create a more fine-grained account of trans experience (Callahan, 2014). However, the taxonomy allows us to make sense of 3 different trans gender-identities, i.e. gender-identities that conflict with assigned sex/gender.

First, one's sexed-biology may include having certain biological traits (e.g. a vagina), with a sex-identity wishing to have different traits (e.g. a penis). This is what is meant by "body dysphoria", and I will call this trans-sex. Second, one may be raised to practice gender in a certain (e.g. feminine) way, but want to act in a different (e.g. masculine) way. This is a component of "social dysphoria", and I will refer to this as being trans-practice. Third, one may be assigned a certain sex/gender category from birth (e.g. be AMAB), while having a conflicting category-identity (e.g. identify as a woman, or as non-binary). This is another component of "social dysphoria" and I will refer to this as being trans-category.

When speaking about "trans" people, I have been speaking about trans-category people, and only trans-sex and trans-practice people insofar as they are incidentally trans-category. With my additional inter-identity associations (i.e. $A(I_c, I_p)$, $A(I_c, I_s)$ and $A(I_s, I_c)$), one can see how social norms may lead pressure who are trans in only one respect to be trans in all respects. For example, it is plausible that some trans-category AFAB people who identify as men may feel the need to identify with trans-practice masculine activities in order to legitimise their category-identity in the eyes of wider society. This is not to pathologise trans experiences, but merely to note that trans people are subject to the same pressure to conform all three nodes and respective identities as cis people.

Under my definitions, many people potentially count as trans who would not under most conceptualisations of trans-ness (Overall, 2012). For example, cis-men who want penis enlargement surgery and cis-women who receive breast reduction surgery have trans-sex traits, while all non-heterosexuals have trans-practice traits. This difference between my conception and the more ubiquitous, exclusive conceptions of trans-ness may appear to be a weakness of my approach. However, under my approach we are best able to recognise and respect the nuanced multiplicities of self-reported trans-experience, whilst simultaneously allowing for a more fluid, scaling and less othering relationship between those who are generally understood to be trans, and those generally understood to be cis.

I do not have space to explore the relation between people's trans-x traits and their status as being trans-x. Most people are trans-x in some way, to some degree. Few people are happy with *every* aspect of their sexed-biology or assigned

sex/gender. However, in this essay, I am interested in people who are trans-x to a high degree.

4.2 Types of Feminism¹⁰

The taxonomy can help us model a variety of possible feminist approaches. As understood here, feminism is united in rejecting $A(s,p)$ and $A(I_s,I_p)$ associations connecting sexed-biology with practice. It is uncontroversial (as a feminist) to hold that reproductive biology should not determine social-roles, behaviours or dress codes (or, indeed, the reverse). However, if we were to attempt a minimalist feminism and *only* reject $A(s,p)$ and $A(I_s,I_p)$, then sex would still determine practice. This is due to the combination of $A(p,c)$ and $A(s,c)$: even if we dispose of the norm “people with testicles should be dominant”, by retaining the norms “people with testicles are men” and “men should be dominant”, we are left with the same result. I call this the Problem of Transitivity. In order for an effective feminism, it must be overcome.

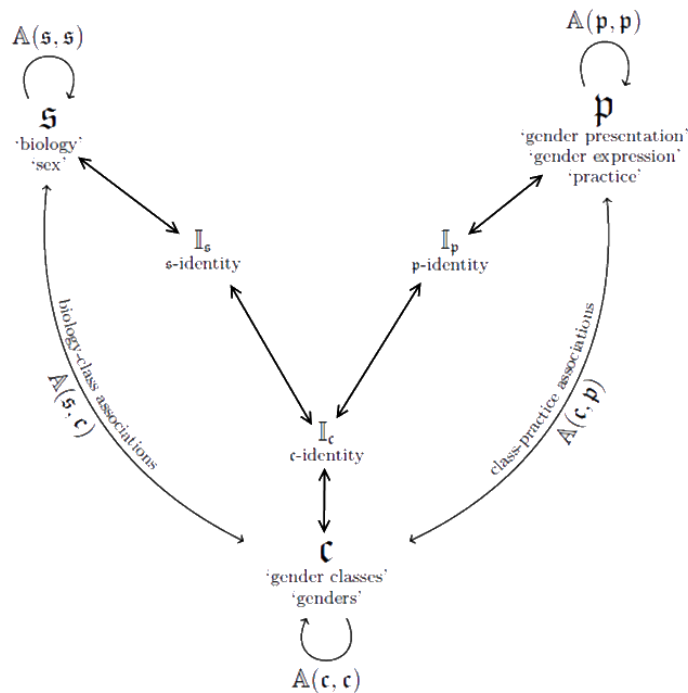


Figure 3. Minimal Feminism

¹⁰ These models are oversimplified caricatures, constructed for primarily discursive reasons. I do not mean to imply that everyone I link to a type of feminism advocates every aspect of that type. Additionally, it is worth noting that (likely) all feminists would agree that the diagram at the end of §3 (for the most part) correctly *describes* current practices. The *key* differences between the feminisms I present are which associations they *prescribe*.

All feminists would celebrate trans-practice rejections of $A(s,p)$ associations – the freedom to be trans-practice is an essential aim of feminism.¹¹ However, radical feminists reject trans-category people's claims that they *are* members of their preferred categories. Some radical feminists have also criticised trans-sex people, but this has only been insofar trans-sex people have used biological/medical changes to legitimise trans-category identities.¹² Similarly, trans-practice people have been problematised for using practice to legitimise trans-category identities.

4.2.1 Radical Feminism in the Taxonomy

Radical feminists solve the Problem of Transitivity by rejecting $A(p,c)$ norms connecting category and practice. Radical feminists retain $A(s,c)$ associations connecting biology and gender-category, e.g. women produce ova, males have testicles, etc. They do so because they believe that women are identified and thus oppressed on the assumption of their reproductive role. Babies are assigned membership in the subjugated “woman” gender-category on account of their genitals. If womanhood is detached from biology (or assumed biology), then issues relating to female reproductive roles (concerns such as contraception, abortion, tampon-taxes, childcare, etc.), can no longer be conceptualised as “women’s issues”. Therefore, they hold that correct analysis of patriarchal society recognises the existence of norms defining womanhood via reproductive biology.

¹¹ Rejecting $A(s,p)$ associations is implicitly trans-practice, because gender-practices are imposed on people on the basis of their sexed-biology.

¹² For example, Sheila Jeffreys (2014) and Janice Raymond (1994).

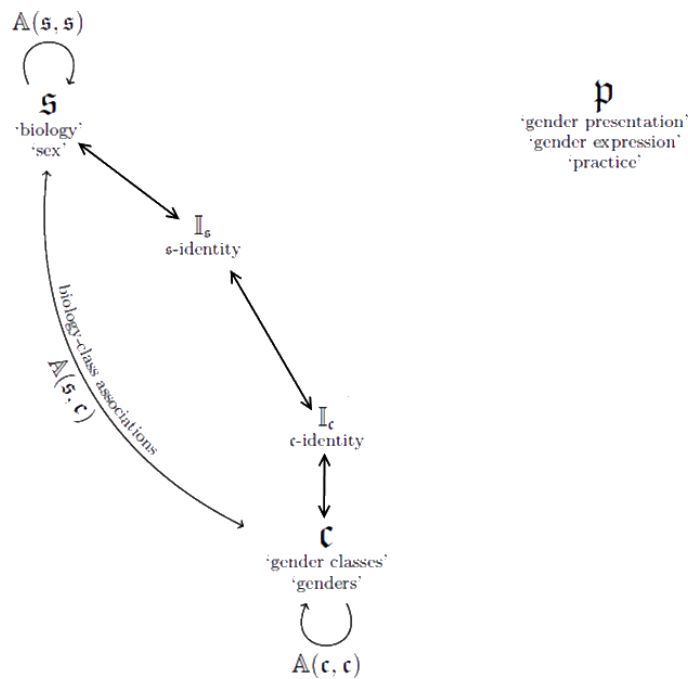


Figure 4. Radical Feminism

Many radical feminists see $A(s,c)$ associations as being non-normative matters of definition – females are defined as those with certain biological features, and women are, by definition, adult females (Jeffreys, 2014; Reilly-Cooper, 2015). This definitional approach has been criticised for not accounting for the social aspects of womanhood (Bettcher, 2009; George & Briggs, 2016). A radical feminist may respond that they do account for such aspects, but that these aspects are what they are fighting to abolish. If traditionally feminine, subordinating practices are contained within the concept of womanhood, then there can be no free women. This approach has famously been taken by Haslanger (2012), but notably her use of “female” would similarly come under attack for not accounting for the social aspects of female-ness.

Anyhow, the radical feminists stipulating a definition of “woman” that tautologically determines the outcome of a political debate is suspicious, even when it is ostensibly “non-theoretical”. Radical feminists are better off staking a claim in the political debate, rather than pretending that it does not exist. It is most charitable towards radical feminism to understand the radical feminist

holding of $A(s,c)$ as politically motivated: the motivation is addressing the connection between women's oppression and their reproductive biology, class consciousness must centre the means of reproduction. Many radical feminists aim to eventually dismantle gender-categories altogether, and thus create a wholly non-binary society: $A(s,c)$ associations are pragmatically and temporarily advocated as a means towards this end. Until this end is met, radical feminists advocate a clear grasp on who is where in the gender class system, and they argue that this is determined on the basis of sexed-biology.

Radical feminism does not accommodate FPA about gender or the existence of non-binary genders. The only gender-categories to be recognised in their analysis are those imposed by society. So, as previously noted, radical feminism leads to misgendering. As we have reason to think that misgendering is problematic, we should attempt to find an alternative approach: FPA feminism.

4.2.2 Associative FPA Feminism

FPA feminism centres a norm of accepting gender-category self-identification, which is an $A(I,c)$ norm. In order to respect trans self-identifications, all FPA feminists reject defining womanhood/manhood via $A(s,c)$ sex-category associations.

It is *prima facie* plausible that for gender-categories to be politically meaningful, they must have some steady, politically potent attributes. As FPA feminists reject $A(s,c)$ sex-category associations, some FPA feminists may accept practice-category norms $A(p,c)$ and define "women" and "men" as categories demarcated by certain practices. Thus, they hold that men should be defined as those who (want to) act in a certain way, and women should be defined as those who (want to) act in a certain way. At their most extreme, they may conceptualise butch lesbians as trans men in denial, and effeminate gay men as trans women in denial (crashchaoscats, 2013). I call this position associative FPA feminism.

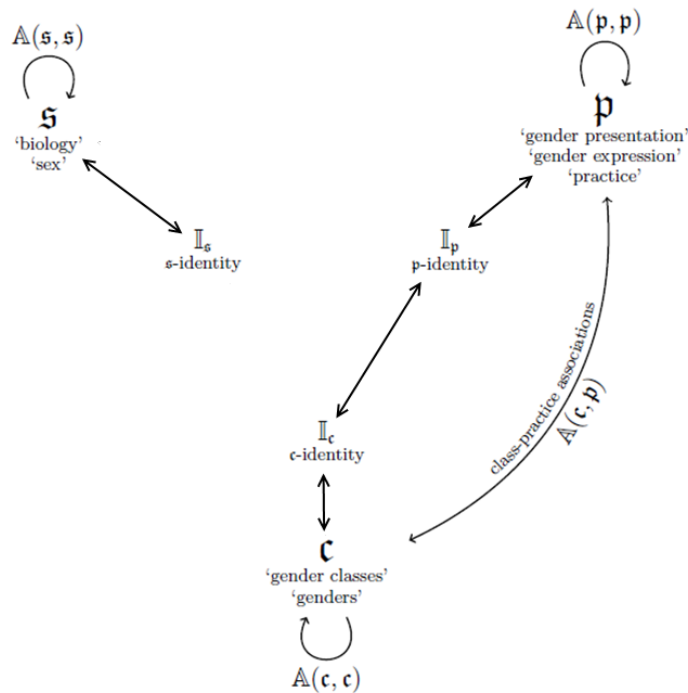


Figure 5. Associative FPA Feminism

Associative FPA feminism is the (arguably straw) position many radical feminists implicitly attribute to trans activists and FPA feminists. Some trans people have ostensibly endorsed this type of view autobiographically.¹³ It is on the basis of associative FPA feminism that radical feminists make the first argument against non-binary identities that I will discuss.

Radical feminists argue that non-binary identities create a new binary between non-binary and binary gendered people (Cox, 2016; Reilly-Cooper, 2016). If, as associative FPA feminism holds, those who do not wholly subscribe to a binaristic gender practice should identify as non-binary, then those who identify with a singular, binary gender are normatively constrained by their identification with that gender. The non-binary category implies that *only* non-binary people may escape binary gender norms – everyone else must be a traditional man or woman. In particular, AFAB non-binary people have been thought to be motivated by an attempt to escape the oppressive norms of womanhood. If non binary people’s non-conformity with their assigned

¹³ For example, Tyler Ford (2015) and Jack Monroe (2015).

sex/gender is not coupled with identification as that sex/gender, they do not challenge gender norms. Instead, non-binary identities multiply the number of gender-categories without challenging their normativity.

An associative FPA feminist may counter that they *are* feminist despite accepting $A(p,c)$ norms. They avoid the Problem of Transitivity by rejecting sex-category associations, and therefore do not hold that sex dictates destiny. Thus associative FPA feminism allows anyone to do whatever they want, as long as they choose the correct category-identity. In an ideal associative FPA feminist world, AMAB people, intersex, and AFAB people are equally able to choose between male, female and non-binary gender-identities, and thus are free to live whatever lives they want, in whichever way they wish.

However, something about associative FPA feminism is not right. People are not allowed to simply live within their assigned gender while not conforming to it. In the associative FPA feminist world, one cannot unproblematically be a woman while simultaneously being a short haired, muscular welder. Such a person would contravene the $A(p,c)$ norms dictating that women should be traditionally feminine. If we are to believe (as many FPA feminists do) that gender-identities are not entirely voluntary, this is obviously problematic. Either people must opt for gender-identities that conflict with the one they actually feel comfortable with, or live up to norms that they may not affirm. Such an approach would misgender those who are trans-practice but not trans-category, such as butch women. Hence, associative FPA feminism doesn't fulfil its aims: it does not accept people's category-identities, come what may.

In response to the radical feminist argument, it is notable that non-binary people are likely aware that "coming out" will increase their tension with prevalent gender norms, rather than ease their lives under the patriarchy. Additionally, some non-binary people report not feeling "trans enough", and therefore feel they are failing to uphold norms associated with *their* genders (Vincent, 2016). It seems that being non-binary is not a way of escaping gender roles, as non-binary people may still feel constrained by certain gendered expectations. These points illustrate that associative FPA feminism is flawed, but also that the radical feminist argument against non-binary identities is misguided.

4.2.3 Jenkins' FPA Feminism

While some trans people may define their category-identities via association with their gender-practice, this is not a necessary facet of being trans. Some non-binary people practice gender in a seemingly cisgender fashion, e.g. there are non-binary AMAB people who dress in a masculine way. One's gender-identity and practice may diverge: a trans man, by which I mean someone who is AFAB and identifies as male, may dress and act in a traditionally feminine manner.

Associative FPA feminism is not the only conceivable FPA feminism. Katharine Jenkins (2016) argues that we should understand gender-identity to be a person's recognition that certain practice-category norms are applicable to them, even if they choose to reject those norms. For example, someone has a woman's gender-identity, in part, if they understand the norm that women should shave their legs to apply to them, even if they choose to reject it. Therefore, Jenkins allows for people to act in ways that are not traditionally associated with their gender-identity. Additionally, Jenkins argues that it is best to categorise people on the basis of their gender-identity, so her feminism is a kind of FPA feminism.

While Jenkins rejects $A(p,c)$, $A(s,c)$ and $A(s,p)$ norms, her notion of gender-identity recognises the current existence of $A(p,c)$ norms. Her notion of gender-identity is $I_{A(p,c)}$, i.e. an identification with a practice-category norm.

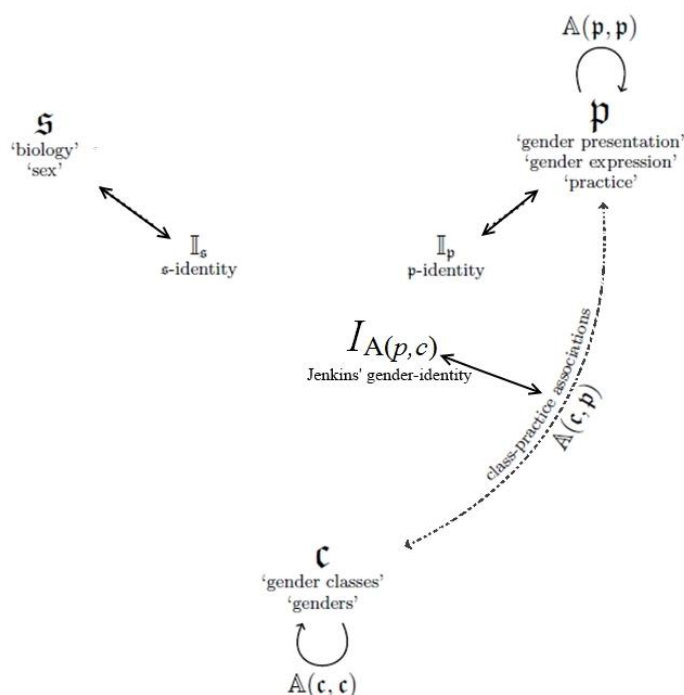


Figure 6. Jenkins' FPA Feminism¹⁴

However, there are problems with Jenkins' approach. According to Jenkins, a trans woman has a female gender-identity if she understands certain dominant cultural norms to apply to her. However, an essential element of these norms' content is *who* the norms are applicable to. The norms do not dictate that anyone should do *x*, but rather that some specific people should do *x*. As Bettcher (2007; 2009) notes in an argument Jenkins' theory relies on, in mainstream society "women" is elliptical for "people with vaginas". Norms directed at "women" by mainstream society are actually directed at those with vaginas. Therefore, a trans woman would be incorrect in understanding such norms to apply to them, regardless of their preferred gender-category. In accordance with mainstream practices, the target of the norm is individuated on account of their biology, not their self-conception. Jenkins' analysis, against her intentions, implies that there is something incorrect about trans gender-identities: the trans person who thinks that a norm directed at "women" refers to them has simply not understood the norm properly. This is problematic because FPA feminists do not wish to label

¹⁴ This model is intended to *only* illustrate Jenkins' notion of gender-identity, and nothing about her understanding of other norms, gender-class or sex.

trans gender-identities as being incorrect – this would be a version of what Bettcher (2009) calls the “Basic Denial of Authenticity” (p.99), a fundamental form of transphobia.

Ignoring this, the association between category membership and practice in Jenkins’ account remains troublingly important. The account implies that the radical feminists’ argument in §4.2.2, that non-binary genders strengthen gender norms, is correct. If gender-identity is about gender norms in the way Jenkins argues, then becoming non-binary is a legitimate way to escape restrictive norms. And if becoming non-binary is the best way to escape such norms, then AFAB non-binary people really are throwing women “under the bus”, by identifying-out of the oppressed “women” category, instead of collectively counteracting those norms (Cox, 2016).

Furthermore, Jenkins’ analysis makes gender-identities dependant on $A(p,c)$ practice-category norms. This means that without $A(p,c)$ norms, which we have shown to be problematic, there is no gender-identity. So gender-identities are dependent on sexism. This is an unnecessarily strong claim, and a notion of gender-identity without this baggage would be preferable.

4.2.4 Dissociative FPA Feminism

Jenkins takes us just short of dissociative FPA feminism, which holds that gender-category and category-identity should be completely unrelated to gender-practice (Reading, 2014). Dissociative FPA feminism holds that gender-categories, e.g. “women”, “men” and “non-binary”, should not be demarcated on the basis of biology or practice: being a woman, man or non-binary person implies nothing about your biology, how you dress, or how you act. In terms of the taxonomy, it rejects inter-node associations between category, sex, and practice. It also rejects reflexive practice-practice, category-category and sex-sex associations, *and* the inter-identity associations between category-identity, sex-identity and practice-identity.

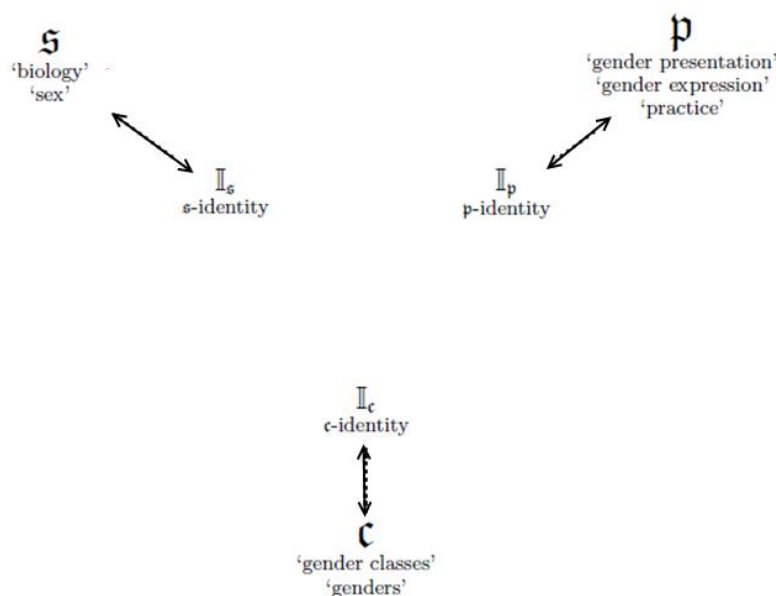


Figure 7. Dissociative FPA Feminism

The only prescribed associations are $A(c, I_c)$, $A(s, I_s)$ and $A(p, I_p)$, connecting nodes to their corresponding identities. These associations allow for norms supporting trans-practice, trans-sex and trans-category people. Under dissociative FPA feminism, being part of a gender-category has nothing to do with reproductive organs, clothing preferences or societal roles, and identifying yourself as part of a gender-category doesn't involve advocating any specific sets of norms about that category (other than $A(I_c, c)$ ones).

Notably, *completely* ruling out $A(p, c)$ norms is not practical for FPA feminism. FPA about gender-categories necessitates *some* category-practice norms. These norms connect categorisation with public self-identification, e.g. saying "I am a woman", or asking for non-binary pronouns. I take these elements of practice to be the minimal real-world manifestations of gender-category. I will speak of retaining *only* these minimal category-practice associations as equivalent to rejecting category-practice associations.

5. Radical Critiques of Dissociative FPA Feminism

There are several criticisms radical feminists would likely aim towards dissociative FPA feminism (Cox, 2016; Reilly-Cooper, 2016). I will outline them here, before presenting an approach that aims to pacify these concerns: radical FPA feminism.

5.1 What are genders?

If categories are not governed by norms connecting them to sex or practice, then in what sense are they “genders”? There are many different ways to categorise people, and, seemingly, gender-categories should have something to do with sex. If “non-binary”, “women” and “men” are not categories on the basis of biology or practice, it is hard to understand why we should consider them genders, or how they cohere into politically important groups at all. Categories cannot be defined by self-identification alone or they will fall victim to logical regress (Bettcher, 2009).

5.2 Innateness and Phenomenology

According to dissociative FPA feminism, having a certain category-identity implies nothing about biological sexual characteristics, or about the acceptance or rejection of masculine/feminine gender practices. So what is the claim to a non-binary, male or female category-identity about? Some non-binary or trans people may explain or justify their gender-identity by stating that they do not “feel” like either a man or a woman. This problematically implies that all men and women “feel” a certain way. Radical feminists are sceptical of internal senses of essential gender, and are sceptical of claims that it is normal to be comfortable with one’s assigned gender. Similarly, some trans people claim that their gender-identity is innate, ostensibly implying that everyone has an innate gender-identity (Cohen, 2017). If this is true, then it seems to justify differences between genders, and thus the oppressive patriarchal system. Additionally, some radical feminists hold that any scientific basis for innate gender has been debunked (Fine, 2011).

5.3 Political Importance of Biology

Reproduction divides humans into at least two arguably-vague-but-distinguishable groups and radical feminists argue that sex-based categorisation

is the foundation of the patriarchy. Thus an effective feminism cannot shy away from the importance of sexed-biology to women's oppression (Reilly-Cooper, 2015). Dissociative FPA feminism erases the ability for radical feminists to organise women on the basis of sex, and thereby leaves women unable to group together on the basis of the cause of their oppression. By removing the relationship between biology and gender-categories, dissociative FPA feminism obscures the fundamental axis of patriarchal oppression, muddying the waters for anyone wishing to tackle it. Only with a clear grasp on who is oppressing (i.e. "biological men"), who is oppressed (i.e. "biological women") and the means by which they are differentiated (i.e. their reproductive biology) can feminist aims be fully realised.

5.4 Ethical Motivations

If dissociative FPA feminism is correct, it seems that people want to be in arbitrary categories, for arbitrary reasons. If nothing defines the boundaries of "woman", "non-binary" or "man", why should we concern ourselves with who we verbally place in each category? Considering that radical feminists connect the word "woman" to female biology for legitimate, political reasons (i.e. because women's oppression happens on the basis of that biology), can FPA feminism be justified in brushing the radical feminist definition aside? It seems there must be some ethical motivation for choosing an FPA conception of gender over the radical one, especially as I have chosen the achievement of feminist ends to be a key desiderata in theory selection.

6. Radical FPA Feminism

Here I give a brief outline of radical FPA feminism, which develops on dissociative FPA feminism in order to address the radical feminist worries.

6.1 Categories are Historical

We can demarcate gender-categories meaningfully, even without using normative or definitional relationships with practice or sex. B.R. George and R.A. Briggs (2016) have argued that categories such as "women", "men" and "non-binary" are genders due to their *historic* relation to sex-based categories. They argue that there are two Primordial Genders, with membership based on perceived sexed-biology. Our current binaristic gender-categories should be

considered to be Historical Genders, descending from these Primordial Genders. While the Primordial Genders have necessary conditions for membership on the basis of (perceived) biology, the Historical Genders that correspond to our gender-categories need not. Additionally, they need not have any necessary conditions for membership: the categories are individuated via their histories. Non-binary categories can be understood as Symbiotic Genders, which are categories defined primarily in terms of Historical Genders.

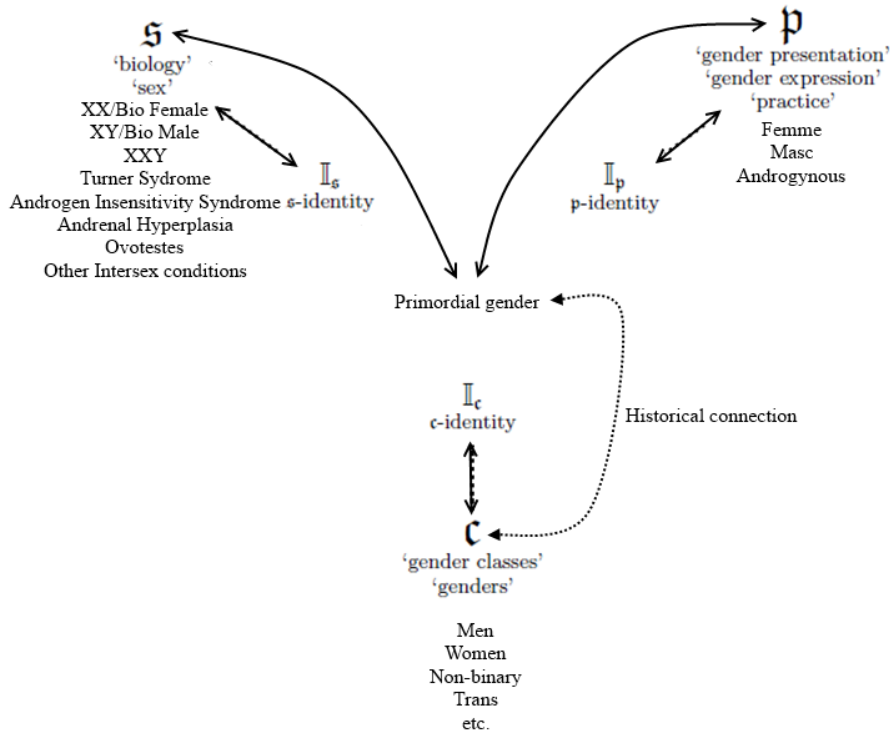


Figure 8. Radical FPA Feminism

Contra Briggs and George, radical FPA feminism demarcates gender-categories as those descended from Primordial Genders that are defined in relation to *both* perceived sexed-biology *and* gender-practice. Otherwise, non-human animals may have gender-categories. Additionally, this allows historic sex-based subjugation to be a defining element of the Primordial Genders. It is notable that this conceptualisation of gender is only broadly relevant for western cultures, as arguably some non-western societies have ancient non-Primordial Genders. Anyhow, to be a woman, man, or non-binary person today in the west, according to radical FPA feminism, is to be part of a historically individuated

gender-category, historically related to sex and practice. As the gender-categories do not have necessary conditions for membership, they allow for FPA.

Gender-categories under radical FPA feminism are politically potent. There are material circumstances connecting all self-identified women or non-binary people – ones that arise from identifying yourself as part of a historically individuated category. For example, the internalisation of stereotypes may affect all people self-identified with a particular gender (McKinnon, 2014).

It seems plausible that if all $A(p,c)$ and $A(s,c)$ associations are not only rejected, but also fade from memory, these categories would disappear. Unlike Jenkins' account, radical FPA feminism is not committed to this supposition, but it is compatible with it.

6.2 Innateness and Phenomenology

Radical FPA feminism is not committed to holding that gender-identities are innate or phenomenally distinct. Someone's gender-identity, at any one time, is simply their preference about which gender-category they would like to be identified as, at that time. Radical FPA feminism holds that we should, in general, treat people as being members of the gender-category they would like to be identified as.¹⁵

It is worth noting that many trans people have used narratives of innateness or "feelings" in order to legitimise their identities to the outside world (Stone, 1991). Such claims are somewhat dependent on the transphobic societies they take place within. Many trans people are critical of claims to innate gender-identities, and innate gender is not implied by trans or non-binary category-identities (Bettcher, 2014; Reed, 2013). Some people move between holding non-binary and binary category-identities, and so they would not claim innate or essential gender-identities (Vincent, 2016). Anyhow, beliefs in innate gender-identities are not particular to non-binary/trans people, cisgender people may hold that their gender-identity is innate. Thus, singling out non-binary people for holding this belief, or non-binary identities for propagating this belief, is suspect.

¹⁵ Whether one's gender-category membership is constituted by or dependant on this preference is a metaphysical question that I am not presently able to explore. Rather, this embryonic version of radical FPA feminism merely instructs us to treat people according to their preference.

Claims about “feeling” like/unlike a gender are epistemically unwarranted and politically unhelpful if taken literally. However, often such claims metaphorically communicate that a person does not feel *comfortable being described as or referred to as being* a woman/man. Rather than implying a widespread phenomenology of gender, such claims are about an individual’s reaction to a society’s particular gendering of them. The discomfort many cis people feel towards their gender is different: it is to do with the norms imposed on their category, not the gender-categorisation itself.

6.3 Practice-Based and Sex-Based Kinds

In order to combat sex-based oppression, our language must be able to reflect sexual dimorphism. Additionally, grouping those with similar behaviours (e.g. effeminate people) may be useful in various struggles against oppression. While radical FPA feminism holds that gender-categories such as “women”, “men” and “non-binary” must not be defined in terms of sex or practice, it groups people who share biological or behavioural traits. I will call these groups “kinds”. According to my definition of gender-categories in §6.1, these behavioural and biological kinds are not gender-categories, because they do not have links to *both* sex and gender.

It may seem that we should use “male”/”female” to refer to biological kinds, and “women”/”men” to refer to cultural kinds. This is unsatisfactory. “Male”/”female” and “women”/”men” are too linguistically intertwined for this kind of division of language to be graceful. “Woman” is defined as adult female; “man” is defined as adult male. Many sentences are rendered clumsy if the division takes place in this way, because “female” and “male” are taken to be adjectival counterparts to “woman” and “man”. The central concern of radical FPA feminism is for these gender-categories to *not* be associated with practice or biology, so we must find different means of categorisation. Here I will make some brief remarks about potential kinds.

Within the queer community, people who identify with traditionally feminine modes of practice are referred to (and often identify as) “femmes”, regardless of their biology or gender-category. Such classificatory tools would be useful in wider discourse – allowing us to describe those exemplifying a cluster of gender-practices without implying anything more. “Masc” is the opposing binary

label, and “feminine” and “masculine” the appropriate adjectives. Non-binararistic kinds must also be conceptualised. A kind such as “androgynous” is a starting point, but may be deemed inadequate for categorising all manifestations of non-binaristic behaviour. There are complex differences between some non-binary gender-categories, i.e. some are not merely on a spectrum between male and female but instead completely reject the spectrum. Analogously, “androgynous” indicates a set of behaviour on a scale between masculinity and femininity, so in order to conceptualise behaviour that does not fit on this scale more non-binary behavioural kinds may be necessary.

If we accept the radical feminist assumption about female oppression, we need a way of categorizing those with wombs, vaginas, and XX chromosomes, and their XY and intersex counterparts.¹⁶ Possibilities include “biological females”, “female-bodied”, “female-read”, “AFAB”, “natal females”, “XX people”, and “vagina/womb-havers”. Each of these have pros and cons. Retention of “female” as a component part allows a neater, more practical continuation with current linguistic practices. However, such options are misleading. The terms “male” and “female” belong to a supposedly unrelated node of the gender taxonomy, i.e. gender-category. As I am holding that “male” and “female” should *not* be understood to be about biology, it is not clear what such phrases would even mean. According to radical FPA feminism, being AFAB or biologically/natally female implies nothing about whether someone is female. Furthermore, and perhaps more importantly, many trans people reject such labels (Yamaguchi, 2016).

Alternatively, there are the biology-referencing options, such as labelling one group “XX” or “uterus/vagina-havers”, another “XY” or “penis/testicles-havers”, with a number of intersex categories for those who have atypical chromosomes, genitals, or combinations thereof. However, some radical feminists have held that naming those with female sexual characteristics anything other than “female” is erasure, and that radical feminists’ self-identification as “female” (understood biologically) should be respected (Reilly-Cooper, 2015). As a result, it may seem that the “female” related labels may be better after all.

¹⁶ I take “intersex” to include trans-sex people who have medically transitioned in some way.

Puzzlingly, different kinds may be more suitable in specific contexts. “AFAB” and “female-read” recognises the social salience of other people’s categorizing judgements. “Biologically female” indicates that people are being grouped on the same basis as other organisms: reproductive biology.¹⁷ In most contexts, however, it should be accepted that “female” refers to a historically individuated gender-category. I am not able to conclusively name sex-based kinds here: more research must be done to find appropriate labels.

As radical feminists argue, these kinds are politically significant. Those in (or assumed to be in) the “biologically female”, “XX” or various intersex kinds are systematically disadvantaged by the patriarchy. Those falling within these kinds have the right to organise amongst themselves and seek justice.

6.4 Ethical Motivations

There are many moral justifications for feminists to adopt radical FPA feminism. Radical FPA feminism is motivated towards fighting the oppression of both trans people and women. We have ethical reasons for accepting gender self-identifications, as explored in §2.1, which radical FPA feminism allows us to do. Additionally, radical FPA feminism enables us to meet central radical feminist concerns, and pointedly take on sex-based oppression. With radical FPA feminism we are able to avoid the conflation of sex-based oppression with oppression based on practice or category, and also able to understand oppressions stemming from specific gender combinations (e.g. the oppression faced by masc, AMAB, XY women or femme, AFAB, XX men). Furthermore, radical FPA feminism allows us to recognise that sex-based oppression historically precedes, is somehow more fundamental than, and/or causes the other gender-based oppressions, without having to define gender-categories via sex.

Furthermore, radical FPA feminism creates an obstacle for sex-based discrimination by removing the traditional properties of gender-categories. If cis women are subjugated on their assumption of their biology, making it so that one cannot tell one’s biology on the basis of one’s dress, name, pronouns, or stated

¹⁷ As an anonymous reviewer has suggested, this may have major implications on feminist theory and practice. For example, it would be pertinent for feminists to redress the oppression of biological females of other species (e.g. dairy cows), regardless of ethical attitudes towards non-human animals more generally.

gender, as per Bettcher, would plausibly make such discrimination more difficult (Bettcher, 2009). Increasing the numbers of women biologically unable to bear children (and the numbers of non-women able to bear children) strikes a blow to the organising logic of the patriarchal system. For example, in a radical FPA feminist society, if one wished to discriminate against those who may become pregnant, one would not be able to do so easily on the basis of dress, pronouns, stated gender, etc.

We can see radical FPA feminism as a potential step towards completely abolishing gender - a goal that many non-binary people, binary trans people, and radical feminists share in common. By ridding gender-categories of all-but-historical associations, they are able to potentially become obsolete. Therefore, radical FPA feminism can be considered part of the central radical feminist program of completely breaking the gender system. Radical FPA feminism is able to have a sex-based analysis of female oppression whilst simultaneously refusing to use the sex-based categorisations in day-to-day life. Additionally, radical FPA feminism is an affront to the compulsory heterosexuality that underpins the patriarchal order. The patriarchal promotion of heterosexuality is dependent on “men” and “women” being defined along reproductive lines.

Radical FPA feminism has further strengths. We can conceptualise the oppressive norm to have conformity between the three nodes and identities as *gender conformism*. Gender conformism is the combination of inter-node associations between category, sex, and practice, *and* the inter-identity associations between category-identity, sex-identity and practice-identity. Gender conformism leads to many problems within mainstream society, as well as leading both radical and associative FPA feminism to misgender people. Gender conformism leads to the marginalisation and invalidation of non-binary people, trans people, butch women and effeminate men alike. Gender conformism is toxic for both cisgender and trans people – enforcing masculinity on men and femininity on women. Trans and non-trans feminists alike should unite to tackle it.

Additionally, instead of only being able to discuss “male privilege”, “binary privilege”, and “cis privilege”, radical FPA feminists are able to discuss “AMAB privilege”, “cis-sex privilege”, “cis-practice privilege”, “cis-category privilege”,

“male biology privilege”, “manhood privilege”, “masculine privilege”, and “conformist privilege”. These are connected but distinguishable, and deserve individuated feminist scrutiny. Radical FPA feminism is uniquely placed to reject problematic gender conformist norms and analyse each of the above privileges, while simultaneously appreciating the historical connection between sexed-biology, gender-practice and gender-category.

7 Conclusion

If we are to take seriously both a radical feminist analysis of female oppression and a norm of gender self-identification, the most plausible route is radical FPA feminism. I have explored a potential solution to this longstanding rift in feminist thought with the hope that trans and non-binary identities need not be untenable to those with radical feminist aims and methods.

There are, however, many alternative avenues to be explored. Indeed, it may seem that my solution has failed both radical and the FPA feminists. Radical feminists may take issue with gender-categories no longer being defined by their most politically salient features, i.e. practice and sex. FPA feminists may take issue with self-identification being unauthoritative for the (more politically salient) sex-based and practice-based kinds. Exploring these, and other, limitations of radical FPA feminism is beyond the scope of this paper. However, if radical FPA feminism is found to be overly problematic, we may have grounds for rejecting the radical feminist analysis and/or gender self-identification. If so, radical FPA feminism is akin to (for want of a less patriarchal allegory) King Solomon chopping the baby in half: if the result is untenably repugnant, then maybe one of those fighting for the proverbial baby should relent.

References

Ansara, G., Webeck, S., Carpenter, M., Hyndal, P., & Goldner, S. (2015). *Re Commonwealth Attorney-General's Department Review of the Australian Government Guidelines on the Recognition of Sex and Gender*. Retrieved 12 27, 2017, from <https://lgbtihealth.org.au/wp-content/uploads/2015/10/FOR->

DISTRIBUTION-AGD-Sex-and-Gender-Guidelines-Review-Advisory-Group-Endorsement-Letter.pdf

Bettcher, T. M. (2007). Evil Deceivers and Make-Believers: On Transphobic Violence and the Politics of Illusion. *Hypatia*, 22(3), 43-65.

Bettcher, T. M. (2009). Trans Identities and First-Person Authority. In L. J. Shrage (Ed.), "You've Changed" (pp. 98-120). Oxford: Oxford University Press.

Bettcher, T. M. (2012). Trans Women and the Meaning of "Woman". In N. Power, R. Halwani, & A. Soble (Eds.), *The Philosophy of Sex (Sixth Edition)* (pp. 233-250). Plymouth: Rowman & Littlefield.

Bettcher, T. M. (2014). Feminist Perspectives on Trans Issues. Retrieved 12, 2017, from <https://plato.stanford.edu/archives/spr2014/entries/feminism-trans/>

Bettcher, T. M. (2014). Trapped in the Wrong Theory: Rethinking Trans Oppression and Resistance. *Signs: Journal of Women in Culture and Society*, 39(2), 383-486.

Bundesverfassungsgericht. (2017). Civil status law must allow a third gender option. Retrieved 12 27, 2017, from <https://www.bundesverfassungsgericht.de/SharedDocs/Pressemitteilungen/EN/2017/bvg17-095.html>

Busby, M. (2017). Canada introduces gender-neutral 'X' option on passports. Retrieved 12 27, 2017, from <https://www.theguardian.com/world/2017/aug/31/canada-introduces-gender-neutral-x-option-on-passports>

Butler, J. (1990). *Gender Trouble: Feminism and the Subversion of Identity*. Abingdon: Routledge.

Callahan, K. (2014). Stop Confusing Gender Dysphoria With Body Dysmorphia Already. Retrieved 03 16, 2017, from <http://roygbiv.jezebel.com/stop-confusing-gender-dysphoria-with-body-dysmorphia-al-1583049920>

Cohen, R. (2017). Trans Women Are Women. Retrieved 03 07, 2017, from <https://www.stonewall.org.uk/our-work/blog/trans-women-are-women>

Cox, S. (2016). Coming out as 'non-binary' throws other women under the bus. Retrieved 1 19, 2017, from <http://www.feministcurrent.com/2016/08/10/coming-non-binary-throws-women-bus/>

crashchaoscats. (2013). Socialized "Trans". Retrieved 03 14, 2017, from <https://crashchaoscats.wordpress.com/2013/08/09/socialized-trans/>

Diamond, L. M., Pardo, S. T., & Butterworth, M. R. (2011). Transgender Experience and Identity. In S. J. Schwartz (Ed.), *Handbook of Identity Theory and Research* (pp. 629-647). New York: Springer.

Fine, C. (2011). *Delusions of Gender*. London: Icon Books.

Ford, T. (2015). Navigating Fashion When You're Agender. Retrieved 03 16, 2017, from <http://www.racked.com/2015/7/1/8838119/agender-shopping>

George, B. R. (2016, 07 02). What even is 'gender'? (Draft). Retrieved 01 25, 2017, from <https://philpapers.org/archive/GEOWEI.pdf>

George, B. R., & Briggs, R. A. (2016, 10 27). Science Fiction Double Feature: Trans Liberation on Twin Earth (Draft). Retrieved 02 16, 2017, from PhilPapers: <https://philpapers.org/archive/GEOSFD.pdf>

Haslanger, S. (2012). Gender and Race: (What) Are They? (What) Do We Want Them to Be? In *Resisting Reality: Social Construction and Social Critique* (pp. 221-247). Oxford: Oxford University Press.

House of Commons Women and Equalities Committee. (2016). *Transgender Equality: First Report of Session 2015–16*. Retrieved 12 27, 2017, from <https://publications.parliament.uk/pa/cm201516/cmselect/cmwomeq/390/390.pdf>

Jeffreys, S. (2014). *Gender Hurts: A Feminist Analysis of the Politics of Transgenderism*. New York: Routledge.

Jenkins, K. (2016). Amelioration and Inclusion: Gender Identity and the Concept of Woman*. *Ethics*, 126, 394-421.

Jones, L. (2016). Women shoulder the responsibility of 'unpaid work'. Retrieved 01 30, 2017, from <http://visual.ons.gov.uk/the-value-of-your-unpaid-work/>

Koyama, E. (2003, 07 26). The Transfeminist Manifesto. In R. Dicker, & A. Piepmeyer (Eds.), *Catching A Wave* (pp. 244-259). Boston: Northeastern University Press. Retrieved 02 06, 2017

McKinnon, R. (2014). Stereotype Threat and Attributional Ambiguity for Trans Women. *Hypatia*, 29, 857–872.

McKittrick, J. (2015). A dispositional account of gender. *Philosophical Studies*, 172, 2575-2589.

Monroe, J. (2015). Being non-binary: I'm not A Girl Called Jack any more, but I'm not a boy either. Retrieved 02 21, 2017, from <http://www.newstatesman.com/politics/feminism/2015/10/being-non-binary-i-m-not-girl-called-jack-any-more-i-m-not-boy-either>

Overall, C. (2012). Trans persons, cisgender persons, and gender identities. In N. Powe, R. Halwani, & A. Soble (Eds.), *Philosophy of sex: Contemporary readings* (sixth edition) (pp. 251–267). New York: Rowan & Littlefield.

Ozturk, B. (Forthcoming). The Negotiative Theory of Gender Identity and the Limits of First-Person Authority. In A. Soble, R. Halwani, & N. Power (Eds.), *The Philosophy of Sex*, seventh edition. Rowman & Littlefield.

Raymond, J. (1994). *The Transsexual Empire: The Making of the She-Male*. New York: Beacon Press.

Reading, W. (2014). Separating Out Gender Identity from Gender Expression. Retrieved 03 16, 2017, from <http://everydayfeminism.com/2014/05/separating-identity-expression/>

Reed, N. (2013). *Born This Way (Reprise): The New Essentialism*. Retrieved 02 27, 2017, from <http://freethoughtblogs.com/nataliereed/2013/03/09/born-this-way-reprise-the-new-essentialism/>

Reilly-Cooper, R. (2015). *Sex and Gender: A Beginner's Guide*. Retrieved 12, 2017, from <https://sexandgenderintro.com/>

Reilly-Cooper, R. (2016). Gender is not a spectrum: The idea that gender is a spectrum is a new gender prison. Retrieved 12, 2017, from <https://aeon.co/essays/the-idea-that-gender-is-a-spectrum-is-a-new-gender-prison>

Stone, S. (1991). *The Empire strikes back: A posttranssexual manifesto*. In J. Epstein, & K. Straub (Eds.), *Body guards: The cultural politics of gender ambiguity* (pp. 280–304). New York: Routledge.

Titman, N. (2014). How many people in the United Kingdom are nonbinary? Retrieved 02 03, 2017, from <http://practicalandrogyny.com/2014/12/16/how-many-people-in-the-uk-are-nonbinary/>

Vincent, B. W. (2016, October). *Non-Binary Gender Identity Negotiations: Interactions with Queer Communities and Medical Practice*. PhD Thesis. Leeds: University of Leeds, School of Sociology and Social Policy.

Williams, C. (2015). The TransAdvocate interviews Catharine A. MacKinnon. Retrieved 03 16, 2017, from http://transadvocate.com/sex-gender-and-sexuality-the-transadvocate-interviews-catharine-a-mackinnon_n_15037.htm

Wiseman, M., & Davidson, S. (2011). Problems with binary gender discourse. *Clinical Child Psychology and Psychiatry*, 17, 528-537.

Yamaguchi, A. (2016). What's the Harm in "Female-Bodied"? Retrieved 03 13, 2017, from <https://the-orbit.net/splainyouathing/2016/01/15/whats-the-harm-in-female-bodied/>

Young, I. M. (2009). Five Faces of Oppression. In G. Henderson, & M. Waterstone (Eds.), *Geographic Thought* (pp. 55-71). London: Routledge.



ARE WE POST-JUSTIFICATION? STOUT'S CASE FOR SELF-KNOWLEDGE, POLITICAL JUSTIFICATION AND PUBLIC PHILOSOPHY

ESTAREMOS NA ÉPOCA DA PÓS-JUSTIFICAÇÃO? O ARGUMENTO DE STOUT A FAVOR DO AUTO-CONHECIMENTO, DA JUSTIFICAÇÃO POLÍTICA E DA FILOSOFIA PÚBLICA

Deven Burks*
deven.burks@uni.lu

Abstract. *Must the participant to public discourse have knowledge of her beliefs, attitudes and reasons as well as belief-formation processes to have justified political belief? In this paper, we test this question with reference to Jeffrey Stout's (2004) approach to public discourse and public philosophy. After defining self-knowledge and justification along the lines of James Pryor (2004), we map thereon Stout's view of public discourse and public philosophy as democratic piety, earnest storytelling and Brandomian expressive rationality. We then lay out Brian Leiter's (2016) naturalistic critique of public philosophy as "discursive hygiene" to see whether Stoutian public philosophy survives the former's emotivist-tribalist gauntlet. Lastly, we find that Leiter's critique proves less radical than it may appear and requires the moderating influence of a public philosophy like Stout's. All in all, Stoutian public discourse and public philosophy powerfully illustrates a strong, necessary connection between self-knowledge and political justification. Post-truth is not post-justification.*

Keywords: *self-knowledge; political justification; public philosophy; public discourse; Jeffrey Stout.*

Sumário. *Será que quem participa no discurso público tem de ter conhecimento das suas crenças, atitudes e razões, bem como dos processos de formação de crenças, de forma a ter crenças políticas justificadas? Neste artigo, testamos esta questão tendo como referência a abordagem de Jeffrey Stout (2004) ao discurso público e à filosofia pública. Depois de definirmos auto-conhecimento e justificação seguindo James Pryor (2004), aplicamos a perspectiva de Stout do discurso público e da filosofia pública como piedade democrática, narração honesta e racionalidade expressiva brandomiana. Em seguida, apresentamos a crítica naturalista de Brian Leiter (2016) à filosofia pública como "higiene discursiva" para determinar se a filosofia pública tal como a concebe Stout sobrevive ao desafio emotivista-tribalista daquele. Finalmente, concluímos que a crítica de Leiter é menos radical do que poderia parecer e requer a influência moderadora de uma*

* University of Luxembourg.

filosofia pública como a de Stout. Afinal, o discurso público e a filosofia pública que Stout defende ilustram de forma poderosa uma conexão forte e necessária entre auto-conhecimento e justificação política. A pós-verdade não é pós-justificação.

Palavras-chave: auto-conhecimento, justificação política, filosofia pública, discurso público, Jeffrey Stout.

0. Introduction

By many accounts, we are inhabitants of a post-truth world. Though truth and justification are separate notions, the question arises whether post-truth does not also entail a post-justification era. Indeed, in a world in which the person may say whatever she pleases, be it untrue or unjustified, it is unclear what role might remain for public philosophy. While this question surpasses the scope of a paper, we hope to make a contribution to the literature on political justification and public philosophy by exploring whether some form of self-knowledge is necessary thereto. To that end, we shall proceed in four parts. We shall begin by defining the notions of self-knowledge and justification pertinent to our enquiry and by setting out four versions of the link between self-knowledge and justification. In the second part, we shall lay out Jeffrey Stout's (2004) vision of public philosophy and the way in which it incorporates the relevant notions of self-knowledge and justification in the guise of "democratic piety", earnest storytelling and Brandomian expressive rationality. Thereafter, our attention will shift to Brian Leiter's (2016) critique of public philosophy from the twofold perspective of emotivism and tribalism, a critique which *prima facie* undercuts Stoutian public philosophy. Finally, we shall examine whether Stout's version of public philosophy is capable of turning Leiter's critique. In the end, we find that Leiter's critique proves less radical than it may appear and requires the moderating influence of a well-framed public philosophy like Stout's. All in all, a post-truth world may still have room, even a need, for justification and public philosophy.

1. Defining self-knowledge and political justification

What relation might obtain between self-knowledge and political justification? Certainly, any answer thereto will depend on the precise understanding of self-knowledge and political justification at issue. The kind of self-knowledge in which we are interested, while introspective, is not that of the

person's senses or pain. Rather, we are concerned with the person's knowledge of her own beliefs, attitudes, reasons and history and the extent to which such knowledge may be more or less elaborate. Crudely, self-knowledge might come in two types. Whereas "thin" self-knowledge designates the person's having knowledge of (a set of) her own beliefs, attitudes, reasons and history, "thick" self-knowledge involves knowledge of the upstream processes (further beliefs, attitudes, reasons and history) at work in the formation of (a set of) her own beliefs, attitudes, reasons and history. By extension, the same distinction applies to the person's knowledge of others: "thin" other-knowledge is secured by the person's knowledge of (a set of) another's beliefs, attitudes, reasons and history while "thick" other-knowledge comes with the person's knowledge of the upstream processes (further beliefs, attitudes, reasons and history) at work in the formation of (a set of) another person's beliefs, attitudes, reasons and history.

In parallel, we propose an understanding of justification which largely follows Pryor (2004).¹ For the latter, justification admits of two basic distinctions: 1.) "what you have justification to believe, and what you're rationally committed to believe by beliefs you already have" (Pryor, 2004, p. 363); 2.) "having justification to believe something, and having a belief that is justified or well-founded" (Pryor, 2004, p. 365). By 1.), Pryor highlights the relation which obtains between two or more beliefs when the person thinks to have "decisive justification" to believe p (Pryor, 2004, p. 364). If decisive justification for belief p is congenial to a further belief q , then decisive justification for p rationally commits her to the further belief q . If, on the contrary, decisive justification for belief p rules out a further belief q , then decisive justification for p rationally commits to her reject the further belief q .

¹ We should make clear from the outset that Pryor (2004) elaborates a twofold notion of justification with regards to Moore's paradox. Consequently, his text and examples bear largely on perception and perceptual evidence. We are, however, concerned with *political* justification. One may reasonably wonder whether we are right to cross-apply his nominally perceptual account to the political. In response, we shall point out only that, on one hand, there is no a priori conceptual reason to rule out general features which concern all instances of justification, and, on the other, Pryor's "justifiable"/"justified" distinction maps onto political justification just as well as onto perceptual. See *infra*. for examples to this effect. Furthermore, we will take this opportunity to emphasize that our account of the relation between self-knowledge and political justification is *Pryorian* rather than Pryor's own and that any distortions which we introduce therein are naturally our own. It bears mentioning that Pryor does not explicitly address the issue of self-knowledge though it appears indirectly in text and references.

To these relations, “rational support” and “rational opposition”, Pryor adds a third relation of interest, “rational obstruction” (*idem.*), wherein decisive justification for the belief p undercuts a further belief q . More precisely, the decisive justification for the belief p undermines the decisive justification which might count in favour of the further belief q . The author cites the case of defective colour perception: decisive justification for the person’s belief that her perception is colour-deficient would undercut decisive justification for her belief that she perceives a given colour at any particular moment. A more political example might be as follows: decisive justification for the person’s belief that the mass media are partisan or biased would undermine decisive justification for her belief that she obtains undistorted information at any given time. In short, the tension is at the level of grounds rather than content, as was the case for “rational opposition”. Any of these three relations between beliefs may obtain within the person’s mental economy without the decisive justification actually attaining the status of justification. It is enough that the justification appears decisive enough to provoke rational commitment of one form or another. Wherefore the importance of Pryor’s distinction between justification and rational commitment.

The second distinction can be summarized as follows: that between a belief’s being justifiable in virtue of an existing justification in the world and a belief’s being justified in virtue of a justification’s application to that belief. To put the point differently, consider that justification for a “belief p ” amounts to “reasons r ” which one might have for holding a “belief p ”. If one does not base one’s “belief p ” on those “reasons r ” or opposes thereto other reasons or beliefs, then the “belief p ” is not justified, though potentially justifiable. Only when a.) one has justification for a “belief p ” and b.) one bases the “belief p ” on those “reasons r ” and c.) one has no other reasons or beliefs which rationally commit one to an opposed belief (*i.e.* rational opposition or obstruction) is the “belief p ” then justified rather than merely justifiable. More simply, having possible justification for a “belief p ” comes apart from being justified in having the “belief p ”. When the latter obtains, we will refer to the person’s being “P-justified”.

Extending these distinctions to political justification consists in transposing the above structures onto recognizably political beliefs. To take but two examples, consider the following political beliefs: a progressive income tax is just; current

electoral systems are unjust. Following Pryor, one might have justification for the belief that a progressive income tax is just while failing to have the belief itself due to ignorance, doubt, weakness of will, or conflicting beliefs. Perhaps the person has access to scholarly findings or political principles which justify that belief but simultaneously believes that scholarly findings are politically compromised, in which case the belief is justifiable rather than justified. Likewise, a belief that current electoral systems are unjust will only be justified when one has justification therefor, bases belief thereon and has no conflicting rational commitments. To reprise the first example's structure, the person has access to scholarly findings or political principles showing that the electoral system is unjust, bases her belief thereon and has no conflicting rational commitments as to the electoral system's justice.

So far, so good. It may, however, seem that an emendation is in order. For Pryor's breakdown of justification and rational commitment, justifiable belief and justified belief remains within a passive framework, even upon transposition. In contrast, political justification, when considered in the form of public discourse and justifying beliefs to others, moves us to an interpersonal and active setting. In this case, successful political justification might take either of the following forms: transmitting a justified belief from one person to another; adapting a belief or justification thereof to another's "cognitive context" to elicit the latter's recognition.² Both cases *prima facie* preserve thin and thick self-knowledge requirements: one's expression of a justified belief must retain the relation between beliefs and reasons which render it justified on the interpersonal level.

Put differently, whether the person seeks another's adoption or recognition of a given belief, the belief to be adopted or recognized must stand in the same relation to justifying reasons if it itself is to be justified rather than merely justifiable. Moreover, another's adopting or recognizing a justified belief may require the person to know something of the other's cognitive context if she is to succeed in having the belief adopted or recognized. In this instance, it will be important for the person to have thin and thick *other*-knowledge: knowledge of

² By this term, we intend a person's relation to a personal history in the broader historical setting of a given time, place and culture. Though not his own term, its formulation is inspired by Stout (1988).

the other's beliefs, attitudes, reasons and history; knowledge of the upstream beliefs, attitudes, reasons and history at work in the development of the former.

Nonetheless, it should be asked whether one can have political justification without self-knowledge. The properly amended Pryorian account seems to require self-knowledge for political justification insofar as one must be aware of one's "belief *p*" and of any other reasons or beliefs which commit oneself to opposed positions. Otherwise, one might simply have a justifiable belief without that belief itself attaining "justified" status. Nevertheless, there are cases in which justification seems to proceed without self-knowledge, thin or thick. One prominent counter-example might be the Rawlsian original position for which, to arrive at a conception of justice for the distribution of primary goods in society or the evaluation thereof, the person has no knowledge of her preferences or beliefs ahead of the resultant distribution (Rawls, 1971). In short, the person would have access to neither the decisive beliefs, attitudes, reasons or history nor those upstream thereof. Consequently, the lack of self-knowledge would in no way hinder political justification and might even facilitate consensus.³

That said, one may reject this depiction as misleading for several reasons. On our Pryorian account of political justification, the class of "beliefs *p*" and "reasons *r*" may be suitably narrowed such that the little which one knows about oneself figures all the same in the justified belief. After all, a person in the original position knows at minimum that she is: a moral person with a sense of the good as well as a sense of justice; a party to society as a system of cooperation; a being in need of certain primary goods. If such pieces of knowledge, however "thin" these may be, serve as justifying reasons to the person's "belief *p*" and are consciously adopted as such, then thin and thick self-knowledge may nevertheless obtain. Conversely, if such knowledge is not applied to justify the person's "belief *p*", then that belief may be justifiable rather than justified. In short, though the original position limits the kinds of self-knowledge available to the person therein, she may still stand in the right formal relations to her beliefs,

³ This consensus-facilitating quality has received some support in the empirical literature. Cf. Frohlich, Oppenheimer, and Eavey (1987).

attitudes, reasons and history, however thinly conceived. Nor does their thinness make the necessity of self-knowledge for justification merely trivial.

Provisionally, this framework leaves us with four possible answers to our starting question, namely, whether self-knowledge is necessary for political justification:

Strong and thick version: Thick self-knowledge is necessary for political justification.

Strong and thin version: Thin self-knowledge is necessary for political justification.

Weak and thick version: Thick self-knowledge is important for political justification.

Weak and thin version: Thin self-knowledge is important for political justification.

2. Stout on self-knowledge and justification: piety, storytelling and expressive rationality

In contrast to Pryor, Jeffrey Stout has a more elaborated, if not systematic, view of the relation between self-knowledge and justification, particularly its political subspecies. In this section, we aim to determine the extent to which Stout might draw on a Pryorian framework in his treatment of self-knowledge and justification. Stout's (1988; 2004) view holds that justification, broadly conceived, proceeds or should proceed with a view to the person's "cognitive context", *i.e.* her relation to a personal history within the broader historical setting of a given time, place and culture. Insofar as justification is indexed to a given context and successful justification turns on knowledge of that context and a given context includes the person justifying and her horizon of reasons, then justification's success is indexed to knowledge of the person justifying and her horizon of reasons. When justifying one's views to another, justification then entails thick self-knowledge (*strong and thick version*). Likewise, it calls for thick-other knowledge in situations where one aims to have another adopt or

recognize a “belief *p*”. Hence, Stout’s account preserves the broad strokes of a Pryorian framework.⁴

Over four books, Stout attempts to work out the insight that justification is indexed to a cognitive context. Political justification receives its most theoretically complete treatment in Stout (2004) wherein the author confronts the reader with a view of political deliberation as earnest reason-giving. In short, when engaged in political deliberation, participants thereto should give their real reasons for a given political position, whether those reasons be tied to a political conception of justice or to a comprehensive doctrine, religious, philosophical or moral (Rawls, 1993). As Stout puts it, “cultivate the virtues of democratic speech, love justice, and say what you please” (Stout, 2004, p. 85). On one hand, we should not require persons to voice reasons other than their own.⁵ On the other, he deems it epistemologically unreasonable on our part to maintain that persons giving comprehensive beliefs and reasons for a political position be excluded from deliberation as such beliefs and reasons can be responsibly held or P-justified.⁶ Recall that a “belief *p*” is P-justified only when a.) one has justification for a “belief *p*” and b.) one bases the “belief *p*” on those “reasons *r*” and c.) one has no other reasons or beliefs which rationally commit one to an opposed belief (*i.e.* rational opposition or obstruction).

Naturally, for the person to voice those beliefs and reasons and to be entitled thereto, she must first know what they are. This requires the person’s taking stock

⁴ For a more complete overview of justification, see notably Stout (1988), pp. 28-30 and Stout (2004), pp. 231-237.

⁵ This is a requirement which Lafont (2013) has dubbed, with reference to Habermas’ work on religion, as “*the right of all democratic citizens to take their own cognitive stance in public deliberation*” (p. 243, emphasis in original).

⁶ While Stout prefers the language of responsibly holding a belief or “epistemic entitlement” (Stout, 2004, pp. 65-73), these notions demonstrate considerable similarity with being P-justified. A person who is epistemically entitled to a belief has potentially justifying reasons therefor, is aware of these reasons as being such and conforms to her broader epistemic responsibilities (Stout, 2004, p. 71). In this, a Pryorian account and Stout’s approach are similar, and we will use “epistemically entitled” and “P-justified” interchangeably. Nevertheless, the authors are not without their differences. Most notably, they would charge each other with confusion over whether it is a “belief *p*” which is justified or whether it is the person, who in holding “belief *p*”, is justified. Stout is keen to differentiate the way in which a person “is justified in believing a claim if he or she is entitled to be committed to it, given his or her discursive context and cognitive conduct” from the way in which a belief or claim is “justified in some discursive context if *everyone* in that context is justified in believing it” (Stout, 2004, p. 99). Pryor would likely see therein two situations in which the same conditions apply to justification, the latter’s context-wide justification being merely an aggregation of the former’s person-by-person instances. In the end, whether one adopts Stout’s distinction or rebuffs it as confusion does not alter the underlying structure of entitlement.

thereof, for which the person may avail herself of “democratic piety” (Stout, 2004, p. 9). Stout understands this notion as the person’s making an inventory of the different moral and social sources responsible for the shape of her life at a given moment in time and giving appropriate expression thereto. In that these intertwine with the person's history, they provide her with a horizon of reasons which may justify or fail to justify her political position (the state of “being justified”) and which she may use to justify, successfully or unsuccessfully, her political position (the act of “justifying”). In a word, the person’s contribution to public discourse should manifest in some form the different social resources at work in her ethical and political reasoning. Notably, that form may vary depending on the context. Stout refers most often to a “democratic” context, an “Emersonian” sub-context:

[F]rom a democratic point of view, the only piety worth praising as a virtue is that which concerns itself with *just* or *fitting* acknowledgment of the sources of our existence and progress through life (...) Imagining or conceiving of those sources and choosing ethically and aesthetically apt expressive means of acknowledging dependence on them are both things for which an Emersonian poet or essayist expects to be held responsible discursively (Stout, 2004, p. 30).

Indeed, these contexts are themselves capable of greater precision depending on the moral and social sources at issue. On several occasions, Stout point the way towards democratic piety by describing the moral and social sources on which he himself depends for his moral and political reasoning. To take but one example:

In the days of my adolescent sublime, Martin Luther King, Jr., was the hero of my humanitarian cause, and Jesus was one of three personifications of my loving divinity. Nowadays things have become more complicated, because I have come to know more about these figures of virtue than their hagiographers and publicists wanted me to know (...) Love and justice remain virtues (...) but now the relation between the persons and the virtues is more complicated. It requires a different, less doctrinal, more improvisational kind of explication. To the extent that King and Jesus exemplify virtues in my imaginative life, they now do so imperfectly and defeasibly. I therefore need an open-ended way to think the relation through, as it were, from both sides at once. Neither doctrine, nor

principle, nor system, nor overarching plot, knowable in advance, constrains the course of thinking (...) We all have our examples, after all, and we all make something of them sooner or *[sic]* later. We do not, however, make the same thing of them. Neither do they make the same thing of us (Stout, 2004, p. 173).⁷

Herein, we see democratic piety at work. Stout tells his story and takes inventory in “critical” fashion: recognizing moral and social sources as such; refusing one-sided tendencies to nostalgia, wishful thinking or idealization; introducing between person and sources two-sidedness via reflection, research and questioning. Just such an earnest stocktaking leaves the person ready to advance reasons, be it in the form of structured argument or personal storytelling.⁸

Indeed, one reasonably wonders how the person may in public discourse articulate for the audience her sources, reasons and commitments and, hence, her cognitive context, of which she will have taken stock via democratic piety. Stout's (2004) clearest answer, worked out in greater detail in Stout (2010a), lies in the idea that persons should voice their deepest reasons and personal histories. In criticizing both universalist and particularist thinkers, he remarks “that neither [Seyla Benhabib nor Stanley Hauerwas] has imagined the possibility, let alone the desirability, of a loosely structured democratic conversation in which variously situated selves tell their own stories on their own terms” (Stout, 2004, p. 179).

In other words, apart from structured argument, a key way for epistemological formations such as beliefs, attitudes, reasons and history to come to light consists in the person's telling her own story and development. Indeed, storytelling enables the audience better to grasp the real horizon of reasons and commitments within which the person is working (supposing that the person knows the story elements and tells it accurately and non-coercively). Moreover, not only should the person's right to self-expression and storytelling merit respect from the audience; her beliefs and reasons, as well as the person herself, deserve our respect, on the condition of their being P-justified. Certainly, whether we give

⁷ For another example, see Stout (2004), p. 97.

⁸ For a review of the deliberative democratic literature on rhetoric and storytelling as modes of deliberation, see Dryzek (2000), pp. 50-56, 62-70. For empirical analysis of one instance of discourse and modes of deliberation therein, see Bächtiger and Gerber (2014).

such respect is another matter. Stout drives this point home when enjoining the audience to respect both the believer's right to express and the content expressed: "Insofar as they [those who differ from us religiously] do acknowledge that dependence [piety] appropriately, given their own conceptions of the sources of existence and progress through life, they may be said to exhibit an attitude that is worthy of our respect, if not our full endorsement" (Stout, 2004, p. 34).

Provided that the person abides by the virtue of democratic piety and appropriately takes stock of and reflects her context or beliefs, attitudes, reasons and history in discourse, the person merits our respect as a fellow interlocutor whose beliefs, attitudes, reasons and history are worth examining at greater length. All the more so in that such expression sets the audience up to unearth the beliefs, attitudes, reasons and history motivating the person to support a given position; it may also grant the person an opportunity at greater thin or thick self-knowledge through give-and-take with the audience. More importantly, listening to stories and according respect as due does not inevitably lead to a mere *modus vivendi*. By censoring them, not only do we "remain ignorant of the real reason that many of our fellow citizens have for reaching some of the ethical and political conclusions they do", but we "also deprive them of the central democratic good of expressing themselves to the rest of us on matters about which they care deeply" (Stout, 2004, p. 64). In so doing, we "lose the chance to learn from, and to critically examine, what they say", and they "have good reason to doubt that they are being shown the respect that all of us owe to our fellow citizens as the individuals they are" (*idem.*).

Far from isolating the audience into incommensurable spheres of discourse, Stout's take on respect lays the first stepping stone to critical examination and exchange of beliefs, attitudes, reasons and history. For without the knowledge afforded by listening and respect, critical examination would otherwise have no object on which to work. And, without materials for examination, no way out of an impasse will present itself to participants in public discourse, perhaps even more so than in cases of deep disagreement.

In addition, respect serves a second, more practical purpose. If, by listening to stories, respect may grant the listener knowledge of the person's real beliefs, attitudes, reasons and history informing a given position (thereby promoting

other-knowledge), this practical exercise also makes the person more amenable to the subsequent exchange of reasons with the audience. In one important sense, through respect, the audience may recognize the person as a full-fledged member of public discourse and, given the necessary conditions, as P-justified in believing that to which she has just given expression. This comes out even more strongly in Stout (2010a), a case-study of broad-based citizens' organizing centring on "house meetings" and "one-on-ones", *i.e.* "individual conversations" and "small gatherings".⁹

Yet a person herself may not always be able to take stock of or to draw out what follows from the critical inventory made during the phase of piety and aired out in the phase of storytelling. Accordingly, the person may require a spokesperson, be this a philosopher to work out the inferential commitments underlying the reasons for her political position or a community organizer to extract the issue taking shape therein. In both instances, another will bring greater clarity to what links certain commitments or issues. Wherefore Stout's view that public philosophy consists in making explicit and scrutinizing the commitments and norms implicit in public discourse and reasons: in short, it requires an exercise in Brandomian "expressive rationality" (Stout, 2004, pp. 12-14; Brandom 1994).

More concretely, this may entail taking norms or reasons, often expressed as material inferences "given x , I shall y " for which we ordinarily acknowledge x as a legitimate premise for the conclusion y , and working out the premise needed to make them formally valid (Stout, 2004, pp. 188-190). So, to statements like the following: "(a) Going to the store is my only way to get milk for my cereal, so I shall go to the store; (b) I am a lifeguard on the job, so I shall keep close watch over the swimmers under my protection; (c) Ridiculing a child for his limp would humiliate him needlessly, so I shall refrain from doing so"; we would need to append further premises: "(a) a statement expressing my *desire* to have milk for my cereal; to (b) the conditional that if I am a lifeguard, it is my *responsibility* to

⁹ For an overview of these organizations and interactions, see Stout (2010a), pp. 2-3; for a detailed examination of one house meeting or small gathering, see Stout (2010a), pp. 151-156.

keep a close watch over the swimmers under my protection; or to (c) the principle that one *ought* not to humiliate people needlessly” (Stout, 2004, p. 188).¹⁰

This exercise presents the “advantage of putting the formerly implicit material inferential commitment in the explicit form of a claim, which in turn allows it to be challenged or justified inferentially in light of other considerations” and takes on still greater importance “when conflicts arise among different material inferential commitments that we have undertaken” (Stout, 2004, p. 189). Hence, Stout’s account of expressive rationality incorporates a requirement like that of examining rational commitments on Pryor’s view in order to arrive at epistemic entitlement or P-justified belief.

More important for our purposes is the way in which working out inferential commitments proves both means and hindrance to the link between self-knowledge and political justification outlined above. For, if our capacity for self-knowledge may help to secure the justified quality of our political positions by working out the entailments of our beliefs, our cognitive failings may also hinder arriving at (thick) self-knowledge and, hence, P-justified political belief. Such that the person may need to rely upon the public philosopher to arrive indirectly at the thick self-knowledge necessary for a P-justified political belief. Indeed, the person may lack entirely the expressive resources necessary to render those commitments explicit or even the cognitive resources necessary to work the latter out (Stout, 2004, p. 193).

Before concluding this section, it will be helpful to consider, with reference to our examples above, how Stoutian self-knowledge and political justification intersect in democratic piety, personal storytelling and expressive rationality and create a need for a public philosophy. In the first case, a person may have the belief that a progressive income tax is just but initially prove unable to say why she holds that belief. To that end, she would, on Stout’s approach, take stock of the moral and social sources informing her belief, be those moral beliefs about fairness or religious beliefs about charity. From there, she would find the mode of expression most fitting for sharing her belief and framing it with reasons. That mode of expression may include personal storytelling, *e.g.* her own experiences

¹⁰ For Brandom’s exposition of this point, see Brandom (1994), 243-253 and Brandom (2000), ch. 2.

as a member of a society with a progressive income tax. Finally, both prior to and during public discourse, she and others would work out whether her beliefs and reasons commit her to other beliefs and reasons, and, if so, which. In this way, she might arrive at a P-justified political belief on the justice of a progressive income tax.

If the above stands as a successful, direct instance of Stoutian political justification, we may consider a second instance wherein political justification succeeds, though only indirectly. Specifically, a person may have the belief that the current electoral system is unjust but initially prove unable to say why she holds that belief. Accordingly, she would then take stock of the moral and social sources informing her belief, perhaps in the form of moral beliefs about autonomy or religious beliefs about temporal power. If she is capable of this inventory, she may nonetheless be unable to frame her belief in terms of deeper beliefs or reasons or to work out the various commitments to which her belief commits her. In such a case, she would fail to achieve thick self-knowledge and, on Stout's view, require a public philosopher's help to work out her upstream beliefs, reasons and commitments in order to attain a P-justified political belief on the injustice of the electoral system.

Finally, it remains to be seen how this impacts our main question on the relation between self-knowledge and political justification. In this section, we have laid out Stout's view of self-knowledge and justification in the threefold form of democratic piety, personal storytelling and expressive rationality. We have further shown the ways in which that threefold form resembles the Pryorian framework, necessitates thick self-knowledge for political justification and thereby subscribes to the strong and thick version of the relation between self-knowledge and political justification. Furthermore, we have seen how Stoutian justification creates a need for public philosophy. For, if public philosophy advances self-knowledge in public discourse, it also advances political justification. On the other hand, public philosophy's advancing self-knowledge hinges on thick other-knowledge and on the person's taking responsibility for how she forms beliefs before and after the public philosopher's work. Otherwise, the public philosopher's work on the raw material of piety and storytelling and on logical entailments of the beliefs, attitudes, reasons and history exposed therein

is for nought. We may likewise question whether public philosophy and its practitioners are themselves capable of articulating thick other-knowledge and thereby advancing indirect self-knowledge. Both hinge on public philosophy's efficacy. If public philosophy and public philosophers are incapable of occupying this role, then Stout's view of justification is mistaken or misleading.

3. Leiter on the paradoxes of public philosophy: Emotivism and tribalism

Brian Leiter would likely suspect just this about Stout's view. Leiter (2016) scrutinizes any public philosophy which sees itself as "contribut[ing] philosophical insight or knowledge or skill to questions of moral and political urgency in the community in which it is located" (p. 51). As this *prima facie* concerns Stout's appeal to public philosophy as an exercise in expressive rationality, Stout must then wrestle with the two paradoxes laid out by Leiter. The first holds that what expertise philosophers may "offer can not [*sic*] consist in any credible claim to know what is good, right, valuable, or any other substantive normative proposition that might be decisive in practical affairs (Leiter, 2016, p. 53).

The second paradox builds off the first. If philosophers do not bring to public discourse substantive normative knowledge, they may nevertheless offer "a *method or way of thinking* about contested normative questions that they offer" (*idem.*, emphasis in the original). In this method, public philosophers engage in an activity quite close to Stout's "expressive rationality". Namely, such philosophers set out from a set of normative commitments and "work out their entailments, demonstrating claims of the form, 'If you believe X, then you ought to believe Y,' and, 'If you believe Y, you should not do Z.'" (*idem.*). In so doing, they engage in "discursive hygiene" by which the author denotes "parsing arguments, clarifying the concepts at play in a debate, teasing out the dialectical entailments of suppositions and claims" (*idem.*). The paradox lies in that, while public philosophers may claim to bring discursive hygiene to public discourse, that hygiene "plays almost no role in public life, and an only erratic, and highly contingent, role in how people form beliefs about matters of moral and political urgency" (Leiter, 2016, p. 55).

If, as suggested above, we associate Stout's view of public philosophy as expressive rationality with discursive hygiene but find ourselves obliged to admit the lack of discursive hygiene in public discourse, we might nevertheless maintain that interlocutors are amenable to the work of "discursive hygiene" either at their own or another's behest. Leiter forestalls this possibility insofar as the psychological phenomena of emotivism and tribalism should temper our optimism about the state of public discourse and public philosophy as "discursive hygiene". These phenomena require further explanation.

Of the first phenomenon, emotivism, Leiter holds that it acts as a limiting case on public philosophy as "discursive hygiene". Notably, Leiter associates emotivism with Charles Stevenson's seminal position which Leiter presents as follows: "Ethical disagreements are at bottom a function of disagreement in attitudes, rather than disagreements about beliefs" (Leiter, 2016, p. 53). To this, he adds that "the connection between particular facts and our attitudes is just a contingent *psychological/causal fact*", to wit "it is just a psychological fact about many creatures like us that if our beliefs change, our attitudes often change too" (Leiter, 2016, p. 54, emphasis in original).¹¹ On this reading, public discourse may take the form of conflict between either attitudes or beliefs. If between beliefs, though rare, then the conflict may be brought to an end by ensuring convergence between beliefs. If between attitudes, then the conflict admits of no clear-cut solution in that attitudes do not seem reason-responsive in the same way as beliefs.

Certainly, we may allow that our beliefs influence our attitudes; we cannot, however, maintain with any certainty how or which beliefs influence attitudes. For we are simply unable to plot the causal mechanisms by which such changes are effected. Moreover, attitudes may alter in light of beliefs which we ordinarily deem "ethically irrelevant" because "self-serving" (Leiter, 2016, pp. 54-55). All in all, our inability to pin specific attitude changes to certain belief changes "includes

¹¹ It bears mentioning that the author also finds emotivism's main thesis in principle compatible with "discursive hygiene" (Leiter, 2016, p. 53). In short, both could in theory exercise causal power in belief-formation. In practice, emotivism tends, however, to exclude the causal efficacy of "discursive hygiene".

changes in belief about the logical or inferential relations between beliefs or between beliefs and attitudes” (Leiter, 2016, p. 55).

This should be worked out more concretely. Reprising our second example, a person may have the belief that the electoral system is unjust due to her moral beliefs about autonomy, religious beliefs about temporal power or her underlying social attitudes about democracy and politics. To change her belief that the electoral system is unjust, another person might try one of several strategies. The latter may appeal to ethical considerations counter to her moral or religious beliefs (*e.g.* the temporal-spiritual power divide is necessary), to unethical considerations about that which best advances her self-interest (*e.g.* the electoral system best preserves her own or group’s freedom of belief) or to opposed social attitudes (*e.g.* a renewed sense of public service). Supposing that, following the exchange, the person holds the belief that the electoral system is just, Leiter’s point is that neither she nor the other will be able, introspectively or conceptually, to ascribe the change in belief to a given strategy or the specific considerations advanced therein rather than non-discursive factors.

On Leiter’s account, there are thus no rules, inferential or otherwise, governing the transformation of and causal interaction between beliefs and attitudes. A given belief may provoke change in another belief or attitude without one’s being aware thereof. Likewise, one may attribute change in a given belief to some other belief or attitude but do so in error. Even if beliefs have causal traction over attitudes, this casts doubt on a.) whether thin knowledge of one’s beliefs can induce changes in one’s attitudes and b.) whether thick knowledge of how one came by those beliefs can bring on change in one’s attitudes. For the author, there is to be no conceptual apparatus capable of reliably tracking or predicting those changes.

Leiter highlights two examples to further his claim: the 2014 Steven Salaita controversy at the University of Illinois Urbana-Champaign suggests that discursive hygiene on clear-cut legal norms can fail; Jonathan Haidt's 2001 "incest experiment" tests a social intuitionist model of practical reasons, “according to which in most ordinary situations, moral judgments are produced by emotional or affective responses, the reasons adduced in their support being post-hoc: they do not explain the judgment, as evidenced by the resilience of the

judgment even in the face of the defeat of the proffered reason” (Leiter, 2016, pp. 57-58).¹² While such examples do not disprove the effect of discursive hygiene on moral and political judgments, they act as a limit to public philosophers’ optimism.

With this, Leiter’s account of emotivism is in place: a person’s moral and political beliefs, reasons and judgments are a function of emotional attitudes or affective responses, and changes therein are subject to no systematic exposition. Since the person is mistaken on the real reasons and processes behind her “belief *p*”, that belief cannot be P-justified. At this point, the author is ready to move to the second limiting case to “discursive hygiene” in public discourse. If “prejudice and bias are dominant forces in human life”, this owes to “tribalism”, whereby Leiter means “the propensity of creatures like us to identify with those ‘like themselves,’ and to view others as unacceptably different, deficient, depraved, and perhaps dangerous” (Leiter, 2016, p. 59).

Although the author sees in tribalism “the dominant force in public life” (*idem.*) and sketches how it may determine whom the person deems worthy of moral concern in virtue of their similarity or dissimilarity, more can be said of how tribalism directly impacts discursive hygiene and public discourse. First, the person tracks not the inferential relation between beliefs, attitudes and reasons highlighted by discursive hygiene but, instead, her own similarity with other interlocutors. Second, it predisposes the person to adopt divergent discursive stances towards those interlocutors similar and dissimilar to her. While she will be more likely to accept a similar person’s beliefs, attitudes, reasons and history as good and to deem the latter a rational actor or a locus of human dignity, she will be more likely to dismiss a dissimilar person’s beliefs, attitudes, reasons and history as bad and to view the other as less than a rational actor or a locus of human dignity.

If, as Leiter concedes, such tribalistic mindsets have somewhat given way with the rise of international and transnational bodies, the fact remains that such institutions as the United Nations and notions as universal human rights first

¹² That said, Brandom and Stout are also sensitive to rationalizing of this kind. See Brandom’s discussion of acting *with* or *for* reasons (Brandom, 2000, pp. 83-84, 95-96). See *infra.* for Stout’s breakdown of types of rationalization and opposition to same-sex marriage.

emerged following horrendous, widespread conflict. To this end, Leiter contends that "argument played little or no role" therein; rather, such progress is to be attributed to "emotional revulsion at barbarity" (*idem.*). Linking this to his emotivist point on attitudinal recalcitrance, he remarks that "[t]he key point, however, is that we philosophers must recognize that moral change depends fundamentally on the emotional attitudes of people, and that these attitudes tend in a strongly Tribalistic direction" (Leiter, 2016, p. 60). For the author, this comes out most strongly in whom one considers deserving of one's moral concern: those with whom one feels an emotional connection, such as kinship, merit one's moral regard. All the same, we find that Leiter's main point might be extended to the discursive level: one's beliefs, attitudes and reasons are most responsive to pressure coming from those who are similar to oneself, yet those similar to oneself are precisely those least likely to pressure one on one's beliefs, attitudes and reasons.

To return to our first example, consider a person who holds the belief that a progressive income tax is just, be it for moral reasons about fairness, religious reasons about charity or social attitudes about wealth. Leiter's main point is that she will judge the taxation system fair insofar as it benefits those similar to herself and penalizes those dissimilar to herself, for the former, rather than the latter are worthy of her moral concern. At the discursive level, were another to challenge the person on her belief that the taxation system is just or on her attitudes or reasons supporting that belief, her likelihood of engaging the former's challenge and related beliefs, attitudes and reasons would hinge on the former's similarity with herself. In such conditions, she may not seek knowledge of her own beliefs, attitudes, reasons and history necessary to arrive at P-justified belief.

For his part, Leiter contends that public philosophers are most likely to broaden interlocutors' moral regard for other persons or beings when they appeal to interlocutors' emotions rather than to the inferential connections between beliefs, attitudes and reasons. Case in point, Peter Singer, perhaps today's foremost public philosopher, appeals to emotional attitudes with no deeper rational basis (*i.e.* the moral salience of suffering rather than species) and makes his argument most effectively when relying on "moral perception", *e.g.* his grisly description of factory farming (*cf.* Leiter, 2016, pp. 60-62). Similarly, Leiter sees

the appeal to non-rational considerations such as “theoretical simplicity, methodological conservatism, and consilience” (Leiter, 2016, p. 60) as still more reason to temper our expectations for public philosophy. If public philosophers count on emotional attitudes and affective responses in lieu of discursive hygiene to do the philosophical heavy lifting, the author rightly wonders why we should expect better of lay audience members in public discourse.

Despite his vigorous contestation of discursive hygiene, Leiter does not mean to sound the death knell of all public philosophy. On the contrary, he marshals three considerations in its favour:

Being unable to contribute meaningfully to urgent ethical and political matters in no way diminishes the importance of finding an answer to those matters.

If we do not understand well the causal linkage between beliefs and attitudes at the time of discursive hygiene, this does not mean that discursive hygiene might not track the evolution of beliefs and attitudes. One should thus go on providing such hygiene.

Law, the discipline closest to philosophy, practices discursive hygiene and recognizes the need therefor in the sense that logical entailments can constrain attitudes when the time comes to rationalize the reasons proffered and the attitudes adopted.¹³

In this way, Leiter hopes to moderate the thoroughgoing critique of public philosophy which his treatment otherwise suggests. Indeed, it is worth emphasizing the most important lesson which the author draws from 3.). Namely, he contends that law has understood something which philosophy has not: rhetoric or “the art of persuasion apart from appeal to what follows from discursive hygiene” may play a predominant role in determining whether and what beliefs a person holds (Leiter, 2016, p. 63). Put differently, “‘belief fixation’, the process by which certain beliefs take hold in the cognitive and affective economy of the mind and thus yield action, does not necessarily track evidential, inferential and logical relations that interest philosophers” (Leiter, 2016, p. 64).

¹³ For lengthier exposition of these points, *cf.* Leiter (2016), pp. 62-63.

If this suggests that philosophers aiming to provide clarity on ethical and political issues should add rhetoric to their argumentative toolbox, they must not abandon the careful work of discursive hygiene, for “rhetoric does not tell us what *beliefs* we should try to produce with our rhetorical tools” (*idem.*, emphasis in the original). In a word, Leiter contends that philosophical argumentation must, at least at one level, track emotional attitudes and affective responses if it is to retain currency within public discourse. Before assessing whether this does enough to moderate Leiter’s critique and whether Stout can take this requirement on board, it is worthwhile to step back and to take stock of whether and to what extent Leiter’s account modifies the connection which we have sketched between self-knowledge and political justification.

His emotivist charge purports to show that the person first has an emotional or affective response to a given set of circumstances for which she only afterwards adduces reasons. This would cast doubt on her claim to thick self-knowledge in that she deceives herself on the processes behind that belief. As to her thin self-knowledge, if she is aware of her belief on a given political position, she nonetheless mistakes her immediate reasons therefor. Were she aware of the emotional source of her belief, she would recognize the adduced reasons as post-hoc additions. Accordingly, her thin self-knowledge is incomplete in part because of the opaque character of the thick, and her belief is, at best, justifiable, though not P-justified.

What then of Leiter’s tribalist charge? If emotional attitude or affective response determines political attitudes and emotive attitude or affective response is divided along tribal lines, then political attitudes are themselves divided along tribal lines. Such that the self-knowledge breakdown for the emotivist charge seems to cross-apply: while the person may know her belief on a position and the purported reasons for that belief, she also ignores how she came to hold the latter and what reasons genuinely support it. Again, thick and thin self-knowledge fail in that she ignores why she considers certain of moral concern rather than others, certain as offering valid beliefs, attitudes or reasons rather than others and her real reasons for judgments of worth and validity. Thus, her political position is at best justifiable in light of other considerations but not P-justified by way of her own beliefs.

Given this failing of thick self-knowledge, Leiter could contend that we must forego thick self-knowledge since the person is constitutively incapable of recognizing the emotivist and tribalist cast of her attitudes, beliefs and reasons. Instead, we should apply pressure to thin self-knowledge through rhetoric and “moral perception” targeting the reasons adduced for a given belief. In the Pryorian terms outlined above, Leiter would, on this reading, play down the possibility of political justification in that few, if any, will hold P-justified (as opposed to justifiable) beliefs, however attached we might be to attaining P-justified political belief. Does Stout have an answer to these challenges or is his view of public philosophy as expressive rationality condemned? For that matter, can Leiter both maintain such a thoroughgoing critique and hold out hope for public philosophy without some moderating influence?

4. What role for Stoutian public philosophy in the face of Leiter’s challenge?

Certainly, Leiter’s challenges to public philosophy as discursive hygiene are considerable, and no panacea is likely to be found therefor. Moreover, a full rebuttal thereof would swiftly surpass the scale of a paper. We will content ourselves then with sketching either of two tacks which Stout might use to bridge the gap between his public philosophy and Leiter’s naturalistic critique.

The first consists, first, in reconciling post-hoc rationalization to self-knowledge and Stoutian public philosophy and, then, in managing one’s discursive expectations accordingly. Stout (2010b) takes a step in this direction by elaborating a typology of opposition to same-sex marriage. More precisely, he outlines how the beliefs of religious opponents to same-sex marriage may be analysed with an eye to the role that reasons play in rationalizing their opposition. For those “sadistic homophobes” who use religious rationalization wittingly as a cover for the emotional attitude or affective response motivating their opposition, they are not tracking reasons for that opposition, and there is consequently little hope that they will react to pressure from reasons. In short, one should not hope to sway them through public philosophy, be it in the form of discursive hygiene or rhetoric.

For those “unwitting homophobes” who use religious rationalization unwittingly as a cover for the emotional attitude or affective response motivating their opposition, neither are they tracking reasons for that opposition nor do they have any helpful thin self-knowledge, but there is hope that they will react to pressure from the right kind of reasons. That is, they may react to pressure in that they are hateful but unaware of their rationalizations as such and may otherwise consider themselves decent people, which would provide an argumentative foothold for rhetorical pressure, such as appeals to their decency or face-to-face interaction with a person whom they nominally condemn.

Finally, there are those “well-intentioned opponents” who, while not homophobic, have a negative emotional attitude or affective response and find religious teachings a plausible explanation therefor. For these opponents, “reasons are playing a greater role in the formation of their political position in the first place” (Stout, 2010b, p. 533). Accordingly, even if they are mistaken on the reasons for their opposition and are unaware of the upstream processes underlying their opposition (and so lack thin and thick self-knowledge), their opposition tracks or is keyed to reasons rather than affective responses or similarity with their own person. Hence, well-intentioned opponents are more likely to be responsive to pressure on the reasons for their opposition. One would have to “show them that their scriptural reasons for opposing same-sex marriage fail to cohere with other commitments they hold with equal or greater confidence” (*idem.*). Such rhetorical pressure may increase the chances of thin self-knowledge and open a progressive path towards thick self-knowledge through inventorying the person’s religious sources, if not through exposing the upstream belief-formation processes of emotivism. In short, thick self-knowledge and political justification can accommodate a certain degree of post-hoc rationalization provided that interlocutors, public philosophers included, can distinguish the kinds of ill-intentioned, mis-intentioned or well-intentioned reasons adduced and adjust their expectations and interactions accordingly.

Notably, Stout (2010b) also suggests that one might call interlocutors’ attention to the role which non-discursive factors play in upstream belief-formation. For an audience of mis- or well-intentioned opponents, if one were to call attention, for example, to how the distinction between genders and its

importance to the division of labour and inheritance rules, this might also help move the discussion forward (Stout 2010b, pp. 533-534). All the same, changes in the social structure itself over time may do more to move the conversation along than discursive factors employed by public philosophy. In this way, Stout, like Leiter, deems changes in attitude in part attributable to non-discursive factors. In other words, Stout is aware, at least implicitly, of one challenge close to the emotivist, tribalist version posed by Leiter and is ready to adapt.¹⁴

Finally, if Stout admits that “the intuitions from which moral reasoning proceeds are not the same” in different social structures (Stout, 2010b, p. 534), that is, different social structures evince divergent emotional attitudes or affective responses, the author’s insistence on pursuing discussion with mis- or well-intentioned opponents suggests that discursive hygiene may nonetheless play a role in public discourse by shedding light on the upstream belief-formation processes.¹⁵ What is then at issue is whether Stout accepts the scope of the problem traced by Leiter, namely that emotivism and tribalism are constitutive of all instances of attitude- and belief-formation and *prima facie* block thin and thick self-knowledge and justified political belief. Leiter’s own attempts to moderate his critique and Stout’s partial integration of post hoc rationalization suggest that it need not be so.

Turning now to the second tack, we contend that Stout could push Leiter on the conditions necessary to pursue the latter’s vision of efficacious public philosophy. More simply, if the latter’s critique is as thoroughgoing and effective in rebutting public philosophy as it appears, his own attempts to moderate that critique may require buttressing from an outside source. To that end, Stout could point to the way in which successfully applying rhetorical pressure depends in part on discursive hygiene and other-knowledge (either thick or thin). Put differently, the public philosopher or interlocutor may draw on discursive hygiene to arrive at (thin or thick) knowledge of the person’s perspective or inferential commitments from which the former may identify the kinds of

¹⁴ Cf. the enlightening discussion of the alternative popular modes of moral inculcation and development (*e.g.* “poems, novels, essays, plays, and sermons” (Stout, 2004, p. 163) and their importance relative to moral theory (Stout, 2004, pp. 162-168).

¹⁵ In truth, whether Leiter intends to do so, his own exposition of the paradoxes of public philosophy may serve a similar purpose in the end by bringing attention to the way in which tribalism and emotivism shape beliefs, attitudes and reasons.

reasons, at the level of form and content, most likely to apply the right kind of rhetorical pressure to the person's beliefs, attitudes or reasons.

In short, efficacious public philosophy entails knowing one's audience, and knowing one's audience may be partially advanced by discursive hygiene. Accordingly, either Leiter's picture of public discourse, public philosophy and discursive hygiene is right on the broad outlines of the problem but unable to furnish conditions under which one might fruitfully pursue the former or Leiter's picture is imperfect on the broad outlines of the problem but able to furnish conditions for meaningful public discourse, public philosophy and discursive hygiene. In the end, this revised picture seems to dovetail with the moderated version of Leiter's critique upon which the author settles in the essay's closing pages. After all, he allows that logical entailments can, at times, constrain beliefs, attitudes and reasons.

That moderating influence might be further carried out by Stout's extended treatment of alternative argumentative techniques such as "moral perception" in which discursive hygiene and inference nonetheless have an important role to play. Stout (2004) scrutinizes one such exchange in the Edmund Burke-Thomas Paine controversy over democracy and custom (*cf.* pp. 216-224). He concludes thereof that, while the emotional attitudes or affective responses constitutive of moral perception are "noninferential, they are inferentially connected to moral passions, like awe and pity, and the actions for which they serve as warrant" (Stout, 2004, p. 217). In other words, inferential relations unearthed by discursive hygiene may allow us to get a discursive grip on non-discursive factors underlying beliefs, attitudes and reasons. In the end, should we allow, with Leiter, that emotivism and tribalism are principally constitutive of attitude- and belief-formation, this does not preclude their responsiveness to the right kinds of reasons depending on the circumstances, for which discursive hygiene may yet serve a purpose and which Stout is better situated to expose.¹⁶

To conclude, while Stout sketches a stronger connection between self-knowledge and political justification than Leiter thinks reasonable, Stout is ready

¹⁶ For an instance in which Stout's account outpaces Leiter's, compare Leiter's cursory remarks *supra.* on horror and revulsion with Stout's own analyses of anger and grief and the sacred and the horrendous in political deliberation (Stout, 2010a, Chaps. 5, 17).

to work within, at least to an extent, the limits which Leiter sets. Moreover, insofar as Leiter wishes to preserve some role discursive hygiene and endorses philosophical rhetoric, he has need of a position like Stout's to flesh out his new discursive practice.¹⁷ While thick self-knowledge and justified political belief will prove challenging for most and wholly out of reach for others, this in no way prevents interlocutors in public discourse from making full use of a mixed argumentative strategy to advance the end of self-knowledge and political justification.

5. Conclusion

Rather than attack post-truth and post-justification head-on, we set out from the question of whether self-knowledge and political justification evince a necessary connection. To that end, we sought to test one version of that connection through its illustration in Stout (2004) as well as its implications for public philosophy. To clarify ideas, we began in the first part by laying out a framework for understanding self-knowledge as thin or thick and justification as rational commitment, justifiable belief and justified belief, the latter inspired by Pryor (2004). From there, we formulated four possible theses for the connection between self-knowledge and political justification and tentatively identified Stout (2004) with the strong and thick version whereon thick self-knowledge is necessary for political justification. In the second part, we linked the strong and thick thesis to the Stoutian justificatory devices of democratic piety, earnest storytelling and Brandomian expressive rationality. After underlining Stout's need for a strong public philosophy, we turned in the third part to Leiter's (2016) vigorous emotivist and tribalist critique of public philosophy as discursive hygiene and explored why these forces hinder public discourse, public philosophy and discursive hygiene. In the fourth and final part, we worked out two potential Stoutian replies to Leiter's critique on which public philosophy and discursive hygiene can adapt to public discourse's emotivist and tribalist tendencies. Ultimately, we judged that Leiter's critique, in its unadulterated form, undercuts the more moderate public philosophy which he envisions and lacks the tools to

¹⁷ Indeed, Stout and Leiter's views seem complementary to the point that one could well have begun this paper by presenting Leiter's qualified view of public philosophy before subjecting it to Stout's own caveats rather than the chronological presentation favoured here. We owe this observation to Brandon Robshaw from a question-and-answer session.

advance that vision. In light of the above, we find that Stout provides a powerful and grounded illustration of the role played by self-knowledge in political justification. All in all, the strong and thick version of the link between self-knowledge and political justification remains possible albeit fraught with the complications and hazards of actual belief-formation. Whatever might be said for the post-truth world, it does not yet appear to be post-justification, nor post-public philosophy.

References

Bächtiger, A. and Gerber, M. (2014). 'Gentlemanly Conversation' or Vigorous Contestation?: An Exploratory Analysis of Communication Modes in a Transnational Deliberative Poll® (Europolis). In K. Grönlund, A. Bächtiger, and M. Setälä (Eds.), *Deliberative Mini-Publics: Involving Citizens in the Democratic Process* (pp. 115-134). Colchester, UK: ECPR.
http://press.ecpr.eu/book_details.asp?bookTitleID=57

Brandom, R. (1994). *Making It Explicit: Reasoning, Representing and Discursive Commitment*. Cambridge, MA: Harvard University Press.
<http://www.hup.harvard.edu/catalog.php?isbn=9780674543300>

Brandom, R. (2000). *Articulating Reasons: An Introduction to Inferentialism*. Cambridge, MA: Harvard University Press.
<http://www.hup.harvard.edu/catalog.php?isbn=9780674006928>

Dryzek, J.S. (2000). *Deliberative Democracy and Beyond: Liberals, Critics, Contestations*. Oxford, UK: Oxford University Press.
<http://www.oxfordscholarship.com/view/10.1093/019925043X.001.0001/acpr of-9780199250431>

Frohlich, N., Oppenheimer, J.A., & Eavey, C.L. (1987). Choice of principles of distributive justice in experimental groups. *American Journal of Political Science*, 31(3), 606-636.
https://www.jstor.org/stable/2111285?seq=1#page_scan_tab_contents

Lafont, C. (2013). Religion and the Public Sphere: What are the Deliberative Obligations of Democratic Citizenship?. In C. Calhoun, E. Mendieta, & J. VanAntwerpen (Eds.), *Habermas and Religion* (pp. 230-248).

Cambridge, UK: Polity Press.

<http://eu.wiley.com/WileyCDA/WileyTitle/productCd-074565326X.html>

Leiter, B. (2016). The Paradoxes of Public Philosophy. *Indian Journal of Legal Theory*, 1(1), 51-64.

http://centreforlegaltheory.in/Mindfields_files/issues/issue1.pdf

Pryor, J. (2004). What's Wrong with Moore's Argument?. *Philosophical Issues*, 14(1), 349-378. <http://onlinelibrary.wiley.com/doi/10.1111/j.1533-6077.2004.00034.x/abstract>

Rawls, J. (1971). *A Theory of Justice*. Cambridge, MA: Harvard University Press. <http://www.hup.harvard.edu/catalog.php?isbn=9780674000780>

Rawls, J. (1993). *Political Liberalism*. New York, NY: Columbia University Press. <https://cup.columbia.edu/book/political-liberalism/9780231130899>

Stout, J. (1988). *Ethics after Babel: The Languages of Morals and Their Discontents*. Boston, MA: Beacon Press.

<http://press.princeton.edu/titles/7049.html>

Stout, J. (2004). *Democracy and Tradition*. Princeton, NJ: Princeton University Press. <http://press.princeton.edu/titles/7667.html>

Stout, J. (2010a). *Blessed are the Organized: Grassroots Democracy in America*. Princeton, NJ: Princeton University Press.

<http://press.princeton.edu/titles/9279.html>

Stout, J. (2010b). Rorty on Religion and Politics. In R.E. Auxier & L.E. Hahn (Eds.), *The Philosophy of Richard Rorty* (pp. 523-545). Chicago, IL: Open Court Press. http://www.opencourtbooks.com/books_n/philosophy_rorty.htm



HOBBS: A VOLUNTARIST ABOUT THE PERMISSIBILITY OF STATE ENFORCEMENT?

HOBBS: UM VOLUNTARISTA ACERCA DA PERMISSIBILIDADE DA AUTORIDADE ESTATAL?

Daniel Guillery*
d.guillery.12@ucl.ac.uk

Abstract. *I take up the question of what argument, if any, Hobbes has for state legitimacy, which term I stipulatively use to mean the general, exclusive permission to enforce compliance with their directives or laws that states are standardly taken to have. I will argue that, contrary to what one might imagine, the ground of state legitimacy for Hobbes is not to be found in the social contract or the authorisation of the state's subjects, but rather in the sovereign's simply not being subject to the kind of laws that rule out enforcement for subjects. The sovereign's right to enforce is based in exactly the same sort of right that all have when not subject to any higher sovereign power. Though this must be nuanced (the sovereign does not literally retain its right to all things from the state of nature, since no sovereign existed in the state of nature), the permissibility of enforcement for Hobbes is to be found simply in the lack of anything that might make it impermissible.*

Keywords: *Hobbes, legitimacy, voluntarism, enforcement, political obligation.*

Sumário. *Trato a questão de saber que argumento tem Hobbes a favor da legitimidade do Estado, um termo que uso para designar a permissão geral e exclusiva para impor a obediência com as directivas e leis que os Estados tipicamente têm. Irei argumentar que, ao contrário do que podemos imaginar, o fundamento da legitimidade do Estado para Hobbes não se encontra no contrato social ou na autorização dos súbditos do Estado, mas antes no facto de o soberano não estar sujeito ao tipo de leis que retiram aos súbditos a autoridade de impor. O direito do soberano impor a lei baseia-se exactamente no mesmo tipo de direito que todos têm quando não estão submetidos a um poder soberano superior. Embora isto deva ser qualificado (o soberano não mantém literalmente o seu direito natural a todas as coisas, uma vez que não existia qualquer soberano no estado de natureza), a permissibilidade da autoridade encontra-se, em Hobbes, simplesmente na falta de algo que poderia tornar essa autoridade inaceitável.*

Palavras-chave: *Hobbes, legitimidade, voluntarismo, imposição, obrigação política.*

* PhD student in philosophy at University College London.

o. Introduction

There are at least three important dimensions of moral evaluation of states. I will call these *obligating power*, *legitimacy* and *justification*. By the first two terms I mean the following:

Obligating power. A state X has obligating power over an individual Y in some domain D if and only if X's commands/directives in D create duties/obligations/pre-emptive reasons for Y to act as commanded/directed, independent of their content (within some limits).¹

Legitimacy. A state X is legitimate with respect to individual Y if and only if X has a general and exclusive moral permission to enforce Y's compliance with its commands/directives/laws independent of their content (within some limits).²

I mean these definitions to be purely stipulative.³ Obligating power is just the flipside of what is often called 'political obligation': if subject Y has a political

¹ On pre-emptive reasons see Raz (1988) ch. 3 and (1979) 18; a pre-emptive reason to ϕ is a reason that pre-empts, excludes from consideration or replaces certain other reasons for or against ϕ -ing. (My label might be slightly misleading since one might have a pre-emptive reason to ϕ without having an obligation to ϕ). X's having obligating power over Y involves (in Hohfeldian terms) it having a *moral power* to change Y's normative situation, but it is consistent with X's also having a *claim right* to Y's obedience or not. For discussion of this distinction see Copp (1999) 10-21. For X's commands/directives to *create* a duty (etc.) need not require that X's command be in any sense the *ultimate* source of the duty. For instance, if Y promises to obey whatever commands X makes, and then on some subsequent occasion X commands Y to ϕ , Y now has a duty to ϕ that she did not have before. X has *created* this duty, even though, in some sense, its ultimate source is Y's promise, not X's command; Y only has the duty to ϕ as a part of the more general duty to do whatever X commands.

² What is meant here by 'content-independence' is slightly more straightforward for obligating power than for legitimacy. In the former case, X's having content-independent obligating power over Y can be understood as it being the case that, for any ϕ , *that X commanded it* is sufficient for, and the reason for, Y's being required to ϕ *so long as* some conditions are met (and these conditions give the limits mentioned in the definition above). For legitimacy, content-independence will have to be cashed out as something like its being the case that X gets to decide what commands it enforces compliance with, that is, what it is permissible for X to enforce is not wholly determined by some other criterion (other than X's decision), such as the independent requirements of morality.

³ The term 'legitimacy' is a highly contested term whose use is much debated by philosophers. I do *not* intend my definition as an analysis of the term, nor as a contribution to the debate about how it should be used. I use the term because the property I mean to pick out is one that has often been closely associated with that term and I think some

obligation (a content-independent obligation to obey the commands of her state), the state has obligating power over Y. To have obligating power over Y is just for it to be the case that one's making a command or directive creates a duty/obligation/pre-emptive reason for Y to act accordingly.

The third dimension, which I am referring to as the question of 'justification' asks whether the *existence* of a state is, in some sense *a good thing* or *morally acceptable*. (This is, I think, what A. John Simmons has in mind when he famously distinguishes justification from legitimacy, though he uses the latter term differently from me).⁴ The word 'justification' is often used in connection with actions, as when we ask whether someone is *justified in doing* something. This is not the way I mean to use the term here. Existing is not an action, and entities cannot be *justified in existing* in this sense. It may be that when people talk about *justifying the state*, or showing that a state is *justified*, they mean to refer elliptically to showing that a state is *justified in doing* certain things (perhaps those things essential to or characteristic of states). Since enforcement is often taken to be essential to states, showing that a state is *justified in doing* things essential to its statehood may end up amounting to establishing what I above called its *legitimacy* (which I will turn to below). But I think that talk of states being justified may also refer to something different, having to do with the moral status of the state's *existence*. When I refer to a state's *justification* or being *justified*, I refer exclusively to this latter thing. There are still several different things that might be meant here, but I think there is a family of questions having to do with the quality of the state and whether it is something whose existence we should accept or be glad for: it could be that it is justified if and only if it is better that it exist than not, or if it meets some evaluative threshold, or something else along these lines. This family of questions is quite separate from what I have called obligating power and legitimacy.⁵

writers have used the term as a name for this property, but if the use of the contested term is confusing, then any other name can be substituted.

⁴ Simmons (2001) 122-57

⁵ Simmons also talks about justifying the state by 'showing that some realisable type of state is on balance morally *permissible*' (2001, 125-6). Institutions are not really candidates for permissibility; it is actions that are either permissible or not, but a state's being justified could be understood as its being morally permissible to create that state

It should be clear that these three dimensions of evaluation are at least distinct *conceptually*. Hobbes might be read with an eye to any of them. The easiest to discern is justification and Hobbes also takes himself to have an argument for obligating power, but here I want to consider what materials there are in Hobbes for a defence of *legitimacy*.

Hobbes's well-known state-of-nature argument is most obviously associated with justification. The familiar argument, very roughly, is that human nature is such that 'during the time men live without a common power to keep them all in awe, they are in that condition which is called war'.⁶ Thus, life without a state is awful and so 'it is a precept, or general rule of reason, *that every man, ought to endeavour peace, as far as he has hope of obtaining it*'.⁷ However, the only way to ensure that this is acted on is to have a state (indeed, an absolute state). It seems quite clear how this can give us a justification of the state, what David Schmidtz calls a 'teleological justification'.⁸ A state is teleologically justified by its fulfilling certain goals. Here, the state is justified by its being necessary for the achievement of peace. A teleological justification of an institution gives us reason to think of it as good in some way, and it may become a premise in an argument that we *ought* to create (or promote or preserve) that institution. Hobbes also takes himself to have given an argument for obligating power. I will

or impermissible to destroy it or something similar. Indeed, Hobbes's justification argument might well be interpreted as an argument to the conclusion that it is permissible to create the state. (It is plausibly interpreted as an argument that it is permissible to create *any* state because any state is better than the state of nature). Hampton (1986, 269) argues that social contract arguments (including Hobbes's) are supposed to show just such a thing. I take it, though, that if this is Hobbes's conclusion, it is supposed to follow from an intermediate conclusion more like how I have characterised a state's being justified above. If Hobbes thinks that it is permissible to create a state this is presumably because of its being *good* in some way, its possessing certain virtues (at least comparatively with the state of nature). I do not think that a conclusion about the permissibility of creating a state follows directly from its meeting some evaluative threshold, its being better that it exists than not, or its being better than the state of nature (as I have argued elsewhere). Thus, I will treat Hobbes's justification argument as just an argument that the state is *good* in some way rather than an argument for the permissibility of creation.

⁶ Hobbes (1996) 84 (13:8:62). I make reference to the Oxford edition that I am using, but I also provide in parentheses references in the following form: (chapter: paragraph: page number of the 1651 edition).

⁷ Ibid. 87 (14:4:64)

⁸ Schmidtz (1990) 102

not attempt to identify what exactly it is, but it must be something additional to the argument for justification. If we take seriously what Jonathan Wolff calls Hobbes's 'self-assumption principle',⁹ the principle that there is 'no obligation on any man which ariseth not from some act of his own', a teleological justification of the state is not sufficient to establish an obligation to obey on the part of its citizens.¹⁰ This is the basis of Wolff's objection to a hypothetical contract interpretation of Hobbes: 'if I am to have an obligation to obey the sovereign', given the self-assumption principle, '[i]t cannot be a consequence of something I merely would have done'.¹¹ More generally, if the self-assumption principle holds, obligation to obey will not follow directly from a state's being justified, its possessing certain virtues.¹²

Legitimacy is (at least conceptually) independent of both. That one has an obligation (or duty or pre-emptive reason) to ϕ does not immediately entail that anyone is permitted to enforce your ϕ -ing. Certain moral requirements may be such that, in virtue of being under such a requirement, one is liable to be permissibly coerced by certain others, but this is not a general feature of moral requirement. For instance, one may have an obligation to keep a promise to a friend consistently with nobody else being permitted to enforce your keeping the promise. In particular, that one has an obligation to obey some institution does not, without additional premises, entail that that institution is permitted to enforce your compliance.¹³ Additionally, legitimacy does not *immediately* follow

⁹ Wolff (1994) 271

¹⁰ Hobbes (1996) 144 (21:10:111)

¹¹ Wolff (1994) 274

¹² Simmons, in the titular essay of his (2001), argues that what he calls 'legitimacy' (something like the conjunction of what I call 'legitimacy' and obligation, with the key part being obligation) does not follow from justification. Hume perhaps *does* argue for obligation from justification (see Hume (1994), e.g. essays 4 and 23), but he does not hold the self-assumption principle, and he will need additional premises to get from justification to obligation.

¹³ Both Green (2004) and Wellman (1996), who note that the question of obligating power is separate from the question of legitimacy, seem to think that political obligation *does* entail the legitimacy of state coercion; that is, they think that it is not possible for a state to have obligating power but no permission to enforce. Neither, though, offers any argument for this claim and there seems to be no reason to think that it must be permissible for someone to enforce a set of rules in order for one to have an obligation to obey those rules.

from a state's being justified in the above sort of sense (roughly, its existence being a good thing) that it is permitted to enforce its directives over its subjects. A state's *quality* and what it is permitted to do are two separate things. Thus, we cannot expect Hobbes's argument for legitimacy, if he has one, to be just identical to his arguments for either justification or obligation. What, then, can we find in Hobbes?

In this paper, I address this question. I will begin by asking what argument Hobbes has for the legitimacy of states *actually* formed by the sort of contract he describes, and then I will ask how this transfers to states not formed in this way. I first consider what might seem to be a natural, voluntarist interpretation of his argument for legitimacy, which bases it in the voluntary *authorisation* of the covenanters. However, I argue, this cannot be Hobbes's argument for legitimacy. I then offer an alternative interpretation, according to which Hobbes has, and needs, no special argument for state legitimacy, since there is simply nothing that might make state enforcement *impermissible*.

1. Actual contract

Let us begin by considering Hobbes's account as describing an actual contract. One way of interpreting Hobbes's argument from the state of nature is as giving a historical account of how states were formed. If Hobbes gives an account of how states acquire legitimacy when formed by actual contract, then this will apply to actual states insofar as the historical picture is an accurate description of those states. However, if, as seems more plausible, Hobbes's argument should not be given a historical interpretation, we can first see how, on Hobbes's view, states acquire legitimacy in the hypothetical scenario in which they are formed by contracting out of the state of nature, and then ask how this transfers to the question of legitimacy for actual states.

1.1 Authorisation

There is something in Hobbes's story of the formation of states that looks somewhat like an account of legitimacy, that is, the permissibility of state enforcement. This is his story of authorisation. Artificial persons, who represent the actions of another person or other people, may 'have their words and actions

owned by those whom they represent'. In this case, the owner of the actions is their *author*, and the actions are performed by authority: 'by authority, is always understood a right of doing any act: and *done by authority*, done by commission, or licence from him whose right it is'.¹⁴ Thus, we can see what looks like the shape of a Hobbesian argument that certain actions by certain agents are *permissible*: if A is *authorised* by B to ϕ and B has a right to ϕ , then so does A.¹⁵ Hobbes goes on to employ this notion of authorisation in his account of the formation of a commonwealth. The latter he defines as '*one person,¹⁶ of whose acts a great multitude, by mutual covenants one with another, have made themselves every one the author, to the end he may use the strength and means of them all, as he shall think expedient, for their peace and common defence*'.¹⁷ The covenant is made 'in such manner, as if every man should say to every man, *I authorize and give up my right of governing myself, to this man, or to this assembly of men, on this condition, that thou give up thy right to him, and authorize all his actions in like manner*'.¹⁸

Thus, it seems that Hobbes takes the state in his state of nature contract story to be *authorised* by its subjects and perhaps this is his view about how it comes to be permissible for the state to enforce its directives on its subjects. Enforcement is permissible because the state's actions, including its coercing its citizens, are *owned* by all of its subjects. Indeed, Hobbes argues that the sovereign's actions, whatever they may be, cannot constitute injury or injustice to its subjects:

For he that doth anything by authority from another, doth therein no injury to him by whose authority he acteth: but by this institution of a commonwealth, every particular man is author of all the sovereign doth: and consequently he that

¹⁴ Hobbes (1996) 107 (16:4:81)

¹⁵ By 'right' here I mean a liberty right (in the Hohfeldian sense) or a moral permission. It seems that this is also how Hobbes uses the word 'right': see Hampton (1986) 51.

¹⁶ An *artificial* person may be constituted by a multitude of *natural* persons.

¹⁷ Ibid. 114 (17:13:88)

¹⁸ Ibid. 114 (17:13:87)

complaineth of injury from his sovereign, complaineth of that whereof he himself is author; and therefore ought not to accuse any man but himself; no nor himself of injury; because to do injury to one's self, is impossible.¹⁹

Anything the state may do to its citizens is permissible, it might seem, because its actions are *their* actions. (At least this will be so if we take the authorisation of the state by its subjects in the Hobbesian contract to be complete, rather than an authorisation to perform *certain* actions only). Thus, we might think that there is an authorisation argument for legitimacy in a historical interpretation of Hobbes that goes something like this:

1. Nothing done to oneself can be impermissible.²⁰

2. In the formation of the state by social contract, its subjects *authorise* it to enforce its directives on them.

3. If A authorises B to do X, B's future actions of X-ing become *owned* by A, that is, they become A's actions.

4. Therefore, (from 2 and 3), when a state formed by social contract enforces its directives on its subjects, they enforce these directives on themselves.

5. Therefore, (from 1 and 4), it cannot be impermissible for a state formed by social contract to enforce its directives on its subjects.

6. Therefore, a state formed by social contract has a general moral permission to enforce its directives on its subjects.

This, then, is an argument for legitimacy given an actual contract that stands independently of the argument for justification (which is the argument that it is *rational* to authorise the state).

In order to see if this is indeed Hobbes's argument for legitimacy within an actual contract, we should consider in more detail what exactly he takes to be involved in authorisation. Hobbes is not too clear on this point, but it seems that,

¹⁹ Ibid. 117-8 (18:6:90)

²⁰ This is one way of reading Hobbes's claim that 'to do injury to one's self, is impossible'. It is to read 'injury' as 'impermissible treatment'. If we are to read the paragraph quoted above as Hobbes's argument for legitimacy, we will need to interpret 'injury' in this way.

on the most plausible interpretation, authorisation will be some sort of giving up of right, a transfer or surrender of right. Since all have a natural right of self-preservation, Hobbes thinks, in the state of nature, which is a condition of war, 'every man has a right to every thing'.²¹ In authorising the sovereign, each subject either transfers or surrenders some part or the entirety of this right to all things. This is suggested by the passage quoted above in which Hobbes comes closest to giving a definition of 'authorisation', saying that 'done by authority' means 'done by commission, or licence from him whose right it is'.²² Additionally, when Hobbes describes the covenant that creates the commonwealth, he describes it as a covenant to *authorise* the sovereign and he is quite clear that this covenant is one in which the covenanters give up their right to all things, or their right of governing themselves.²³ However, *which* of these two things it is (transfer or surrender) is a point of dispute in the literature.

David Gauthier claims that 'authorisation must involve some translation of right'. He and Gregory Kavka understand authorisation along the lines of a principal-agent model where the principal temporarily grants or loans the agent some power or right she has.²⁴ On this interpretation, authorisation involves A's bestowing a right on B, a right that A retains. A gives B the right to represent A by granting B the *use* of A's right: 'authorisation (...) enables *you* to act in my place, and so with *my* right'.²⁵ Since B, being authorised by A, is acting as A's representative, that is, with A's right, A *owns* B's actions in the normal, modern sense. B's actions, because done with A's right, are (as if) A's actions. The same obligations, responsibilities, rights, commitments that A would incur by ϕ -ing herself are incurred when B ϕ s under A's authority. On this interpretation, the authorisation of the sovereign is seen as an *agreement* with some normative force (the agreement alters the normative landscape, by transferring rights from one

²¹ Ibid. 87 (14:4:64)

²² Ibid. 107 (16:4:81)

²³ See pg. 114 (17:13:87-8) and pg. 87 (14:5-6:64-5)

²⁴ Gauthier (1969) 124; Kavka (1986) 389-90.

²⁵ Gauthier (1969) 124

person to another). (Hampton, by contrast, sees the authorisation of the sovereign to be a purely behavioural abandonment of right, constituted by the subjects' coming to obey the sovereign.) Such an understanding of authorisation lends itself well to the form of argument above. There is a sense in which, when B's ϕ -ing is authorised by A, A owns B's ϕ -ing and is *responsible* for it. It might be thought that whatever permissibility facts apply to actions performed by an agent also apply to actions that agent is responsible for. If this is the case, it becomes possible to make something like the argument given above.

However, Jean Hampton disputes this interpretation, claiming instead that authorisation involves a *renunciation* or *surrender* of right.²⁶ She argues that such a 'loan' of right could not be the basis for Hobbes's absolutist theory and points to passages in which Hobbes describes the covenant as involving 'laying down' or 'divesting oneself' of rights.²⁷ Understanding authorisation as Hampton does gives a clear picture of how authorisation *creates* the sovereign. By abdicating their right to all things, the covenanters cease (at least given certain conditions) to pursue whatever means they see to be most appropriate for their own protection and instead obey the commands of the sovereign. In Hampton's view, authorisation of the sovereign just is constituted by one's coming to obey the sovereign's punishment commands.²⁸ The sovereign is created as it gains the ability to enforce its commands thanks to the generalised obedience resulting from the covenanters' abdication of their right to all things.

If we accept a Hampton-style interpretation of authorisation, the question arises what authorisation has to do with ownership; in other words, what grounds

²⁶ Hampton (1986) ch. 5. Tom Sorell (1991) similarly interprets the laying down of right as involving simply giving 'way to another in his pursuit or enjoyment of a thing' (112). He makes a different distinction between transfer and renunciation to the one made here, where the former involves giving way to only one particular person, while the latter involves giving way to *anyone*. In either case, though, what he takes to happen is a surrender in the sense under discussion here, not a temporary loan.

²⁷ Hobbes (1996) 114 (17:13:87-8) and 87 (14:5-6:64-5)

²⁸ She says it involves '1) participating with the other inhabitants of the state of nature in a process in which one of them is selected as sovereign ... and 2) obeying the punishment commands a) of only this individual b) to refrain from interfering in the punishment of another and c) to actively assist in the punishment of another, insofar as these commands are (or have been made by the sovereign to be) individually rational' (Hampton (1986) 186).

are there for asserting premise 3 of the above argument if authorisation is just surrender of right? How does one come to *own* the actions of another by simply surrendering one's right to all things to them? Hampton's answer is that in 17th century English, the words 'authorise' and 'own' could have senses slightly different from their ordinary modern senses. She points out that 'to authorise' could mean 'to set up as authoritative; to acknowledge as possessing final decisiveness ... to give final approval to; to sanction, approve, countenance', while 'to own' could mean 'to acknowledge as having supremacy, authority or power over one; to profess, or yield, obedience or submission to'.²⁹ The former is a sense still possessed by the term 'to authorise', but the latter is obsolete. Thus, the subjects do not *own* the actions of the sovereign in the usual sense; rather, they acknowledge its actions as having supremacy, or something along these lines. How do these senses of 'authorise' and 'own' work with the idea that the covenanters authorise the sovereign by surrendering their right to all things? Perhaps the thought is that *by* surrendering their right to all things the covenanters set up the sovereign as authoritative; the surrender *expresses* final approval or something similar.

Can we make an argument like the above given such an interpretation? Surrender of right as simple factual obedience of commands on its own will not get us such an argument. That is, it will not get us such an argument *unless* it is thought that this obedience of commands constitutes in some way something like an acknowledgement of the supremacy or authority of the actions of the commander. Thus, the expression of approval or acknowledgement of authority or some such similar thing that Hampton takes Hobbes to be referring to is essential. These may be constituted by the obedience, but if they are, it is in its role as an expression of approval or acknowledgement of authority that the obedience will be relevant to the argument. However, even when we add these understandings of 'authorise' and 'own' to de facto surrender of right, we still cannot quite make the argument as presented above. Once we understand 'own' as Hampton suggests Hobbes does, we can no longer make the equivalence between A's owning B's actions and B's actions being A's actions. That I

²⁹ Ibid. 126-7. She cites the *Oxford English Dictionary* for the senses of these words.

acknowledge your actions ‘as having supremacy, authority or power over’ me does not make your actions my actions. The best version of the above argument we can make when we understand ‘authorise’ and ‘own’ as Hampton does is something like the following:

1. If A expresses approval of/acknowledges as permissible/engages to support... B’s doing ϕ to A, it is permissible for B to do ϕ to A.³⁰

2. In the formation of the state by social contract, its subjects *authorise* it to enforce its directives on them.

3. If A authorises B to do X, A expresses approval of/acknowledges as permissible/engages to support B’s future actions of X-ing.

4. Therefore, (from 2 and 3), when a state formed by social contract enforces its directives on its subjects, its doing so is approved of/acknowledged as permissible/supported by the subjects.

5. Therefore, (from 1 and 4), it is permissible for a state formed by social contract to enforce its directives on its subjects.

The problem with this version of the argument is that it depends on the very questionable new version of premise 1. While the original version of premise 1 (‘nothing done to oneself can be impermissible’) may be disputable, it has at least *prima facie* plausibility. This one seems to lack even this: it does not seem plausible to think that people are always infallible about how it is permissible to act towards them.³¹ I cannot, though, see any better way to construct an argument for legitimacy from the chapter 18 passage reading ‘own’ as Hampton does. This need not show that Hampton is wrong to interpret Hobbes’s use of ‘authorise’ and ‘own’ as she does. If she is right about the meaning of ‘own’, we can alternatively read the passage as not being about permissibility. (This, I think, as will become clear below, is in fact the correct way to read this passage).

³⁰ The OED definition Hampton cites does not include acknowledgement as permissible or engagement to support, but these seem in the spirit of it and might do better for getting to legitimacy.

³¹ Cases of adaptive preferences seem to be clear cases where victims of impermissible treatment wrongly see that treatment as permissible.

In any case, though, whichever of these is the correct interpretation of Hobbes's understanding of authorisation, Hobbes cannot hold that the legitimacy of the sovereign is derived from authorisation (and so the passage from chapter 18 quoted above cannot be Hobbes's argument for legitimacy).³² Hobbes thinks that, since the purpose of entering into a covenant is to achieve some good for oneself, there are certain rights that are inalienable by covenant.³³ Thus, 'the promise of not resisting force, in no covenant transferreth any right; nor is obliging'.³⁴ For this reason Hobbes says that 'no man is supposed bound by covenant, not to resist violence; and consequently it cannot be intended, that he gave any right to another to lay violent hands upon his person'; and then, 'it is manifest therefore that the right which the commonwealth (that is, he, or they that represent it) hath to punish, is not grounded on any concession, or gift of the subjects'.³⁵ This seems to be a fairly clear statement that Hobbes does not take any version of the above argument to be the source of the state's liberty right/permission to enforce. The above argument takes the source of the permission to be the subjects' individual voluntary authorisation of the sovereign through covenant (whether through surrender or transfer of right), but Hobbes here appears to be denying that such a thing could ground any permission to punish, which is an essential part of the permission to enforce. It is not possible, on Hobbes's view, to give up the right to self-defence to the benefit of another. As Hobbes recognises, the right to punish will not amount to much without the right to use *violence* to punish. Similarly, the right to enforce will not amount to much

³² I do not think it is necessary for me to decide between these two interpretations; the account I will give below is consistent with both. For what it is worth, though, I think that there is probably an element of both in Hobbes. I think without Hampton's behavioural *surrender* of rights, the sovereign would not have the *ability* to enforce that makes it what it is. And, I think, either *in surrendering* the right to all things, or *as well*, the covenanters *authorise* the sovereign. That is, they take ownership of the actions of the sovereign. Each subject permits the sovereign to represent her, to act in her name. I don't think this should be understood as a transfer of the right to all things, but as an authorisation to decide what *I* do. The sovereign doesn't gain a permission to do any acts that it wasn't permitted to do before, but it can now (it has the ability or power to), by acting, determine what *I*, as a covenanter, do, because its actions represent me. It gains 'the *right to present* the person of them all' (115 (18:1:88)).

³³ Ibid. 88-9 (14:8:65-6)

³⁴ Ibid. 93 (14:29:70)

³⁵ Ibid. 205-6 (28:2:161)

without the right to punish. Thus, Hobbes appears to be telling us explicitly that the above is not his argument for legitimacy. Furthermore, even if these explicit statements could be explained away somehow, as will become clear below, it would not *make sense* for Hobbes to attempt to derive the state's legitimacy from its authorisation by its subjects. Though he seems to be a voluntarist about obligation, he is not a voluntarist about the permissibility of enforcement.

1.2 State enforcement rights by default

If, then, what looked like an argument for legitimacy from authorisation (at 18:6) was not in fact that, are we to conclude that Hobbes does not think states have a moral permission to enforce their commands, or that there is some other way in which the creation of the sovereign bestows upon it a permission to enforce its commands on its subjects? In fact, we should conclude neither of these things. Hobbes does not need any special argument to establish the state's moral permission to enforce. There is nothing at all special for him about this permission.

In the state of nature Hobbes thinks there are no moral constraints on what it is permissible for people to do to each other.³⁶ The right of nature is the right to all things. It is permissible to pursue one's own interest through whatever means of violence or coercion are conducive to it.³⁷ When inhabitants of the state

³⁶ Ibid. 85 (13:13:63)

³⁷ Hobbes appears to be entirely sceptical about the existence of state-independent and subject-independent moral facts (at least if moral facts are to be understood as categorical imperatives). Some have denied this but I think it is hard to see Hobbes in any other way. He seems to think that whatever moral facts exist are dependent on an individual's desires: 'whatsoever is the object of any man's appetite or desire; that is it, which he for his part calleth *good*: and the object of his hate, and aversion, *evil*' (Hobbes (1996) 35 (6:7:24)). See, for instance, Hampton (1986) and Tuck (1996) for different interpretations of Hobbes's moral views in line with this. Hampton's interpretation, which is quite appealing, would have the consequence that there *are* state-independent moral facts in Hobbes's sense, but for Hobbes on Hampton's view moral facts are just hypothetical imperatives, instrumental requirements imposed by our own self-interest. Most importantly, though, however we understand Hobbes's moral views, it seems fairly clear that if there are any moral requirements, the only ones that apply in the absence of a common coercive power are the requirements to seek peace and to be prepared to covenant into a state (his first and second laws of nature). The laws of nature, Hobbes thinks, apply to everyone and are binding. (On Hampton's view, this is because the desire they are hypothetical on is a universal desire, the desire of self-preservation). If Hobbes thinks there are moral imperatives, they are presumably given by these. However, they are all (including the second) derived from the first (the requirement to seek peace) and

of nature covenant together to establish a sovereign, they leave the state of nature, as there is now an effective coercive power over them that can establish effective binding rules, decide disputes and so on. In covenanting, they give up their right to all things and what they are permitted to do becomes determined by the sovereign.³⁸ However, the sovereign does not covenant with its subjects, nor with anyone else and so retains its right of nature with respect to them (and with respect to other sovereigns). Thus, the sovereign does not come to have obligations in a commonwealth in the same way that the subjects do, since their obligations are just what the sovereign determines them to be (and the sovereign is not bound by its own laws).³⁹ The sovereign retains the right to all things; unlike the subjects, the sovereign never surrenders or transfers its right of nature to anybody else. This makes clear also how the sovereign gets an *exclusive* right to enforce (or to determine who may enforce): everybody else has given up their right to all things, and so what they are permitted to do is now determined by the laws the sovereign enforces.

It would not make sense for Hobbes to base legitimacy on authorisation, since in surrendering or transferring their right to all things the subjects do not give the sovereign any right to do anything it does not already have a right to do: it already has a right to do anything. The surrender or transfer of right is not without significance since the behavioural obedience that goes along with it is what makes the sovereign's enforcement *possible* and it is relevant to the subjects' *obligations* and *attitudes*. But it cannot be the basis of what the sovereign is *permitted* to do, since it adds nothing in this respect.⁴⁰ Hobbes puts the point thus:

(excluding the second) only binding when a common power is in place. So long as there is no common power, there are no restrictions on how one individual may treat another.

³⁸ The laws of nature are derivable from reason alone (a derivation not dependent on any choices of the sovereign). However, they are only binding on action if the sovereign enforces them.

³⁹ See Hobbes (1996) 176 (26:6:138)

⁴⁰ Gauthier (1969) claims that authorisation *does* give the sovereign a new right, namely the right to act for another (124). However, I cannot see what this is supposed to amount to, since there is no resulting change in the extension of the sovereign's permissions: the range of actions that is permissible for the sovereign is not altered by authorisation. In particular, authorisation can make no change to the permissibility of enforcement.

I have (...) showed formerly that before the institution of commonwealth, every man had a right to every thing, and to do whatsoever he thought necessary to his own preservation; subduing, hurting, or killing any man in order thereunto. And this is the foundation of that right of punishing which is exercised in every commonwealth. For the subjects did not give the sovereign that right; but only in laying down theirs, strengthened him to use his own, as he should think fit, for the preservation of them all: so that it was not given, but left to him, and to him only; and (excepting the limits set him by natural law) as entire, as in the condition of mere nature, and of war of every one against his neighbour.⁴¹

Thus, the sovereign is permitted to enforce its commands because it is not in a condition in which the sort of constraints that might make it *impermissible* to enforce one's commands on others exist. (This must be nuanced slightly, as suggested by Hobbes's caveat 'excepting the limits set him by natural law', which I will discuss below). These latter only exist between people subject to a common coercive power and a sovereign, by definition, is subject to no such thing.⁴² The sovereign gains the *ability* to coerce *you* from *other* people's abdication of the right to all things, which amounts to a commitment to obey (or just the behavioural fact of obedience), but the abdication of right does not give the sovereign any permission right that it did not have before. In addition, the abdication of right could yet be the basis of Hobbes's argument for *obligation*, but not his argument for legitimacy.

If this is correct, how are we to interpret the passage from chapter 18 that suggested an authorisation-based argument for legitimacy? I think we can understand this passage if we do not interpret 'injury' as meaning simply 'impermissible harm'. It must be something that it is possible to do in the state of nature, where everything is permissible, since if injury were only possible under a sovereign, we would expect Hobbes's argument for the impossibility of injury by the sovereign to be based on the fact that the sovereign creates the conditions for the possibility of injury (echoing his argument that the law cannot be unjust). On the other hand, it cannot be simply harm, since it obviously *is* possible to harm oneself, *contra* what Hobbes says in the passage. Perhaps then it is something like 'harm at which one can rightfully feel aggrieved, or to which one can rightfully retaliate'. If we take literally Hobbes's talk of the sovereign's actions to its subjects being equivalent to actions they do to themselves, then clearly one

⁴¹ Hobbes (1996) 206 (28:2:161-2)

⁴² The requirements of justice are entirely established by the law, to which the sovereign is not subject (176 (26:4:137)).

cannot feel aggrieved at what one does to oneself. If we do not take this talk literally, we can instead read the passage as claiming some sort of inconsistency in recognising or acknowledging the sovereign's actions as authoritative or having supremacy while feeling aggrieved at them. On this picture, the covenanters surrender their right to all things, expressing acknowledgement of the sovereign as being authoritative or having supremacy, in order to allow the sovereign to govern and establish order. They cannot then complain of injury by the sovereign, since it is *they* who empowered the sovereign and recognised (thereby creating) its authority.⁴³

1.3 Problems and nuances

One seeming problem with the way I have explicated the sovereign's right to enforce is that it may suggest a picture where the sovereign exists in the state of nature and then a group of covenanters come together and agree to obey this sovereign, who retains the right to use violence that it had all along. This is the wrong picture because Hobbes's sovereign is an *artificial* person, not a natural person, and it is *created* by the act of covenant. By covenanting to obey a sovereign the covenanters *create* that sovereign. So the sovereign, in fact, did not exist in the state of nature. It could not have done: what characterises the state of nature is that there is no sovereign. The language of retention may suggest a picture where the sovereign persists from the state of nature, but, so long as it is not understood in this temporal way, there is no problem for my picture. The right of nature (the right to all things) is a right possessed by all persons not living in a commonwealth (that is, under a sovereign) with civil laws forbidding and requiring certain things. For all persons not living in such a state, there are no constraints on permissible treatment of others. There is no reason why this should be true only of natural persons and not of artificial persons. In

⁴³ An alternative possibility is that Hampton is right to point to these non-standard uses of 'authorise' and 'own' but wrong to equate authorisation with surrender of right. In chapter 17 Hobbes says that it is as if the covenanters say 'I authorise *and* give up my right of governing myself ...' (Hobbes (1996) 114 (17:13:87), my emphasis). It could be that the authorisation is just an expression of approval or recognition of authority/supremacy and it is with this that considering oneself to be injured by one's sovereign is inconsistent, while the surrender of right is something separate. This, though, would raise the question why the authorisation, as distinct from the surrender of right, is necessary to create the sovereign, or for the purposes of the covenanters.

covenanting with each other, the covenanters give up their right of nature. In doing so, they create a sovereign who does not covenant with them, who does not live under the power of a sovereign. Thus, this newly created artificial person has the same right to all things that any person has when they do not live under a commonwealth.

Now, it might be objected that the distinction Hobbes makes between punishment and hostility seems to pose a challenge for this interpretation. In chapter 28 ('Of Punishments, and Rewards'), the same chapter in which the passage quoted above occurs where Hobbes bases the right to punish on the right of nature, he clearly distinguishes between punishment and acts of hostility and sets out a number of conditions that an infliction of harm by the sovereign on a subject must fulfil in order to qualify as a punishment and not an act of hostility. For instance, he says that 'the evil inflicted by public authority, without precedent public condemnation, is not to be styled by the name of punishment; but of an hostile act'.⁴⁴ Hobbes defines punishment as 'an evil inflicted by public authority, on him that hath done, or omitted that which is judged by the same authority to be a transgression of the law; to the end that the will of men may thereby the better be disposed to obedience'.⁴⁵ Punishment is done 'by public authority' and so is not possible in the state of nature; all harms inflicted in the state of nature are acts of hostility. Thus, it seems, Hobbes does not think the sovereign *is* in the state of nature with respect to its subjects or that its acts of enforcement just *are* equivalent to acts of hostility performed in the state of nature. There is a subset of harms inflicted by the sovereign that have a different status, and it is these that his account of legitimacy is primarily concerned to justify.⁴⁶ I think this is right, but I do not think it means that we need to reject the interpretation I have given above. It points, though, to a nuance that must be added to that interpretation.

⁴⁴ Hobbes (1996) 206 (28:5:162). Again, in chapter 30, Hobbes says that those who do not understand the right the sovereign has to make and enforce law 'take [punishment] but for an act of hostility; which when they think they have strength enough, they will endeavour by acts of hostility, to avoid' (223 (30:4:176)).

⁴⁵ *Ibid.* 205 (28:1:161)

⁴⁶ In the same passage quoted above in which Hobbes rules out the possibility of *injury* by the sovereign, he also says that 'it is true that he that have sovereign power may commit *iniquity*' (18:6:90, my emphasis). It is thus not the case for Hobbes that the sovereign's actions are just not open to criticism.

The sovereign's right to enforce (and right to punish) are explained simply by its retaining the same right to inflict those harms involved in enforcing or punishing that it has by virtue of not being subject to a sovereign power. However, this is not just because the sovereign *is* in the state of nature with respect to its subjects.

When a commonwealth is instituted, all those involved are put out of the state of nature. The sovereign is thus not created in a state of nature; the moment of its creation coincides with the end of the state of nature (for the society in question). None involved are in a situation characterised by war and insecurity, but rather one characterised by law and stability. This is true for the sovereign just as much as for the subjects. The laws of nature that could not require action in the state of nature come to be binding on action since there now *is* a power that can ensure compliance by other parties. The laws of nature are derived purely from what is in the interests of an individual's self-preservation; they are binding on an individual because she is presumed to desire her own preservation.⁴⁷ Outside of a commonwealth, all but the first two laws of nature fail to be binding on action when one can rationally fear others' non-compliance, since complying with them will not in fact be conducive to self-preservation in such scenarios. However, it is in the interests of all for the laws of nature to be observed, and so for their observance to be enforced (because their observance is necessary for the persistence of a stable commonwealth). This is true for the sovereign as much as for the subjects: it is (just as much as for anyone else) in the interests of the sovereign for a stable commonwealth to persist; 'the good of the sovereign and people, cannot be separated'.⁴⁸

Thus, there *are* certain constraints on what a sovereign may or should do to its subjects and these are placed by the laws of nature.⁴⁹ The sovereign is required by rationality to enforce the laws of nature, since its own self-preservation

⁴⁷ Ibid. 86 (14:3:64)

⁴⁸ Ibid. 230 (30:21:182)

⁴⁹ See also Hobbes's distinction between *good* laws and *just* laws (230 (30:20:181-2)). All laws are just, justice is simply determined by the laws the sovereign makes. However, it is not possible for the civil laws to conflict with the laws of nature; the laws of nature must be a part of the civil laws. Purported laws which conflict with the laws of nature will not be laws (177 (26:8:138)). A sovereign may permissibly make any laws within this constraint, and so whatever they require will be just, though they may fail to be *good*.

requires the commonwealth to persist, which requires the laws of nature to be enforced. (The sovereign is constrained by the laws of nature in a different way to the subjects: the sovereign must *enforce* the laws of nature; the subjects must *obey* them. Additionally, these requirements of right reason are not constraints in quite the same way as the laws enforced by the sovereign. Only the civil laws impose requirements of justice, since the latter depend on sovereign enforcement. The requirements of right reason could be thought of as advice rather than constraints on permissibility, or we could say that there are just two sorts of permissibility).

Much of the second part of *Leviathan* is taken up by claims about what sovereigns ought and ought not to do. I think this can all be understood in this way: there are constraints on what sovereigns should do, including on how they should treat their subjects, and these constraints are placed by the sovereign's need for the preservation of the commonwealth. Hobbes's distinction between punishment and hostility may be explained in these terms: acts of punishment are those that are conducive to the maintenance of a stable commonwealth; acts of hostility put the recipient outside of the protection of the sovereign and in a state of war with regard to it. The sovereign may do this to some of its subjects, but not to all, since that would amount to the dissolution of the commonwealth.

However, it is still the case that the only source of the permissibility of enforcement for the sovereign is the right of nature. When the commonwealth is instituted, certain actions become impermissible for the subjects: those ruled out by the civil laws, the laws made by the sovereign. Though it does not happen in separate temporal stages, something similar happens for the sovereign. The sovereign is created a person subject to no sovereign and so has a right to all things, but because it is not in a state of nature, there are certain limitations on this right to all things (whether they are thought of as constraints on permissibility or simply as constraints on what it is advisable to do is not important). The sovereign, however, is not subject to the civil laws, but only to the laws of nature (and certain other requirements imposed by the need to preserve the commonwealth). Everything that is not ruled out in this way remains permissible (in both senses of 'permissible'). The sovereign and the subjects alike retain intact their right of nature as far as is left open by the laws that apply.

Unlike the subjects, though, the sovereign is *only* constrained by the laws of nature (and the requirement not to endanger the commonwealth). Everything else remains permissible, including the use of violence against others (except in certain special circumstances where its use threatens the commonwealth).⁵⁰ There is thus no special positive argument for the permissibility of enforcement. This is what is meant by the passage quoted above.

The way I have presented the foregoing picture takes the laws of nature to be constraints on permissibility (imposed by the requirement of rationality to do what is necessary for one's self-preservation), that is, limitations on the right of nature. If this is correct, the right of nature is not quite a right to do anything whatsoever, but a right to do anything tempered by certain requirements of right reason. I think this is the right way to understand the laws of nature, though my overall argument does not turn on it. However, the requirements that apply to inhabitants of the state of nature and to sovereigns (if we consider them to be requirements at all) are clearly of a different sort to the requirements of justice that can only bind subjects of a commonwealth through civil laws created by a sovereign.

That the laws of nature are constraints on the right to all things is suggested by the passage quoted above from chapter 28 in which Hobbes says that the right to all things is left to the sovereign entire '*excepting the limits set him by natural law*' (my emphasis).⁵¹ This suggests that there are certain constraints on the sovereign's behaviour imposed by the laws of nature. These are presumably binding on the sovereign in just the same way that the laws of nature are binding on the inhabitants of the state of nature (though, of course, there it is only the first two laws that are binding).⁵² However, these constraints are *not* binding on

⁵⁰ In addition, as Sorell (1991) says, 'the sovereign's obligations *to* the many are no more binding than obligations *between* the many before there is a sovereign to over-awe them' (119). In other words, whatever constraints *do* exist on the sovereign's treatment of its subjects are binding only in the same way in which the laws of nature are binding on the inhabitants of the state of nature.

⁵¹ Hobbes (1996) 206 (28:2:162)

⁵² Sorell (1991) thinks that there can even be *obligations* in the state of nature for Hobbes, only people can rationally hold back from making covenants or following the laws of nature for fear that others will not honour their obligations (113).

the sovereign in the same way as the civil laws (including the laws of nature) are binding on the subjects. The requirement to enforce the laws of nature is binding on the sovereign *simply* because of its necessity for the preservation of the commonwealth. For the subjects, the threat of punishment at least has an important role to play in explaining the bindingness of the civil laws and laws of nature.⁵³ Only the latter sort of requirements are requirements of justice. The sovereign's enforcement, when within the limits imposed by right reason, is permissible in *both* senses. It infringes neither requirements of justice (because the sovereign can be subject to no such requirements) nor requirements of right reason.

One might find this picture implausible because it leaves the right to all things as not *really* a right to *all things*. This element of the picture, though, is superfluous to my primary purpose. You need not accept it to accept my primary contention. All I need to say to explain the distinction between punishment and hostility consistently with my account of legitimacy in Hobbes is that the sovereign is not literally *in the state of nature* with respect to its subjects, but it has the same right to all things (albeit perhaps with certain limitations) that all persons have when not subject to a sovereign power. (If the laws of nature are not constraints on inhabitants of the state of nature, then there are presumably no constraints on permissibility for sovereigns and the classification of certain acts

⁵³ In justifying the secondary laws of nature Hobbes frequently appeals to the fact that behaviour of the sort demanded is required by the first law, the imperative to seek peace. This might be read as the reason that *sovereigns* ought in all circumstances to *enforce* these laws of nature, but often the language used suggests that this fact is *itself* the reason for *following* the law (e.g. he says that 'he that shall oppose himself against [the fifth law] ... is guilty of the war that thereupon is to follow'; Hobbes (1996) 101 (15:17:76)). This appears to suggest that when people are assured (by the sovereign's enforcement) that others will behave alike, the necessity of compliance with the laws of nature for the preservation of the commonwealth (which is in their interests) is *itself* the reason they ought to comply. On this picture, what is permissible for the subjects of a commonwealth is determined by what laws the sovereign enforces - enforcement is necessary for a law to be binding, but the *reason* they ought to obey certain laws (the laws of nature) if enforced by the sovereign is not *simply* that they will be punished if they do not, but also that their compliance is necessary for the maintenance of the commonwealth, which is necessary for their self-preservation. The problem with this is that no individual's compliance *is* strictly necessary for the maintenance of the commonwealth. Hobbes's answer to this is his answer to 'the fool': failure to comply would make you unfit to be accepted into a civil society and so lead you to be returned to a state of war (or simply to be punished by the sovereign). Thus, the primary reason to obey, it seems, must be to avoid punishment, but it may be that there is also a secondary reason to comply with the laws of nature in particular, which is derived from the requirement to seek peace.

of violence as acts of hostility has nothing to do with their permissibility. Still, the sovereign is not in the state of nature, and so acts of hostility are at least *different* from acts of punishment for this reason). It is the fact that enforcement is not *ruled out* by limitations on the sovereign's right to all things (if there are any) that explains the permission to enforce rather than any special positive argument.

2. Hypothetical contract

Now, thus far, we have been considering Hobbes's argument as if it is a historical argument basing the power of states (and perhaps the obligation to obey and the permission to enforce) on an actual historical contract. An alternative is to interpret Hobbes's contract as a *hypothetical* contract that is imagined by Hobbes in order to serve some role in argument about actual states without itself *being* an argument about actual states. We can see how the hypothetical contract quite straightforwardly provides a teleological *justification* of the state of the sort described at the beginning of this essay. We imagine how things would be in the absence of the state and conclude that it is much better with the state: the state is necessary for the fulfilment of some important goal, here the achievement of peace.⁵⁴ However, it is less immediately obvious how a description of a hypothetical contract could serve as an argument for either *obligation* or *legitimacy*. Considering legitimacy, it should be fairly obvious that the form of argument from authorisation discussed above could not hope to work if the contract is taken to be hypothetical rather than historical (and of course I have rejected it as an interpretation of Hobbes even *within* a hypothetical contract). If the contract is taken to be hypothetical, not actual, there is no claim that the actual citizens of actual states have authorised the state in the way that the covenanters in the hypothetical contract scenario have. That, because they have authorised the state and so *own* its actions, hypothetical covenanters cannot be injured by that hypothetical state is quite obviously no argument that actual citizens of actual states cannot be injured by *those* states, since they have *not* authorised their state and so cannot be assumed to own its actions.

⁵⁴ This may show, if successful, at least that *some* state is justified.

We can also see that interpreting the contract as a hypothetical one makes a major difference to the argument for obligation. As has often been noticed, a hypothetical contract argument is a very different sort of animal to an actual contract argument. In the words of Ronald Dworkin, ‘a hypothetical contract is not simply a pale form of an actual contract; it is no contract at all’.⁵⁵ If a hypothetical contract is to be the basis for political obligation, it cannot provide the sort of voluntarist argument that an actual contract could. If the actual contract is of the right sort, those who have contracted may be said to have consented to the authority of the state and so incurred an obligation by an act of their own, that is, voluntarily. Hobbes’s ‘self-assumption’ principle would be met. But its merely being the case that one *would* have consented to the state’s authority under certain hypothetical conditions has no such implication. Wolff suggests that Hobbes’s hypothetical contract is meant to show (and give the reasons) that we, though perhaps unaware of doing so, in fact *do* consent.⁵⁶ As he notes, though, this cannot provide an account of obligation for all citizens, since some may not accept that the absolute state is better than the state of nature.⁵⁷ It is the *fact* of actual consent that here would provide the basis for obligation (though the consent itself is based on hypothetical reasoning about what one *would* consent to in other circumstances). It is not the hypothetical consent that the hypothetical covenanters would give to the state that grounds this obligation.

However, interpreting Hobbes’s contract as a hypothetical rather than a historical one does *not* in fact make any difference to his argument for legitimacy, unlike his argument for obligation. This is because, even if Hobbes does not take actual states to be founded on the basis of state-of-nature contracts, he does not need any special argument for their being permitted to enforce their commands. The sort of obligations that could make certain sorts of action between people *impermissible* only arise once a common power is established over them. States are not subject to any higher sovereign and so, whatever their history, are not in the sort of condition in which any moral requirements apply to them beyond the

⁵⁵ Dworkin (1973) 501

⁵⁶ Wolff (1994) 275-7

⁵⁷ Ibid. 275-6

laws of nature. Thus, apart from the constraints imposed by those laws of nature that become binding with the existence of a sovereign, they retain intact their right to all things and so are permitted to enforce their commands not because of any special fact about them, but simply because there are no constraints that apply to them that could make it *impermissible*.

3. Conclusion

I think, then, that Hobbes certainly thinks that states have a general and exclusive permission to enforce compliance with their commands (so long, presumably, as they provide the peace and security that justifies their existence). However, although he could easily be read as giving a voluntarist argument for this permission (and he does seem to be a voluntarist about political obligation), this would in fact be a mistake. Rather, he has no special argument for the permissibility of state enforcement; it simply results from the sovereign not being in the kind of condition in which enforcement could be *impermissible*. (And the exclusivity of the permission arises simply from the fact that all other members of the society in which the state is sovereign *are* in a condition where constraints on permissible action are set by the sovereign's laws and commands). Thus, the grounds of state legitimacy in Hobbes's view depend on his scepticism about state-independent morality, that is, on there not being any constraints on morally permissible action except those created by a sovereign power. If we do not share Hobbes's moral scepticism, then, we should not look to his arguments to support state legitimacy. There may of course be Hobbes-*inspired* arguments for state legitimacy: we could use a Hobbesian argument for the *justification* of the state as a *premise* in an argument for legitimacy. But this will need some additional work: we will not find such an argument in Hobbes himself.

References

Copp, David (1999) 'The Idea of a Legitimate State' *Philosophy and Public Affairs* 28: 3-45. <http://www.jstor.org/stable/2672824>

Dworkin, Ronald (1973) 'The Original Position' *University of Chicago Law Review* 40: 500-33. <http://www.jstor.org/stable/1599246>

Gauthier, David (1969) *The Logic of Leviathan* Oxford: Oxford University Press

Green, Leslie (2004) 'Law and Obligations' in Jules L. Coleman, Kenneth Einar Himma and Scott J. Shapiro (eds.) *The Oxford Handbook of Jurisprudence and Philosophy of Law* Oxford: Oxford University Press

Hampton, Jean (1986) *Hobbes and the Social Contract Tradition* Cambridge, UK: Cambridge University Press

Hobbes, Thomas (1996) *Leviathan* (ed. J.C.A. Gaskin) Oxford: Oxford University Press

Hume, David (1994) *Political Essays* (ed. Knud Haakonssen) Cambridge, UK: Cambridge University Press

Kavka, Gregory S. (1986) *Hobbesian Moral and Political Theory* Princeton, NJ: Princeton University Press

Raz, Joseph (1979) *The Authority of Law: Essays on Law and Morality* Oxford: Oxford University Press

----- (1988) *The Morality of Freedom* Oxford: Oxford University Press

Schmidtz, David (1990) 'Justifying the State' *Ethics* 101: 89-102.
<http://www.jstor.org/stable/2381893>

Simmons, A. John (2001) *Justification and Legitimacy: Essays on Rights and Obligations* Cambridge, UK: Cambridge University Press

Sorell, Tom (1991) *Hobbes* London: Routledge and Kegan Paul

Tuck, Richard (1996) 'Hobbes's moral philosophy' in Tom Sorell (ed.) *The Cambridge Companion to Hobbes* Cambridge, UK: Cambridge University Press

Wellman, Christopher H. (1996) 'Liberalism, Samaritanism, and Political Legitimacy' *Philosophy and Public Affairs* 25: 211-37.
<http://www.jstor.org/stable/2961925>

Wolff, Jonathan (1994) 'Hobbes and the motivations of social contract theory' *International Journal of Philosophical Studies* 2: 271-86.
<http://www.tandfonline.com/doi/abs/10.1080/09672559408570795>



BASIC LIBERTIES, THE MORAL POWERS AND WORKPLACE DEMOCRACY

LIBERDADES BÁSICAS, PODERES MORAIS E DEMOCRACIA NO LOCAL DE TRABALHO

Stephen K. McLeod*
skmcleod@liv.ac.uk

Abstract. *The article responds to previous work, by Martin O'Neill, about the Rawlsian case for an entitlement to an element of workplace democracy. Of the three arguments for such an entitlement that O'Neill discusses, this article focuses mainly on the one he rejects (on the grounds of its having an implausible premise): the Fundamental Liberties Argument, according to which the right to an element of workplace democracy is a basic liberty. This article argues that while the argument can be improved to withstand O'Neill's objection, it is invalid. The article sets out a new argument, the Argument from Risk to the Moral Powers. While inspired by the Fundamental Liberties Argument, it is valid. Moreover, its premises are at least as plausible as those of one of the two arguments upon which O'Neill builds his Rawlsian case for an entitlement to an element of workplace democracy.*

Keywords. *basic liberties, economic justice, industrial democracy, liberal egalitarianism, moral powers, John Rawls.*

Sumário. *Este artigo responde ao trabalho de Martin O'Neill sobre o argumento rawlsiano a favor do direito a um elemento de democracia no local de trabalho. De entre os três argumentos para este direito que O'Neill discute, este artigo concentra-se principalmente no argumento por ele rejeitado (por causa de ter uma premissa implausível): o Argumento das Liberdades Fundamentais, segundo o qual o direito a um elemento de democracia no local de trabalho é uma liberdade básica. Apesar de o argumento poder ser melhorado para resistir à objeção de O'Neill, este artigo defende que este é inválido e apresenta um novo argumento, o Argumento do Risco para os Poderes Morais. Embora este seja inspirado pelo Argumento das Liberdades Fundamentais, é válido. Além disso, as suas premissas são, pelo menos, tão plausíveis como as de um dos dois argumentos sobre os quais O'Neill constrói o seu caso rawlsiano a favor do direito a um elemento de democracia no local de trabalho.*

Palavras-chave. *democracia industrial, igualitarismo liberal, justiça económica, liberdades básicas, poderes morais, John Rawls.*

* University of Liverpool.

the form of association (...) which if mankind continues to improve, must be expected in the end to predominate, is not that which can exist between a capitalist as chief, and work-people without a voice in management (...) (Mill, 1965, p. 775, quoted by Christie, 1984, p. 112)

The values and ideas central to liberal democracy (...) call for some degree of workplace democracy. (Vela, 2000, p. 3)

Much work remains to be done if we are to realize justice at work. (Hsieh, 2009, p. 409)

o. Introduction

Rawls (2001, p. 178) indicates that “democracy in the workplace”, more specifically, and strongly, in the form of “Mill’s idea of worker-managed firms” is “fully compatible” with his conception of justice.¹ Rawls (2001, p. 179) raises, but does not attempt to answer, the question of whether the “long-run prospects of a just constitutional regime may depend”, in part, upon the establishment of workplace democracy. A positive answer to that question implies more than the mere *compatibility* of workplace democracy with a Rawlsian theory of justice. Such an answer suggests that either Rawlsian principles of justice themselves, or those principles in conjunction with certain social and/or psychological facts, entail that there is a *right* to an element of workplace democracy (and, therefore, that there is a *moral requirement* to institute democracy in the workplace).

Three questions about the normative and logical connections between a broadly Rawlsian theory of justice and workplace democracy are relevant here. First, do Rawlsian considerations *allow for* an element of workplace democracy? This is the question of *consistency*. Second, do these considerations *favour* an element of workplace democracy? This is the question of whether there is, by Rawlsian lights, (*moral*) *reason* to institute democracy in the workplace. Third, do these considerations *require* an element of workplace democracy? That is, does Rawlsian justice entail that there is a *right* (whether basic or non-basic) to

¹ Rawls (2001, p. 114) considers the theory of justice to be neutral, at the level of fundamental rights, with respect to capitalist and socialist modes of production. The specific claim of Rawls (2001, p. 178) is about the compatibility of the worker-managed firm with “property-owning democracy”, which is the form of social order that, assuming capitalism, and in contrast, for example, with the welfare state, Rawls favours. On the differences between these two forms of social order, with references to Rawls and others, see, e.g., Hsieh (2005, pp. 129–30; 2009). Hsieh (2005, p. 115) begins by quoting more fully and commenting upon remarks from Rawls (2001, p. 178) about, in Hsieh’s words, “worker participation in the governance of economic enterprises”. Cf. Hsieh (2008, p. 397).

an element of workplace democracy, or *decisive moral reason* to institute it?² It is easier to establish that there is *moral reason* to institute workplace democracy than it is to establish, and, from a strategic point of view, to convince the opposition, that there is a *right* to an element of democracy in the workplace. Thus, it would be regrettable to concentrate too much –whether here or at large– on the third question. This article assumes, like the work with which it mainly engages, and the passing remarks of Rawls himself on the topic, that the question of consistency is to be answered affirmatively. Over the article’s course, the main focus shifts from the third question onto the second. This shift has the additional strategic and theoretical advantage of allowing for the possibility that democratic rights are only, in the end, appropriate in some workplaces rather than others.³ The powerful will be more likely to concede that there is moral reason to share some of their power than that their power belongs, largely, and by right, to others. The argument that there is such moral reason, short of an inalienable right, favours gradualism and reform over revolution. It is likely to have greater practical influence, and unifying force, than the appeal to a putatively inalienable right.⁴

Perhaps you are already thinking: to what does workplace democracy amount and what forms might it take?⁵ It is not the purpose of this article to specify any putatively Rawlsian answers to these questions. Rather, it settles for working with the notion of a “measure”, “degree” or “element” of workplace democracy: it is concerned with whether there should be some democracy in the workplace rather than none. It is one thing to answer that question affirmatively and another to work out the details, both normative and practical, of measures

² The distinction made here between consistency, favouring and entitling (as a matter of right) builds on Vega (2000, p. 35) and Hsieh (2008, pp. 72–73). A Rawlsian account of justice recognizes that consequentialist considerations can qualify as morally decisive reasons when acting on them violates no rights. To cut down on clutter, in what follows the word “right” will be used as shorthand, except when the context excludes this, for anything to which all citizens are, as a matter of Rawlsian justice, entitled for decisive moral reasons. Vela’s term “call for”, quoted in the epigram above, accommodates both entitlement claims and favouring claims.

³ Thus, it is consistent with the view of Christie (1984, p. 116), according to which not all small businesses should be required to recognize such rights.

⁴ Norton (2003) has influenced these thoughts. In relation to debates about environmental conservation, Norton (2003, p. 469) distinguishes between the “intellectual debate” about what “the *correct moral stance towards* nature is” (including on the question of whether non-human species and ecosystems have inherent value) and the “strategic debate” about which arguments are “likely to be *effective* in saving wild species and natural ecosystems”. My point is that even if the correct moral stance on workplace democracy should turn out to be more demanding than the stance I defend (e.g., in that the former demands inalienable democratic rights for all employees), the stance I defend seems more likely to be strategically effective in advancing workplace democracy as a progressive cause and practical project.

⁵ You might also be asking: what counts as a workplace? This question must here be left aside. On a separate issue, Vela (2000, pp. 144–145) argues that when it comes to democratizing production, employees should be at the front of the queue. I agree, but the issue has no bearing on the arguments below.

for instituting democracy in the workplace. Nevertheless, some brief observations may be in order to help forestall some over-estimations of what “workplace democracy” might mean here. In this connection, consider some shifts in the usage of “workplace democracy” over the course of a series of articles by Nien-hê Hsieh. The defence of an element of workplace democracy is a more modest project than defence of that which Hsieh (2005, pp. 115–116) describes as follows (while not setting out to defend it):

Workplace democracy, as I understand it, is an economic regime in which workers have a right to participate in the governance of economic enterprises on [the] grounds that they have a right to exercise control over the means of production.

The presence of an entitlement to some degree or other of workplace democracy entails neither workers’ (full or partial) ownership of, nor their (full or partial) control over, the means of production.⁶ On this matter, the conception of workplace democracy with which this article will work converges with the “workplace republicanism” defended by Hsieh. According to it, writes Hsieh (2005, p. 137) “what matters is that the decision-making procedure [in the workplace] incorporates the voice of workers as part of the process”. The view defended here adds that the incorporation of voice should be democratic, rather than, for example, done merely via such mechanisms as consultation, focus groups, and one-to-one discussions between employer and employee.

In a subsequent article, Hsieh (2008, p. 82) works with a less demanding definition under which

a guarantee of the worker’s right to participate in governance is independent of any requirement of ownership on the part of workers. In addition, (...) no constraints [are placed] on whether worker participation is direct or representative. Furthermore (...) workers and providers of capital [might] share control of the enterprise as in a system of co-determination. For a regime to count as one of workplace democracy, workers must be guaranteed at a minimum a say equal to that of the providers of capital.⁷

⁶ Workplace democracy, as understood by O’Neill (2008) and throughout this article, is consistent with the decoupling of democracy in the workplace and workers’ ownership of the means of production. It is neither ownership over nor control of the means of production that is crucial. Rather, it is workers’ democratic participation in the making of decisions that affect their own terms and conditions of employment and their conduct, and that of their bosses, in the workplace and the marketplace.

⁷ Hsieh (2009, p. 399) adopts an even less demanding definition, on which workplace democracy is “worker participation in organizational decision making”. Since democratic voice is not entailed by such participation –as the cases of consultation and the use of focus groups show– this remark seems to have been a slip-up.

On this definition, workplace democracy has three elements: (i) the workers have, as a matter of right, voice in respect of governance;⁸ (ii) their voice is expressed by democratic means; (iii) their voice is equal to that of the owners of the means of production. This article is concerned only with the first two elements. The question is not whether Rawlsian considerations require, or favour, the possession by workers of equal voice, democratically expressed, to that of owners or bosses. Rather, it is whether Rawlsian considerations, require, or favour, that workers should have some democratically expressed voice (rather than none). This is a relatively modest question. The proposal that Rawlsian justice favours, or requires, the institutionalization of some measure or other of workplace democracy is also relatively modest. Therefore, the proposal is apt for philosophical evaluation in advance of any more demanding proposal that entails it.

Martin O'Neill (2008) has set out three Rawlsian arguments, two of which he endorses and the third of which we will shortly examine in detail, for the conclusion that a measure of workplace democracy is a requirement of Rawlsian justice (and therefore morally required). The relative modesty of O'Neill's goal, shared here, enables his arguments to avoid what Hsieh (2005, pp. 116–17) calls “the objection most often [levelled] against workplace democracy, which is that it pays insufficient attention to the need for managerial decision-making in large-scale economic enterprises”.⁹

O'Neill (2008, p. 31) calls these arguments the “Fundamental Liberties Argument”, the “Democratic Equality Argument” and the “Democratic Character Argument”.¹⁰ These arguments differ mainly in the terms of the sub-conclusion

⁸ A putative right that Hsieh (2005, pp. 117, 134–40), discusses is “the right to contest managerial decisions”. This involves a form of voice, but need not be democratic. To see that voice in the workplace does not entail democracy in the workplace note, for example, that individuals could have the right, as individuals, to contest managerial decisions that directly affect them without having any organized role in determining the organization's explicit and documented values, strategy, tactics and policies. Moreover, their right to express their views on such decisions might permissibly, at least in respect of a right to voice that is so weak, elicit nothing more than the managers' classic response: “We hear what you're saying, but we'll be doing it anyway”.

⁹ Christie (1984, p. 115) mentions the more general objection that calls for workplace democracy are, in practical terms, impossible to implement. Since there already are enterprises in which workers have some institutionalized democratic rights (and even economies in which a measure of democratic involvement by workers in decision-making is legally required), this more general objection is also here circumvented.

¹⁰ I have capitalized “Argument”. Throughout, O'Neill uses both the expression “economic democracy” and the expression “workplace democracy”. The former might be thought to encompass not only “democratization of economic enterprises” (O'Neill, 2008, pp. 30–31), at the level of the firm or agency of production, but also democratic political control, at a state or federal level, over production. Christie (1984, p. 113) regards economic democracy as something that involves “a serious attempt (...) to democratize the economic sphere, including workplaces”. This suggests that workplace democracy is a form of, but not identical with, economic democracy. It is the implications of a broadly Rawlsian account of justice for workplace democracy, rather than for other forms or aspects of economic democracy, that is of interest here. For references to some other work, dating from 1978 to 2000, that argues, on

meant to support the conclusion that justice requires a degree of workplace democracy.

The *Fundamental Liberties Argument* moves from the sub-conclusion that a degree of workplace democracy is a Rawlsian “basic liberty” to the conclusion that justice requires it.¹¹ The *Democratic Equality Argument* moves from the sub-conclusion that implementation of the difference principle, when understood as a principle that encompasses all social primary goods, requires that “an economic system” should allocate “decision-making powers in a broadly dispersed way” to the conclusion that a measure of workplace democracy is a requirement of justice (O’Neill, 2008, p. 31).¹² According to the *Democratic Character Argument*, the institutional recognition of a right to a measure of workplace democracy is appropriate because “a society which institutes forms of economic democracy will be more likely to preserve a stable and just basic structure over time, by virtue of the effects of economic democratization on the development of an active, democratic character among citizens” (O’Neill, 2008, p. 30).¹³ While O’Neill (2008, pp. 31, 42–48) supports the Democratic Character Argument and the Democratic Equality Argument, he ultimately rejects the Fundamental Liberties Argument. This article aims to show, first, that we can circumvent O’Neill’s objection to the Fundamental Liberties Argument. Secondly, that, even so, the argument fails to establish the conclusion that workers have, on Rawlsian grounds, an inalienable right to a measure of workplace democracy. Thirdly, that the argument obscures the distinction between two types of requirements of justice: deontological requirements (which follow from the principles of justice alone) and practical requirements (which follow from the principles of justice

Rawlsian grounds, for an entitlement to a measure of workplace democracy, see Hsieh (2005, p. 115, note 4).

¹¹ Previously, Clark and Gintis (1978, pp. 303, 311–313) gave this argument, in its essentials. Unlike O’Neill, they are unreserved proponents of the argument. A right to an element of workplace democracy, if it were a basic liberty, would presumably be a particular instance of a more general basic liberty. Thus, the argument appears to be an example of what Mayer (2000, p. 305) calls “[d]erivative arguments [which] deduce the right to workplace democracy from a prior and more general right”.

¹² Given its reliance on the difference principle, this seems to be an example of what Mayer (2000, p. 306) calls “[d]istributive arguments for the right to workplace democracy [which] do not derive that right from another but appeal instead to a norm of distributive justice”. In contrast with the other two arguments covered by O’Neill, this article neither reconstructs nor discusses further the Democratic Equality Argument.

¹³ The Democratic Character Argument is anticipated, in its essentials, by Clark and Gintis (1978, p. 312): “the extension of democratic principles to the production process would have the effect of strengthening the system of total liberties, since the experience of equal participation in decision-making strengthens individual commitment to principles of justice”. In the terminology of Mayer (2000), this argument appears to be neither derivative nor distributive. This suggests, if the argument is intended to establish an inalienable right to workplace democracy, that Mayer’s distinction, while exclusive, is not exhaustive. Perhaps the argument is not intended to establish an inalienable right to workplace democracy at all. That would be a charitable interpretation, for the argument is sound only if its conclusion is interpreted, as is done in Section 5 below, as *not* entailing that there is such a right. Christie (1984, p. 117) and Cohen (1989) also give arguments that, while not specifically Rawlsian, anticipate the Democratic Character Argument.

only alongside certain facts about actual social conditions that are independent of the principles of justice). Fourthly, that there is nevertheless a strong case for saying that considerations of Rawlsian justice favour (and even, in light of social circumstances, morally require) the institutionalization of an entitlement (albeit alienable) to an element of workplace democracy for employees.

Section 1 explains the background to the Fundamental Liberties Argument and provides a detailed reconstruction of the argument.

Section 2 discusses the most significant Rawlsian objection, previously endorsed by O'Neill, to the content of the Fundamental Liberties Argument.

Section 3 shows how to circumvent this objection via a weakening of some of the content of the argument. More specifically, given that O'Neill's objection works only if the definition of the basic liberties that it uses is tenable, the Fundamental Liberties Argument can be reformulated, in a manner that is invulnerable to O'Neill's objection, by replacing O'Neill's definition of the basic liberties with a tenable sufficient condition upon a liberty's being basic. (It is not necessary to specify necessary and sufficient conditions for a liberty's being basic.) Section 3 then provides a reformulation of the argument that includes this weakening.

Section 4 explains why, rather than being an *ad hoc* manoeuvre designed to save the argument, this weakening is independently justified.

Section 5 raises an objection to the Reformulated Fundamental Liberties Argument that also affects the original version of the argument. Unlike O'Neill's objection, this objection is to the argument's validity. On the one hand, there are *deontological requirements* of Rawlsian justice. Among rights and liberties, these include only those rights and liberties possessed by citizens of sufficiently advanced societies in all relevantly similar morally permissible worlds in which the circumstances of justice obtain. On the other hand, there are *practical requirements* of Rawlsian justice. These include rights and liberties that are not so possessed. The article argues that the (Reformulated) Fundamental Liberties Argument obscures this distinction. The argument embodies either a deontological fallacy or, if you prefer, a fallacy of equivocation. When this flaw is resolved, a valid argument inspired by the (Reformulated) Fundamental Liberties

Argument, namely the Argument from Risk to the Moral Powers, is recoverable. This argument is on a par with the Democratic Character Argument. Both arguments suggest that Rawlsian considerations about justice favour a measure of workplace democracy: given certain actual social conditions (about which parties in the original position are ignorant) its presence would help enable a just social order to be reached and, once established, to be stable (given the continuation of some of those social conditions) over time. There is no suggestion that the practical requirements of justice cannot thereby be moral (and thus deontic) requirements. They are nevertheless distinct from, and more derivative than, those deontological requirements that are either: (i) inalienable rights of citizens (such as the right not to be discriminated against under the law on the grounds of race or sex); or, (ii) otherwise (e.g., via the difference principle) directly entailed by the principles of justice (independently of facts about social conditions that differ across the societies, actual or possible, to which the principles of justice apply).

Section 6 is a concluding summary.

1. The Fundamental Liberties Argument

Before setting out the argument, let us clarify some of its background. The conception of justice as fairness regards citizens as persons engaged in social co-operation who have “what we may call ‘the two moral powers’”: “the capacity for a sense of justice” and “a capacity for a conception of the good” (Rawls, 2001, pp. 18–19; O’Neill, 2008, p. 34). The principles of justice concern the design of the basic structure of society. This basic structure

is the way in which the main political and social institutions of a society fit together into one system of social cooperation, and the way they assign basic rights and duties and regulate the division of advantages that arise from social co-operation over time (...). (Rawls, 2001, p. 10)

The fundamental case in which the capacity for a sense of justice is exercised is in “the application of the principles of justice to the basic structure and its social policies” (Rawls, 2001, p. 112; O’Neill, 2008, p. 34). The fundamental case in which the capacity for a conception of the good is exercised is in “forming, revising, and rationally pursuing such a conception over a complete life” (Rawls,

2001, p. 113; O'Neill, 2008, p. 34). With these preliminaries in place, we can reconstruct the Fundamental Liberties Argument as follows.

Fundamental Liberties Argument

1. A liberty is basic (and thus its institutionalization is a requirement of justice) if and only if it is necessary to the provision of “the social conditions essential for the adequate development and the full and informed exercise of [people’s] two moral powers (...) in the two fundamental cases” (Rawls, 2001, p. 112; O’Neill, 2008, pp. 35–36).¹⁴ (Premise)

2. The “freedom to take part in decisions about economic production” is necessary to the provision of these conditions (O’Neill, 2008, p. 35). (Premise)

3. This freedom requires a degree of workplace democracy. (Premise)

4. This freedom is a basic liberty (and thus its institutionalization is a requirement of justice). (From 1, 2)

5. Justice requires a degree of democracy in the workplace. (From 3, 4)

¹⁴ Though O’Neill seems to interpret what Rawls writes here as, in effect, providing a definition of the basic liberties, this is not the role that the quotation actually plays in Rawls’s own work. Rather, it concerns the question of how the basic liberties should be weighted, in a scheme of liberties, once a list of them is *already* in place. See further McLeod and Tanyi (2017).

According to the argument, as O'Neill (2008, pp. 31, 33) states it, "a right to participate in the democratic determination of decision-making within the social and economic institutions to which citizens belong" is one of the fundamental liberties that must be constitutionally protected in order to enable citizens to exercise the capacity for a sense of justice and the capacity to have a conception of the good. As a consequence, the institutionalization of a degree of workplace democracy is a requirement of justice.

Premise 1 is, at least as it features in the argument, definitional and it goes unquestioned in O'Neill's discussion. However, as we shall see, there are good and independent reasons for taking the definition embodied in Premise 1 to be in need of weakening. Premises 2 and 3 are debatable. In O'Neill's formulation of the argument, Premise 3 is hidden. It may be, in effect, that O'Neill does not spell it out as a separate premise because he takes it already to be entailed by, or included in, Premise 2. However, it is helpful to distinguish between Premises 2 and 3 because the freedom to take part in decisions about economic production could arguably be embodied in a form of state socialism under which there is no element of workplace democracy within enterprises or industries. Workplace democracy involves each worker in a firm having democratic rights within that firm. While this is a form of economic democracy, defined as the freedom to participate in decision-making about production, it is not exhaustive of it. For example, in a form of state socialism citizens who do not work could, via the political process, have such decision-making powers.

In setting out the arguments that underlie the non-definitional premises of the argument, O'Neill (2008, pp. 35–36) appeals, as follows, to the fundamental case in which the capacity for a sense of justice is exercised, namely in making judgements about the justice of the basic structure and its social policies:

(...) with regard to the sense of justice (...) unless individuals have some first-hand experience in the deliberative direction of some collective enterprise (such as a firm), then they will lack the skills that will be needed in order to participate fully in the "free use of public reason" in democratic politics. (...) citizens who are used to the regular exercise of their deliberative capacities with regard to the direction of economic enterprises will be best placed to exercise those capacities in a *full and informed* manner in the political sphere.

2. O'Neill's objection to the Fundamental Liberties Argument

Before we turn to the objection that leads O'Neill to reject the argument, it is worth noting, in the interests of completeness, some of his other critical observations about the argument. O'Neill (2008, pp. 37–41) notes that Rawls himself would have rejected the Fundamental Liberties Argument, though for a reason that O'Neill finds indecisive. This reason can be summarized as follows. Rodney Peffer proposed an amendment to Rawls's principles of justice. Under the amendment, a right to workplace democracy, while not a basic liberty, is a requirement of justice (Peffer, 1990, p. 14; O'Neill, 2008, p. 37). According to Peffer's first principle, persons are entitled to have their "basic security and subsistence rights" met (Peffer 1990, p. 14). According to his second principle, "[t]here is to be a maximum system of basic liberties" (Peffer, 1990, p. 14). Peffer's third principle is:

(3) There is to be (a) equal opportunity to attain social positions and offices, and (b) an equal right to participate in all social decision-making processes within the institutions of which one is part. (Peffer, 1990, p. 14)

Rawls (1993, pp. 7–8, note 7) rejected (3)(b) on the grounds that it appeared to him "to require a socialistic form of economic organization" that should not, in his view, have been a requirement of "the first principles of justice" (that is to say, those principles that are chosen in the original position under the veil of ignorance). Thus, although Peffer's proposal does not entail that a right to workplace democracy is a basic liberty, Rawls's general reason for rejecting Peffer's proposal also applies to any proposal that entails that such a right is a basic liberty (for here it would be included in the very first principle, on Rawls's account, of justice).

For the following reasons, O'Neill does not find this Rawlsian line of the objection to the Fundamental Liberties argument convincing. First, it might be maintained that principles of justice need not be neutral about modes of socio-economic organization and that, accordingly, the theory of justice may quite aptly have a socialistic, rather than a liberal, orientation from the start. As O'Neill (2008, p. 39; cf. Mayer, 2000, p. 304) puts the point, "we should not reject an otherwise normatively attractive account of social justice simply because it can only be realized under some particular system of ownership". Secondly, Rawls, in his response to Peffer, mistakenly bundled up ownership rights over the means

of production with the right to workplace democracy. Therefore, the Rawlsian line of objection to the Fundamental Liberties Argument also suffers from this flaw. While ownership rights are often, *de facto*, bundled up with rights of control, there is no essential moral connection between them.¹⁵ In relation to economic production,

full *ownership* of the means of production need not be a necessary condition for the exercise of some *control* over how production is to take place. For example, one could allow that productive capital could be owned by particular private individuals, whilst nevertheless allowing that the *use* of that productive capital could legitimately be constrained by a requirement that workers (...) were entitled to some degree of participation in decision-making about production. (O'Neill, 2008, p. 39; cf. Vela, 2000, pp. 156–158)

For this reason, provided that by “socialism” Rawls intended a system involving workers’ ownership of the means of production, Rawls’s objection to Peffer’s proposal, which O’Neill applies, appropriately, to the Fundamental Liberties Argument, is indecisive against that argument.

Nevertheless, Premise 2 of the Fundamental Liberties Argument is implausible. Even if it might *enable* the development and full and informed exercise of the moral powers in the two fundamental cases, the freedom to take part in decisions about economic production is not *essential* to this. As O’Neill (2008, p. 41) puts the point, albeit (given that he does not clearly distinguish between Premises 2 and 3) in relation to the specific case of workplace democracy, “there are many other venues (such as in civil society and in private associations) within which people can develop and exercise their two moral powers in the two ‘fundamental cases’”. For this reason, O’Neill (2008, p. 53) does not ultimately rest his Rawlsian case for workplace democracy, even in part, upon the Fundamental Liberties Argument.

3. Resolving O’Neill’s objection to the Fundamental Liberties Argument

Even if Premise 2 is false, it would be a natural piece of advocacy on behalf of the argument to look for a way of weakening it whilst still securing the argument’s conclusion. One way of doing so would be to replace Premise 2 itself

¹⁵ On this point, see also Mayer (2001, p. 231).

with material about mitigating against contingent risk (above a certain threshold) to the moral powers and their exercise in the two fundamental cases. Given the elliptical nature of Premise 2, its weakening can be brought about by weakening the wording of Premise 1. The reformulation of Premise 1 can be accomplished, as Section 4 argues, in a manner that is independently justified rather than *ad hoc*. Moreover, valid inference from Premises 1 and 2 to line 4 of the argument only requires that Premise 1 should state a sufficient condition for a liberty's being basic: a full definition of the basic liberties, stating their necessary and sufficient conditions, is not required. As we shall also see in Section 4, this change is also independently justified: even if the necessity of an entitlement to the provision of "the social conditions essential for the adequate development and the full and informed exercise of [people's] two moral powers (...) in the two fundamental cases" is a sufficient condition upon that entitlement's being a basic liberty, it is not a necessary one. A resultant, weakened, version of the Fundamental Liberties argument can be formulated as follows.

Reformulated Fundamental Liberties Argument

A liberty is basic (and thus its institutionalization is a (Premise)
1. requirement of justice) if it is necessary to the provision of social conditions under which the risk, partly due to social conditions, of stunting or atrophying the adequate development and the full and informed exercise of at least one of the moral powers, in at least one of the two fundamental cases, falls below a reasonable threshold.

The "freedom to take part in decisions about economic (Premise)
2. production" is necessary to the provision of these conditions.

This freedom requires a degree of workplace democracy. (Premise)
3.

This freedom is a basic liberty (and thus a requirement of (From 1, 2)
4. justice).

Justice requires a degree of democracy in the workplace. (From 3, 4)
5.

There are some capacities that are unlikely to atrophy or be stunted just because the person who has them cannot exercise them at work (rather than because, say, of the nature of the work or the length of the working day or week). For example, if one is an office worker working normal hours within reasonable commuting distance of home, then one's job does not interfere with one's capacity to be an excellent cook. Likewise, if I am unable to learn a foreign language at or on the way to work, I might still be able to do it in my leisure time. The workplace, *per se*, and working life in general, pose no threat to capacities and excellences like these.

In the original version of the Fundamental Liberties Argument, the suggestion was that, in a society in which there are firms, workplace democracy is necessary to the provision of "the social conditions essential for the adequate development and the full and informed exercise of the two moral powers". O'Neill's objection was that the moral powers can be developed and exercised, in a manner that fosters their full and informed exercise in relation to the justice of the basic structure and its social policies, outside of the context of making economic decisions about production: for example, in voluntary associations. The objection suggests that the two moral powers are capacities, like being an excellent cook or a good cyclist, that people can adequately gain and maintain outside the workplace. No arrangements in the workplace are therefore *necessary* to their development or their exercise. This analogy is implicit in the objection to the Fundamental Liberties Argument. The moral powers, however, are relevantly *unlike* other some other capacities, in the following respects. First, the moral powers are intimately related to habits of thought, attitude and action that colour our relations and everyday interactions with our fellow persons: some of these habits can only be sustained if practised, more or less consistently, throughout

one's social dealings with others.¹⁶ Those who occupy working situations in which unjust practices are encouraged, or worse, necessary to career success, are to some (perhaps significant) degree *likely* to have their sense of justice corrupted in a manner that risks spillover into areas of life outside the workplace and which might prejudice their capacity to exercise the sense of justice in a full and informed manner in the fundamental case. To put the point another way, thinking, feeling and acting justly are habits that can be lost not only if one does not have the opportunity to exercise them, but if one is regularly put into a situation in which one is under pressure to participate in, or to remain silent about, actions that are contrary to the requirements of justice. Secondly, and relatedly, the hierarchical nature of the typical workplace means that it is a theatre of special threats to the two moral powers. Thirdly, the hierarchical workplace, as a coercive environment, is one in which people are not always free to act with impunity in accordance with their consciences and in which conscientious objections to particular workplace demands are unlikely to be institutionally tolerated or accommodated (particularly when not based on grounds protected by anti-discriminatory legislation). Fourthly, for working people in full-time jobs, work is the predominant mode of social co-operation, outside of the family, in which they are engaged. Many working people with family lives have little or no time, especially if they commute to work, for engagement in the forms of voluntary association that could help them develop and exercise the two moral powers. (They could of course do so during their holidays, but the theory of justice should not have, and cannot coherently have, the result that the cultivation and maintenance of the moral powers ends up being, for some people, a luxury reliant upon their own supererogatory activities.)

It is presumably not enough, if a just basic structure is to be stable, that some people should be able regularly to engage in activities that provide the opportunity to develop and engage the two moral powers while others do not. On Rawls's account, maintaining a stable and just basic structure requires that citizens should *all* be able to develop and engage the two moral powers. Thus,

¹⁶ Compare here Swartz (1982, pp. 636–638), quoting, among others, Adam Smith and Émile Durkheim, though Swartz's discussion is about the relationship between the capacity for autonomy and the habit of exercising it.

despite O'Neill's objection to his formulation of the argument, there seems, at this stage, to be plenty of life left in the argument yet.

4. Weakening Premise 4 of the Fundamental Liberties Argument is not *ad hoc*

In order to show that the above weakening of Premise 1 is not *ad hoc*, this section offers two arguments.

The first begins with the theoretically conservative presumption that it is a condition on the extensional adequacy of a definition of the Rawlsian basic liberties that it should include all the rights and liberties that Rawls included in his lists of the basic liberties. A consequence of this presumption is that the definition used in Premise 1 must, independently of its bearing on the Fundamental Liberties Argument, be abandoned. This point can be illustrated by appeal to the case of freedom of movement. For Rawls (1993, p. 335), “the liberty and integrity of the person” is “violated (...) by denial of freedom of movement”; thus, freedom of movement is a basic liberty. Samuel Arnold (2017, §4.2; cf. Pogge, 2007, p. 87) argues that if a law were enacted that restricted people's freedom of movement to within their metropolitan areas, this would not make it impossible that every citizen should possess, and exercise in a full and informed way, the moral powers in the two fundamental cases. Even if it is very unlikely that, under such conditions, every citizen could be so fortunate as to be able to do this, it does not seem to be impossible. A way to remedy the definition so that it includes freedom of movement as a basic liberty is to weaken it precisely along the lines suggested above: i.e., so that it appeals, instead of necessity, to probability.¹⁷ Given that it is not *ad hoc*, the putative critic of Premise 1 of the Reformulated Fundamental Liberties Argument must show that the sufficient condition on a liberty's being basic that it employs is too permissive.

¹⁷ For a full and sustained defence of this move, see McLeod and Tanyi (2017).

The second argument begins by noting that Rawls jumbles up modality and probability when he writes that:

a liberty is more or less significant depending on whether it is more or less essentially involved in, or is a more or less necessary institutional means to protect, the full and informed exercise of the moral powers in one (or both) of the two fundamental cases. (Rawls, 2001, p. 113)

Necessity, which is a modal notion, does not admit of degrees. Neither does essentiality. For any given liberty, either it is essentially involved in, or a necessary institutional means to protect, the full and informed exercise of the moral powers in the two fundamental cases or it is not. In other words, this is a binary property of the liberty, not a scalar property. The relevant notion that *does* admit of degrees is probability. To retain Rawls's distinction between those liberties that are essential to citizens considered as free and equal and those that relate to the institutional means for protecting the moral powers in the relevant way, while at the same time respecting the analytical point that necessity does not admit of degrees, something like the following amendment to Rawls's wording would be required:

a liberty meets the threshold of significance for being basic if and only if either (i) it is essentially involved in the full and informed exercise of the moral powers in one (or both) of the two fundamental cases (as it is an essential feature of persons, considered as free and equal, that they have this liberty), or (ii) its *absence* is, other things being equal, and for some people, *more likely than not* to jeopardize the full and informed exercise of the moral powers in one (or both) of the two fundamental cases.

5. The Fundamental Liberties Argument is invalid

O'Neill's objection to the Fundamental Liberties Argument is a content objection that the Reformulated Fundamental Liberties Argument, in view of its weaker and more plausible content, avoids. Unfortunately, however, both versions of the argument fall to a formal objection.

The inference from Premises 1 and 2 to the sub-conclusion at line 4 of the (Reformulated) Fundamental Liberties Argument is valid. The other inference, from lines 3 and 4 to the conclusion, embodies a fallacy. If justice requires that *P* and *P*, in turn, requires that *Q* it does not necessarily follow that *justice requires that Q*. To take an example from retributive justice, suppose that justice requires that some crimes should be punishable by imprisonment. Imprisoning criminals

requires that there should be jailors. Now it does not follow that justice requires that there should be jailors: for jailors are required only if criminal acts are actually committed that are serious enough for imprisonment to be an appropriate retributive measure. That there should be jailors is not strictly a deontological requirement of justice: rather, it is a *practical requirement* that follows from a requirement of justice *under certain contingent conditions*. There is no direct analogy between this example and the case of distributive justice, for it is the circumstances of justice that are the relevant contingent conditions in the case of distributive justice.¹⁸ Nevertheless, the example shows that the form of inference employed in deriving the conclusion of the (Reformulated) Fundamental Liberties Argument from lines 3 and 4 is invalid. Moreover, the argument obscures the distinction between requirements of justice that are *deontological* and those that (though they can be moral) are merely *practical* and it can be seen, thereby, as embodying a fallacy of equivocation. On the assumption that an entitlement is a core basic liberty only if it is a deontological requirement of justice that it be extended to all citizens in every society in which the circumstances of justice obtain, the (Reformulated) Fundamental Liberties Argument does not secure its intended conclusion.

The moral powers are not requirements of justice in the sense, for example, that equality before the law, or the distribution of social primary goods according to the difference principle, are supposed to be requirements of justice. If a certain person never happens adequately to develop them, or adequately to be able to exercise them, then this is not necessarily a matter of injustice, for it need not be due to any social disadvantage. Rather, the moral powers are *prerequisites* for justice. A just constitution can only be drawn-up, and can only be sustained, thanks to the moral powers. Their *continuance* is necessary to the *stability* of that constitution. Given that we are not yet close to living in what Rawls would consider to be a well-ordered society, we require that the moral powers should, if necessary, be developed, and should, certainly, be exercised to take us closer. So, full and informed exercise of the moral powers, while necessary to the *continuation*, and thus to the *stability*, of a well-ordered society is also necessary to the *formation* of such a society. On Rawls's account, persons in a well-ordered

¹⁸ On the circumstances of justice, see Rawls (1971, §22).

society should be able to formulate for themselves the principles of justice; the ability to do so rests upon, or is a special case of, the ability to exercise the moral powers in a full and informed way in the two fundamental cases.

Rawls uses the expression “basic liberties” to encompass two forms of entitlement: inalienable freedoms (such as freedom of thought) and inalienable rights (such as the right to equality before the law). An entitlement is inalienable only if it is: (i) due to all persons, considered as free and equal citizens, in societies in which the circumstances of justice obtain; (ii) not apt for trade-offs, made either by its owner or by anyone else, against other goods. We can see that a right to a measure of workplace democracy is not a basic right if we consider a society in which the circumstances of justice obtain but in which the economic circumstances are somewhat different to those of the societies with which we are familiar. For example, we can image a society in which the working week is significantly shorter than is the case for the typical full-time worker in today’s advanced industrial or post-industrial societies. In such a society, a lack of democratic rights in the workplace might not meet the threshold of posing a significant risk of jeopardy to the moral powers or their full and informed exercise in the two fundamental cases. Suppose, alternatively, that legislation granting democratic rights in the workplace was introduced in a current society, but that it came with a proviso allowing employees to trade-in those rights in return for a radical reduction in the duration of the working week, with no corresponding reduction in their remuneration. Such a proviso would, by the lights of the Reformulated Fundamental Liberties Argument, have to be regarded as breaching no requirement of justice provided that the trade-in also effected a sufficient reduction in jeopardy, taking the risk below the threshold, to the full and informed exercise of the moral powers in the two fundamental cases.¹⁹

¹⁹ Cf. Vega (2000, pp. 31–40, 43–44). According to Mayer (2001, p. 301), “the decisive issue upon which the debate” about workplace democracy “ought to turn”, which has “largely neglected by both advocates and critics” of the right to workplace democracy, is that of whether the putative right is alienable. Even if it were valid, the Reformulated Fundamental Liberties Argument could not secure an inalienable right to workplace democracy as long as factors other than an entitlement to workplace democracy might mitigate against the risk posed, by the undemocratic workplace, to the moral powers and their exercise in the two fundamental cases. Mayer (2001, p. 302) argues that “employee claims to a share of power in the workplace (...) are not inherently inalienable”. Mayer (2001, p. 301) is concerned with a more substantial right than O’Neill, namely the right to an equal say in the running of the firm on the part of all workers, but many of the points he makes are of more general relevance, including to arguments intended to establish only an element of workplace democracy (rather than an equal say for workers).

The moral powers and their full and informed exercise in the two fundamental cases are essential aspects of the conception of citizens as free and equal that is integral to Rawls's theory of justice. It is a deontological requirement of justice that no-one who is naturally capable of having and exercising these powers in these ways should, through morally contingent social conditions, be impeded from doing so. It is useful to distinguish between two sorts of social impediments. On the one hand, there are impediments that are *intrinsic* violations of justice: these are impediments that, independently of any other morally contingent social or psychological facts, violate the extension of citizenship to all persons naturally capable of it, or which violate the conception of citizens as free and equal. For example, sexist or racist restrictions on voting rights are intrinsic violations of justice: for racism and sexism are intrinsically anti-egalitarian affronts to the dignity of persons. There is no permissible world in which there is a state in which suffrage is restricted on the grounds of race or sex. On the other hand, there are impediments the removal of which is a prerequisite for justice not because they are intrinsic violations of justice but because of morally contingent social or psychological facts. On a Rawlsian account of justice, the absence of a right to democracy in the workplace is not an intrinsic violation of justice, because (on the assumption that it is indeed a fact) it is a matter of moral contingency that the undemocratic workplace poses a special threat to the moral powers and their exercise in the two fundamental cases.²⁰ If humans were immune to having the moral powers and their exercise in the two fundamental cases impeded by undemocratic workplaces, then there would be no need for a right to a measure of workplace democracy. The distinction that is being made here is similar to the deontologist's distinction between the *intrinsic* features of an act (e.g., that it is an act of promising or of lying) and its *extrinsic* features (e.g., that it secures, but is not constitutive of, a business contract). The effects of an act are among its extrinsic features. So far as arguments for an entitlement to a measure of workplace democracy that appeal to the moral powers are concerned, the existence of permissible worlds

²⁰ Also, forms of social co-operation that are undemocratic need not, merely in virtue of this fact, be unjust. For example, if the members of an amateur sports team defer to their coach, and give the coach the lion's share of decision-making rights in respect of the team's activities off the field, then this does not, at least on a Rawlsian account of justice, entail that they have attempted to hand to the coach some of their inalienable rights. I owe the example to Iñigo González Ricoy. Social co-operation within economic production need not always be an exception to this. For example, the apprentice may defer, and give over decision-making rights concerning production, to the accomplished practitioner for whom the apprentice works, as sole employee. On a Rawlsian account of justice, this does not entail that the apprentice has thereby attempted to forsake an inalienable right.

containing states in which there is no right to a measure of workplace democracy is not precluded: for there are permissible worlds in which persons are not vulnerable to the special threats the workplace poses, in the actual world, to the possession and full and informed exercise of the moral powers in the two fundamental cases. By the lights of Rawlsian arguments that appeal to the moral powers, and in the actual world, it is because of its consequences, not any intrinsic injustice, that the lack of a right to workplace democracy is morally dubious (if not impermissible).

Thus, while the original version of the Fundamental Liberties Argument aimed to establish that a right to a measure of workplace democracy is a basic liberty, neither it nor the Reformulated Fundamental Liberties Argument succeeds in doing this. Nevertheless, if an argument inspired by the (Reformulated) Fundamental Liberties Argument does establish that a measure of workplace democracy is, other prevailing social conditions being equal, a practical (perhaps even moral, even if not fundamental) requirement of justice then this remains an interesting consolation prize.

If it is a *practical requirement*, in order for a just and stable social order *to come about*, that possession of the two moral powers and the effective ability to exercise them in a full and informed manner, should be prevalent in society, or (more weakly) if it is *practically unlikely* that such a social order will come about without this, then this is a good (though not necessarily decisive) moral reason to organize society in ways that will bring this about. Our practical starting position is not the original position, and in order for us to approximate to what Rawls would consider a well-ordered society, the ability for full and informed exercise of the moral powers in the two fundamental cases must be protected and promoted. Existing institutional arrangements, which commonly operate in ways that stifle the moral powers, must be reformed so that the risk of such stifling falls below a certain threshold of significance.²¹ Factional interests, for example, hinder and stifle the full development and exercise of the ability to be impartial. Hence, they hinder the sense of justice. Partly because of conflicting factional

²¹ Even if it is not true, contrary to Schwartz (1982, p. 639), that “persons’ autonomous developments are stunted when their jobs severely restrict their opportunities for rationally framing, pursuing, and adjusting their own plans”, it might still be the case that there is a tendency towards, or a significant risk of, such stunting. While it is, at least partly, a matter for empirical investigation, it is plausible that in the hierarchical workplace there is a tendency towards, or a significant risk of, the stunting of the moral powers.

interests, the workplace is frequently a place of strife. Within contemporary societies, the development of workplace democracy, whilst taking some decision-making power away from senior executives and shareholders in order to redistribute it to employees, would be likely to result in a decrease in factional interests in production. These interests have traditionally been at the fulcrum of political life, including to the point of civil disorder, in industrialized societies. They remain relevant, even if less pronounced, in post-industrial, service-led economies. It is a reasonable conjecture, open to empirical investigation by social scientists, that a levelling out of control in decision-making in the workplace correlates with a decrease in factional tensions in both the economic sphere and the more narrowly political sphere, with a corresponding increase in the extent to which people are able to view policy issues in an impartial, rather than factional, manner.

We have seen that while the Reformulated Fundamental Liberties Argument, with its appeal to probability, is not vulnerable to O'Neill's objection to the Fundamental Liberties Argument, neither argument is in fact sound. While the original argument contains a false premise, the reformulated version is, alongside its predecessor, invalid. Nevertheless, the foregoing observations enable the construction of a new argument, as follows, that is an adaptation, designed to support a weaker conclusion, of the Reformulated Fundamental Liberties Argument.

Argument from Risk to the Moral Powers

- If ensuring the satisfaction of a certain condition is (Premise)
1. necessary to the provision of social conditions under which the risk (partly due to social conditions) of stunting or atrophying the adequate development and the full and informed exercise of at least one of the moral powers (in at least one of the two fundamental cases), falls below a reasonable threshold, then we have moral reason to ensure the satisfaction of that condition.

- Under prevailing social circumstances, ensuring an (Premise)
2. entitlement to “freedom to take part in decisions about economic production” is necessary to the provision of the conditions under which that risk is kept below the threshold.

- Under prevailing social circumstances (including, for (Premise)
3. example, the length of the typical full-time working week and the size and structure of the average firm or agency of production), this freedom requires an entitlement to a degree of workplace democracy.

- We have moral reason, under prevailing social (From 1, 2)
4. circumstances, to ensure an entitlement to “freedom to take part in decisions about economic production”.

- We have moral reason, under prevailing social (From 3, 4)
5. circumstances, to ensure a degree of workplace democracy.

The Argument from Risk to the Moral Powers can be seen as claiming that, other things being equal, the full and informed exercise of the moral powers in the two fundamental cases is put into too much jeopardy in the absence of economic-democratic entitlements, including the entitlement to a degree of workplace democracy. Unlike each version of the Fundamental Liberties Argument, the argument is valid. Critical discussion of its premises would be likely to focus on the issue of where the line is to be drawn regarding what constitutes a reasonable threshold of risk to the moral powers and their exercise in the two fundamental cases and on the plausibility of Premises 2 and 3. The Argument represents a significant advance on the Fundamental Liberties Argument, for, in addition to being valid, its controversial premises are logically weaker and more plausible than those of each version of that argument. Notably, the premises are even more plausible than those of the Reformulated

Fundamental Liberties Argument, which was an argument specifically designed to avoid O'Neill's content objection to the Fundamental Liberties Argument.

There is a parallel between the Argument from Risk to the Moral Powers and the Democratic Character Argument.²²

Democratic Character Argument

1. An entitlement to democratic participation in workplace decision-making is conducive towards and/or supportive of the democratic characters of citizens in that it promotes the full development and/or the exercise of the moral powers (in the two fundamental cases).²³ (Premise)
2. That which is conducive towards and/or supportive of the full development and/or the exercise of the moral powers (in the two fundamental cases) thereby enables citizens to play their full roles as citizens. (Premise)
3. When a just society is one in which citizens can play their full roles as citizens, that society is likely (other things being equal) to be more stable over time than a just society that is less developed in this respect. (Premise)
4. If an entitlement has positive effects on the stability, over time, of a just society, then there is moral reason to secure that entitlement (as a matter of public policy). (Hidden Premise)

²² The reconstruction paraphrases O'Neill (2008, p. 42), makes the content and structure of the argument more explicit and weakens its conclusion.

²³ O'Neill couches his premises in terms of democratic participation itself, rather than the entitlement to it. Since the argument's conclusion is about the entitlement to it, in the reconstruction the premises explicitly mention the entitlement (so as to secure the argument's validity).

5. An entitlement to democratic participation in workplace decision-making enables citizens to play their full roles as citizens. (From 1, 2)

6. A just society in which there is an entitlement to democratic participation in workplace decision-making is, thereby, likely (other things being equal) to be more stable, over time, than a just society that is less developed with respect to the ability of citizens to play their full roles as citizens. (From 3, 5)

7. There is moral reason to secure an entitlement to democratic participation in workplace decision-making (as a matter of public policy).²⁴ (From 4, 6)

The Argument from Risk to the Moral Powers and the Democratic Character argument are two sides of the same coin. While the Democratic Character argument emphasizes the benefits, to moral character, of an entitlement to workplace democracy, the Argument from Risk to the Moral Powers emphasizes the risk of harm to moral character in the absence of such an entitlement. The first emphasizes a positive correlation between two factors and the second a negative correlation between those same factors.

Both arguments establish, at best, that a measure of workplace democracy is a practical requirement of Rawlsian justice. That is, given certain actual social conditions (about which parties in the original position are ignorant), it is

²⁴ At the point at which he presents the substance of the argument, O'Neill (2008, p. 42) does not actually formulate its conclusion. Instead, the conclusion, that an element of workplace democracy is "a requirement of justice", is stated in his abstract (O'Neill, 2008, p. 30), though the element of workplace democracy appears under the moniker "economic democratization". The conclusion in the above reconstruction is weaker so as to render the argument valid and to protect it from the ambiguity of "requirement of justice" that was discussed earlier in this section.

required if a just social order is to be reached and, once established, to be stable over time (assuming the continuation of some of those social conditions). In so far as these arguments support an entitlement to workplace democracy it is only one that is apt to be secured at the legislative stage in the four-stage sequence.²⁵ Thus, this entitlement is not on a par with such core basic liberties as freedom of thought and equality under the law. Nevertheless, Rawlsian considerations of justice, along with certain seemingly plausible empirical hypotheses about the interplay between institutional arrangements in the workplace and the full and informed exercise of the moral powers in the two fundamental cases, provide a strong case for *favouring* an element of workplace democracy over none. Moreover, Rawls's theory of justice allows, at the legislative stage of the four-stage sequence, and at subsequent stages, for entitlements that are merely practical requirements of justice to serve the functional role, within a just social order, of being rights apt for protection under the law. While such rights are of lesser moral and political significance than are the core basic liberties, their significance may still be profound.

6. Conclusion

The arguments considered in this article do not show that a degree of workplace democracy is a deontological requirement of justice in a well-ordered society. However, there seems to be a strong Rawlsian case for the introduction of a degree of workplace democracy into current societies: this is plausibly a practical, and even moral, requirement if a just social order is to come about and be stable over time. Under prevailing social conditions, and alongside some seemingly plausible empirical hypotheses, considerations of Rawlsian justice certainly *favour* the institutionalization of a measure of workplace democracy and might well *require* it too, albeit not as a matter of inalienable right.

²⁵ In the case of the Democratic Character Argument, O'Neill (2008, pp. 31–32) acknowledges this. The four-stage sequence consists of (in this order) the original position, the constitutional stage, the legislative stage and the judicial stage. For details and discussion, see Rawls (1971, §31; 2001, §13), Wenar (2017, §4.9) and McLeod and Tanyi (2017).

Acknowledgements

I thank the organizers of, and members of audiences at, the following conferences, at which some of this material was presented: “Workplace Democracy: Arguments, Policies, Practices” (Lille Catholic University, 2016); VIII Braga Meeting on Ethics and Political Philosophy (University of Minho, 2017); the Society for Applied Philosophy Annual Conference 2017 (University of Copenhagen). For comments on earlier drafts, I thank Ian Dunbar, Andrew Holland, Gregory Miller, Thomas Schramme and Rachael Wiseman. I am especially grateful to Ana Reimão, whom I mistook for her neighbour, for helping a stranger with his Portuguese, to Martin O’Neill, for the work that inspired me to write this article, and to Attila Tanyi, for working with me on, and helping me to think about, the question of which liberties are basic.

References

Arnold, S. (2017). Putting liberty in its place: high liberalism without the liberalism. *European Journal of Philosophy*. Advance online publication. doi: 10.1111/ejop.12215

Christie, D. (1984). Recent calls for economic democracy. *Ethics*, 95(1), 112–128. Retrieved from <http://www.jstor.org/stable/2380553>

Clark, B., & Gintis, H. (1978). Rawlsian justice and economic systems. *Philosophy and Public Affairs*, 7(4), 302–325. Retrieved from <http://www.jstor.org/stable/2264960>

Cohen, J. (1989). The economic basis of deliberative democracy. *Social Philosophy & Policy*, 6(2), 25–128. doi:10.1017/S0265052500000625

Hsieh, N.-h. (2005). Rawlsian justice and workplace republicanism. *Social Theory and Practice*, 31(1), 115–42. Retrieved from <http://www.jstor.org/stable/23558690>

Hsieh, N.-h. (2008). Justice in production. *Journal of Political Philosophy*, 16(1), 72–100. doi: 10.1111/j.1467-9760.2007.00290.x

Hsieh, N.-h. (2009). Justice at work: arguing for property-owing democracy. *Journal of Social Philosophy*, 40(3), 397–411. doi:10.1111/j.1467-9833.2009.01459.x

Mayer, R. (2000). Is there a moral right to workplace democracy? *Social Theory and Practice*, 26(2), 301–325. Retrieved from <http://www.jstor.org/stable/23559088>

Mayer, R. (2001). Robert Dahl and the right to workplace democracy. *Review of Politics*, 63(2), 221–247. Retrieved from <http://www.jstor.org/stable/1408665>

McLeod, S. K., & Tanyi, A. (2017). What are basic liberties? Manuscript submitted for publication.

Mill, J. S. (1965). *Collected works*, Vol. 3, J.M. Robson (Ed.). Toronto, Canada: University of Toronto Press.

Norton, B. G. (2003). The cultural approach to conservation biology. In *Searching for sustainability: interdisciplinary essays in the philosophy of conservation biology* (pp. 467–477). Cambridge, England: Cambridge University Press. (Originally published in 1989, in D. Western & M. C. Pearl (Eds.), *Conservation for the twenty-first century*. New York: Oxford University Press.)

O'Neill, M. (2008). Three Rawlsian routes towards economic democracy. *Revue de philosophie économique*, 8(2), 29–55. Retrieved from

<http://www-users.york.ac.uk/~mpon500/TRRTEconomicDemocracy.pdf>, 23rd January 2018.

Peffer, R. (1990). *Marxism, morality, and social justice*. Princeton, NJ: Princeton University Press.

Pogge, T. W. M. (2007). *John Rawls: his life and theory of justice*. Oxford, England: Oxford University Press.

Rawls, J. (1971). *A theory of justice*. Cambridge, MA: Harvard University Press.

Rawls, J. (1993). *Political liberalism*. New York: Columbia University Press.

Rawls, J. (2001). *Justice as fairness: a restatement*. Cambridge, MA: Harvard University Press.

Schwartz, A. (1982). Meaningful work. *Ethics*, 92(4), 634–646. doi:10.1086/292380

Vela, R. G. (2000). *A theory of economic democracy* (Doctoral dissertation). Retrieved from <http://hdl.handle.net/1721.1/8958>

Wenar, L. (2017). John Rawls. In E. N. Zalta (Ed.), *Stanford encyclopedia of philosophy* (Spring 2017 ed.). Retrieved from <https://plato.stanford.edu/archives/spr2017/entries/rawls/>



FROM NATURAL LAW TO THE GOLDEN RULE: AQUINAS REVISITED

DA LEI NATURAL À REGRA DE OURO: S. TOMÁS REVISITADO

Damiano Simoncelli*
damiano.simoncelli@gmail.com

Abstract. *These days, the Thomistic account of natural law is the object of renewed interest and criticisms. A number of objections are usually lodged against the idea of a human nature and a shared human good, in that it might seem that these ideas are unquestionably culturally related and that cultural boundaries cannot be crossed. At the same time, the concepts of 'human nature' and 'natural law' are often misunderstood to be related to human biology only. To overcome these issues, this paper aims to reinterpret the Thomistic doctrine of natural law as a form of the golden rule ('Do not do unto others as you would not have them do unto you'; 'Do unto others as you would have them do unto you').*

Keywords: *natural law, Aquinas's ethics, golden rule, intercultural ethics.*

Sumário. *Atualmente, a concepção tomista da lei natural é objecto de interesse renovado e de novas críticas. As ideias de natureza humana e de um bem humano partilhado são alvo de várias críticas, na medida em que pode parecer que essas ideias são, definitivamente, relativas à cultura e que os limites da cultura local não podem ser ultrapassados. Ao mesmo tempo, os conceitos de 'natureza humana' e de 'lei natural' são frequentemente mal compreendidos, como se estivessem apenas ligados à biologia humana. Para ultrapassar estas questões, esta comunicação visa reinterpretar a doutrina tomista da lei natural como uma forma da regra de ouro ("Não façam aos outros o que não queres que te façam a ti").*

Palavras-chave: *lei natural, etica tomista, regra de ouro, etica intercultural.*

*PhD Student, University of Genoa-FINO Consortium.

o. From the Golden Rule to Natural Law¹

Enquiring into natural law and the golden rule ('Do not do unto others as you would not have them do unto you'; 'Do unto others as you would have them do unto you') means investigating the primary sources of human action. The golden rule, which is part of the heritage of most cultures (Wattles 1996; Vigna 2005; Neusner & Chilton 2008; Gensler 2013; Puka n.d.) and encompasses the structure of an ethics of mutual recognition, shares with natural law a common history, at least within the Christian tradition. We can find the origins of this strong connection in the Patristic custom of a synoptic reading of two biblical settings (du Roy 2008, 2012). The first is St. Paul's letter to the Romans, at 2:14–15, in which the apostle holds that the Gentiles are a law unto themselves in that they do by nature what is ordered by the law because they have the law written in their hearts.² The second is Matthew 7:12, where the positive version of the golden rule is said to be the Law and Prophets.³ This interpretation goes back even to Justine the Apologist (100–165) and to Origen (185–245) and is accepted by the following fathers of the Church (Sciuto 2005; du Roy 2008), including the authorities Basil the Great and Augustine of Hippo.⁴ Thus, it becomes a common reference point for the thinkers that followed, from the medieval to the modern scholastics.⁵

Given these historical connections, in this paper I will revalue them theoretically by reinterpreting the Thomistic account of natural law as a form of the golden rule. To do this, my reference points will be the thought of the Italian ethicist Carmelo Vigna and that of Alasdair MacIntyre. The aims of this theoretical move are (a) to look for the idea of 'common human nature' without

¹ I am grateful to Prof Alasdair MacIntyre, Prof Carmelo Vigna, Prof Angelo Campodonico, Prof Harm Goris, Prof Adrian J. Reimers, Dr Alessandro Biasini and Dr Maria Silvia Vaccarezza for their helpful comments. I also thank the audience at VIII Braga Meetings on Ethics and Political Philosophy and the anonymous reviewer.

² 'For when the Gentiles who do not have the law by nature observe the prescriptions of the law, they are a law for themselves even though they do not have the law. They show that the demands of the law are written in their hearts, while their conscience also bears witness and their conflicting thoughts accuse or even defend them' (Rom. 2:14–15 New American Bible Version).

³ 'Do to others whatever you would have them do to you. This is the law and the prophets' (Mt 7:12).

⁴ For a systematic perspective on the golden rule in Augustine, see Catapano 2005.

⁵ The case of the Spanish jurist Alfonso de Castro (1495–1558), who even rooted the golden rule in the proper structure of the natural law, seems noteworthy: according to him, the first precept of the natural law is the positive version of the golden rule (1568, book I, chap. I).

neglecting the differences in which it is embodied, and (b) to make room for the contemporary attention to intersubjectivity in the context of a discourse on natural law. First, I will sketch a brief anthropological outline. Second, my enquiry will test the potential for relating the golden rule to the resources of Thomistic natural law. Third, I will focus on objections to this reinterpretation. Finally, I will highlight the strengths of my proposal.

1. Anthropological Premises

The Thomistic doctrine of natural law should be understood within its specific metaphysical context. Indeed, considering natural law means considering the human being's peculiar way of being subjected to the divine government of the entire creation. This way is rooted in their metaphysical constitution itself: unlike inanimate things, plants, and animals, human beings are able to head towards their end by themselves, thanks to the dynamism of their freedom.

As just seen, the human being is part of the divine government in a specific way. The human being's specificity depends on its exceptional metaphysical status. It is important to understand the etymology of the adjective 'exceptional', which comes from the Latin verb *excipere*: 'taking (*capere*) out (*ex*)'. Accordingly, by its very constitution, the human being is 'the out-taken'. If the human being is the out-taken, then where is it taken out of? The answer may appear paradoxical: it is taken out of the self-closure of every nonrational creature, being at the same time everywhere and nowhere (ST, I 75.5; Pagani 2012, 2014). Far from the nonsense this expression may appear to be, it intends to highlight that the human being is entirely related to an infinite horizon—that is, the horizon of being, and of truth and good, as such (QDV, 1.1).

At this stage, we should note two things. First, here we can find the grounds of human freedom: indeed, being open to an infinite horizon entails that none of the finite realities that can be found within this horizon are bound to compel any human choice. The gap between this opening and the finite status of everything one can find within one's experience allows one to not be entirely captured by every finite reality and to be able to focus on a different choice (ST, I–II 10.2).

The only thing to which the human being is bound is that infinite horizon; therefore, it implies that the human being is looking for an infinite object, even when its eyes lay on a finite one (ST, I–II 10.1; QDV 25.1). Second, it is important to stress that the human being is entirely related to this horizon. This relation does not only involve humans' reason and will, but somehow all their faculties (ST, I 76 and 77.7). Thus, it appears impossible to divide, within the human being, a pure animality from a pure rationality: everything within it is marked by its openness to all being (DE II; De Finance 1993; Vigna 2015, 2016).

This infinite horizon attracts the human being. If we consider the relation in such a way, we can call it 'desire'. Nonetheless, it is a common experience that human desire cannot find a totally fulfilling object within its historical boundaries: such an object must be the plenitude of being in all its aspects. In Thomistic terms, this object must be 'what everyone calls God'. But there is an object in which the human being can find a trace of their appropriate one: another human being, one who is somehow infinite, if their own infinite openness is considered (Vigna 2015). Moreover, the other human being not only is one's historical end, but, in a certain way, one's principle: in fact, the other's sight, when it is a true sight, is what enables one to be aware of one's own 'ex-ceptionality'—that is, transcendental (Vigna 2015; Zanardo 2017).⁶

To summarise: (a) the human being is structurally and totally related to the infinite; (b) when this relation occurs as an appetite, it can be called 'desire'; (c) historically, human desire cannot find a totally fulfilling object; (d) nonetheless, the human being, within intersubjective relationships, can find a partner who makes them aware of their ontological stature and can partially fulfil their desire.

2. From Natural Law to the Golden Rule

Revaluing the historical connection that I have previously brought to light in a theoretical perspective could present many advantages concerning, for instance, the possibility of setting intersubjective and cross-cultural grounds for

⁶ The Hegelian teaching on intersubjectivity given in chapters IV and VI of his *Phenomenology of Spirit* appears to be an essential reference point to develop the Thomistic anthropological account (Hegel 2000).

doing ethics. However, one might ask why we need to deal with common human nature and with appealing to the ancient category of natural law when we could consider individual traditions with their specific sets of values and differences. My answer is threefold. First, if we want to consider the traditions in their uniqueness, we should also consider the premises of every social reality⁷—among them, the existence of others, given that people want to be in social relationships with them. Second, if we want to explain any cross-cultural relationship,⁸ we should admit to a common grammar of every cultural narrative: if two—or more—cultures were completely enclosed worlds, each exclusively governed by their specific logic, any encounter would be inexplicable since any ‘common ground’ (in the case of a cooperative relationship) or any ‘common bone of contention’ (in the case of a conflictual relationship) would be lacking.⁹ Third, if every tradition is to experience cross-cultural encounters as a tool of enhancement, it must be true that every different culture expresses the same common nature.¹⁰

Therefore, my starting point will be the same as Aquinas’s: ‘Bonum est faciendum et prosequendum, et malum vitandum’ (ST, I–II 94.2). That is, ‘Good is to be done and pursued and evil is to be avoided’. Indeed, according to Thomas Aquinas, this principle encompasses the entire natural law. Every other specific tenet is nothing more than an expression of this fundamental precept, a precept

⁷ It is a fundamental point highlighted by MacIntyre: ‘If we are to achieve an understanding of good in relation to ourselves as being, as animal, and as rational we shall have to engage with other members of the community in which our learning has to go on in such a way as to be teachable learners. And thus we accomplish the first realization of our good by the most elementary way respecting the good of those others in encounter with whom we have to learn. What we grasp initially in understanding the binding force of the precepts of the natural law are the conditions for entering a community’ (1990, 136–37).

⁸ It is important to stress that I speak of any relationship, regardless of its quality; indeed, both agreement and disagreement presuppose a common reference point (MacIntyre 1988, 1990, 2009).

⁹ MacIntyre writes: ‘Aquinas’ account of the precepts of the natural law, far from being inconsistent with the facts of moral disagreement, provides the best starting point for the explanation of these facts’ (2009, 26); ‘When confronted by some immediate disagreement as to what you or I or we should do here and now, reason requires us to ask who is in the right, and the argument then proceeded by our further noting that, if we are to enquire effectively who is in the right, we must do so in the company of others and more especially of those others with whom we are in disagreement.... It is a condition of the rationality of shared enquiry that the social relationships of those engaged in it should be structured by certain norms, norms that find expression in the primary precepts of the natural law’ (2009, 24–25; see also 2000).

¹⁰ From this perspective, the existence of a common human nature seems to be the theoretical ground of the MacIntyrean dialectic among different traditions, in which every tradition can compare its resources with those of a different one. If a specific tradition were not able to recognise in a diverse one a different way of enquiring into the same aims, the comparison of their resources would be nonsensical (MacIntyre 1988, 1989, 2006).

that reveals the structural human tendency towards good in general.¹¹ As already seen, we can also call this tendency ‘desire’, which, conceived as a tendency towards good *in general*, has an unlimited capacity. Being the appetite for a boundless object, desire brings about the possibility of encompassing every action, practice, or project of life, for they remain, in any case, in the realm of the finite being.

It might seem this conception of desire can be easily accepted these days because of its structural indefiniteness. On the contrary, it appears problematic to posit that this boundlessly open desire expresses itself through a number of more definite tendencies that are universally shared by all human beings as such. Therefore, nowadays there is a need to focus on which particular needs are truly universal.¹² Furthermore, even if we agreed on a number of universal needs, the question of how to live them well would remain open. Finally, the current awareness of the cultural dimension requires that we pay more attention to the role that the concrete ethos plays in knowing and appreciating the human goods.

The Thomistic doctrine of natural law allows room for the above three issues. Aquinas holds that (a) there are a number of human ontologically rooted needs¹³ (according to Aquinas’s jargon, *inclinationes naturales*; ST, I–II 94.2, co), (b) there is a just—that is, rational—manner of living them in the context of

¹¹ ‘All these precepts of the law of nature have the character of one natural law, inasmuch as they flow from one first precept’ (ST, I–II 94.2 ad 1, trans. Fathers of the English Dominican Province).

¹² Here ‘universal’ means ‘transcendental’—that is, ‘what is common beyond the differences’.

¹³ ‘I answer that, As Boethius says (De Duabus Nat.) and the Philosopher also (Metaph. v, 4) the word “nature” is used in a manifold sense. For sometimes it stands for the intrinsic principle in movable things. In this sense nature is either matter or the material form, as stated in Phys. ii, 1. In another sense nature stands for any substance, or even for any being. And in this sense, that is said to be natural to a thing which befits it in respect of its substance. And this is that which of itself is in a thing. Now all things that do not of themselves belong to the thing in which they are, are reduced to something which belongs of itself to that thing, as to their principle. Wherefore, taking nature in this sense, it is necessary that the principle of whatever belongs to a thing, be a natural principle. This is evident in regard to the intellect: for the principles of intellectual knowledge are naturally known. In like manner the principle of voluntary movements must be something naturally willed. Now this is good in general, to which the will tends naturally, as does each power to its object; and again it is the last end, which stands in the same relation to things appetible, as the first principles of demonstrations to things intelligible: and, speaking generally, it is all those things which belong to the willer according to his nature. For it is not only things pertaining to the will that the will desires, but also that which pertains to each power, and to the entire man. Wherefore man wills naturally not only the object of the will, but also other things that are appropriate to the other powers; such as the knowledge of truth, which befits the intellect; and to be and to live and other like things which regard the natural well-being; all of which are included in the object of the will, as so many particular goods’ (ST, I–II 10.1 co).

the moral life,¹⁴ and (c) there is a direct proportionality between the knowledge of natural law and the context of practices (be they individual or social) in which it is comprehended. Indeed, Aquinas holds that, on the one hand, virtuous practices deepen people's knowledge and appreciation of human good,¹⁵ and, on the other, vicious practices (*ex mala consuetudine*) are bound to involve a number of difficulties in knowing and appreciating what is really good.¹⁶

However, many difficulties arise in any attempt to articulate and single out specific human needs. In this respect, the golden rule, considered from an intersubjective philosophical perspective, could provide us with valuable ways of rearticulating this Thomistic doctrine while offering solutions to its problems.

In its primary function, the golden rule can help us to discover the structural needs of the human being, untangling them from different customs and individual tastes. Our starting point can be the positive formulation of the golden

14 'All the inclinations of any parts whatsoever of human nature, e.g. of the concupiscible and irascible parts, in so far as they are ruled by reason, belong to the natural law, and are reduced to one first precept, as stated above: so that the precepts of the natural law are many in themselves, but are based on one common foundation' (ST, I-II 94.2 ad 2).

15 'I answer that, We may speak of virtuous acts in two ways: first, under the aspect of virtuous; secondly, as such and such acts considered in their proper species. If then we speak of acts of virtue, considered as virtuous, thus all virtuous acts belong to the natural law. For it has been stated that to the natural law belongs everything to which a man is inclined according to his nature. Now each thing is inclined naturally to an operation that is suitable to it according to its form: thus fire is inclined to give heat. Wherefore, since the rational soul is the proper form of man, there is in every man a natural inclination to act according to reason: and this is to act according to virtue. Consequently, considered thus, all acts of virtue are prescribed by the natural law: since each one's reason naturally dictates to him to act virtuously. But if we speak of virtuous acts, considered in themselves, i.e. in their proper species, thus not all virtuous acts are prescribed by the natural law: for many things are done virtuously, to which nature does not incline at first; but which, through the inquiry of reason, have been found by men to be conducive to well-living' (ST, I-II 94.3 co; see also Campodonico 2013).

16 'We must say that the natural law, as to general principles, is the same for all, both as to rectitude and as to knowledge. But as to certain matters of detail, which are conclusions, as it were, of those general principles, it is the same for all in the majority of cases, both as to rectitude and as to knowledge; and yet in some few cases it may fail, both as to rectitude, by reason of certain obstacles (just as natures subject to generation and corruption fail in some few cases on account of some obstacle), and as to knowledge, since in some the reason is perverted by passion, or evil habit, or an evil disposition of nature; thus formerly, theft, although it is expressly contrary to the natural law, was not considered wrong among the Germans, as Julius Caesar relates (*De Bello Gall. vi*)' (ST, I-II 94.4 co). 'I answer that, As stated above, there belong to the natural law, first, certain most general precepts, that are known to all; and secondly, certain secondary and more detailed precepts, which are, as it were, conclusions following closely from first principles. As to those general principles, the natural law, in the abstract, can nowise be blotted out from men's hearts. But it is blotted out in the case of a particular action, in so far as reason is hindered from applying the general principle to a particular point of practice, on account of concupiscence or some other passion, as stated above. But as to the other, i.e. the secondary precepts, the natural law can be blotted out from the human heart, either by evil persuasions, just as in speculative matters errors occur in respect of necessary conclusions; or by vicious customs and corrupt habits, as among some men, theft, and even unnatural vices, as the Apostle states (*Romans 1*), were not esteemed sinful' (ST, I-II 94.6 co).

rule:¹⁷ ‘Do unto others as you would have them do unto you’. At first sight, the golden rule might appear an empty tenet. However, if we proceed to invert its maxim, we will have an effective heuristic tool: ‘As you would have them do unto you, so do you unto others’. If we take into account the inverted maxim in its first part, the golden rule basically orders one to put oneself in the other’s place and to look at oneself from the perspective of one who recognises us (Vigna 2005, 2008, 2015). Although this discovering process lives within concrete relationships, we can attempt to outline how it works.

Why do we need to invert the maxim? The answer is based on the anthropological account outlined before: far from being only a set of categorial needs, the human being, qua human being, has a transcendental—that is, boundlessly open—desire that looks for a fulfilling object. This object—within our actual experience—can only be another human being (Vigna 2015). This fact reveals the need for intersubjective relationships to fulfil human desire. However, not just any relationship can satisfy human desire, but only a relationship of mutual recognition, since it is necessary that the other be allowed to maintain their stature.¹⁸ If one is considered a mere object, one will be reduced to a completely finite being and bound to disappoint the other’s infinite desire. Given this ontologically grounded starting point, we can consider what one who recognises us has to do unto us. First of all, one must allow our *existence* in our *dignity*—the former since the existence of a partner is a necessary condition for any relationship, and the latter since one has to grant oneself a partner in their stature as a transcendental subjectivity. Furthermore, to have a balanced relationship, one must accept a mutuality. In this way, the other must recognise our relationality and safeguard our equality. What is more, such a relationship entails wishing to deepen the true good of the partner in order to be oriented towards reciprocal flourishing. Hence the knowledge of one’s own true good is considered vital. In the end, a balanced relationship calls for the possibility of

¹⁷ Indeed, the negative one (‘Do not do unto others as you would not have them do unto you’) depends on this positive formulation, whose is the safeguarding side (Vigna 2015).

¹⁸ Vigna writes: ‘The mutual-recognition relationship would seem to be *the only practical intersubjective relation in which two (or more) subjectivities are able to coexist in all the magnificence of their universality/transcendentality*. Each subjectivity needs to be recognized as an unsurpassable horizon of meaning, that is as intentionally unconditioned (and this is so on account of the universality/transcendentality present in it)’ (2008, 217).

cultivating a religious dimension, understood in the broadest sense. As we have already seen, human desire has unlimited width and looks for a fulfilling object—that is, within our actual experience, another human being. Nonetheless, the other human being is not entirely unlimited, for they are also situated within specific boundaries (first and foremost, within the boundaries of their mortal life). Thus, the possibility of an enquiry into the existence of a subjectivity that really is unlimited regarding all aspects must be left open.

By this deepening of the relationship of mutual recognition, we have been able to bring to light the same fundamental spheres of needs as those Aquinas sketches out in several places in his works—existence, life, knowledge, and so on¹⁹ (see, e.g., ST, I-II 10.1; QDV, 22.5; QDM 16.4 ad 5; ST, I-II 94.2).

The golden rule provides us not only with a tool of discovery, but with a regulative one. Having highlighted the basic objects of human desire, we should ask ourselves how they ought to be sought. For instance, we saw that life is one of these fundamental needs. We should take care of it by staying healthy, by avoiding dangers, by being concerned for our own safety, and so on. However, this inclination ought to be lived not as if it were the only human good, but within the overall architectonic structure of human needs. The standpoint of the other who recognises us is able to help us to reach an ordered desire. The otherness of the other's perspective guarantees less involvement with our own specific situation. Therefore, the perspective can provide an impartial view of what we need. At the same time, this other perspective does not turn out to be extraneous and cold, for it is coloured by an attitude of care.²⁰

¹⁹ The goods concerning the human relationships are, within this perspective, at the same time a class among the human goods and the context in which all the human needs can be discovered.

²⁰ In this way Vigna stresses the regulative role of the (friendly) mediation of one's own desire: 'Since we are usually aware that passions affect judgement, we ask a friend for a counsel, because a friend not directly involved in the passions that affect us in that given moment is in a better position to cast an "objective" eye on the matter, or he is simply in a better position anyway.... [B]ut there is also a second reason to be considered. If the object of my desire is not for the sake of me, but of others, I will surely be inclined to avoid any kind of excess in using my strengths, since the strengths are mine whereas the object is for others, but not for me. In short, a certain wise economy in the energy I put into my efforts is definitely to be expected.... A third reason is that, when we do something for the other's sake, we generally do so in response to a more or less explicit request. But this request always comes with, and looks like, a burden, a limitation on our freedom of movement; we quickly perceive the excess contained in the requests coming from the others, and we feel deprived of our freedom to decide correctly on the appropriate response. We do want to do something for the others, but at the same time we make an effort to understand what they *really* need, precisely so as not to live that constriction as violence brought to bear upon ourselves. This is why, when acting for the benefit of others, we naturally tend to follow a reasonable

Finally, as partially said before, the golden rule reminds us of the essential role of involvement in a concrete relationship in order to know human needs. First and foremost, the relationship has to be concrete: if it were only envisaged speculatively, the golden rule would be bound to be a self-centred mental experiment, whereas one of the main features of this principle is the mediation of one's own desire through another's perspective. Moreover, this concrete relationship also has to be oriented towards mutual existence and flourishing. Using classical jargon, we can say that one's relationship must involve a virtuous²¹ partner. Conversely, being involved with a vicious²² partner (that is, one who does not recognise us) would contradict the very nature of the golden rule. Indeed, in this situation, if one wants to know one's own good, one will refer to the perspective of another who is disinterested or, even worse, hostile; as a result, one will turn out to desire what one in truth would not desire.

3. Objections

Besides the opportunities for an intercultural and intersubjective approach to natural law, it seems that the objection can be raised that this interpretation is able to work only within virtuous practices (or ethos). It appears that the possibility of referring back to some sure guidelines for behaviour is, in a certain way, undermined because it depends on a previous favourable attitude towards the other. In Hegelian terms, it depends on the willingness to recognise the other as a transcendental subjectivity (Vigna 2015).

Conversely, it can be also noted that no strong justification of the structural human tendencies can be found in Thomistic works unless we consider the metaphysical framework of the human soul.²³ To sketch his anthropological

consideration of their needs, which we never abandon. So it is that we arrive, almost physiologically, at a measure of "normality" contained in the [Golden] Rule both as an objective and as a result' (2008, 220).

²¹ Within the Thomistic perspective, the virtue in question is that of 'friendliness' (*affabilitas*), a part of the virtue of justice. See ST, II-II 114.

²² Generally speaking, according to Aquinas, vices opposed to friendliness are 'flattery' (*adulatio*; ST, II-II 115) and 'quarrelling' (*litigium*; ST, II-II 116).

²³ On this metaphysical structure many natural-law theorists of the modern age used to base directly the main principles of the natural law: the abstract essence of the human being was considered the reference point for a deduction (e.g., Suárez 1872, book II, chap. VIII ; Grotius 1712, book I, chap. I, § XII). This attitude contributed in tightening up the doctrine of natural law, which certainly turned

model, Aquinas starts from the shared ethos of his time, which does not require any particular effort in identifying what the specific fundamental needs of human beings are. This might explain also why Aquinas makes a sketch, rather than a precise list of these human needs.

Returning to the objection, does this interpretation require too stringent conditions, since a relationship of mutual recognition is needed? It does not seem so, if we refer to the nature of human desire, which is boundlessly open to all things. As said earlier, its fulfilling object—within our concrete experience—is another human being, who has to be treated with regard to his or her ontological stature. If the other is treated like a mere finite being, a very significant gap between the request of desire and the subsequent response will remain. Thus the attitude is bound to bring about dissatisfaction, which might open the possibility of reconsidering one's approach to the other from a different perspective: that of recognition.

To summarise: It is true that this proposal can work only within a virtuous intersubjective context. Nonetheless, it seems also that every vicious practice (or ethos) is unstable because of the disequilibrium in relationships it brings about. Thus, the requests of human desire may turn the tide towards mutual recognition, which is the starting point for building virtuous human practices.

4. Positive Suggestions

Turning to the strengths of our interpretation, we can first of all say that it can play a significant role in building an intercultural ethics, avoiding at the same time abstract perspectives and relativistic solutions. In fact, since the golden rule is part of the heritage of most cultures, every human tradition is internally provided with a resource that is simultaneously particular (that is, formulated in a certain way, with unique connections to its symbolic universe) and universal. This resource can become a vital starting point for a shared ethics, following the

out to lose its historical plasticity—for example, within the system of Christian Wolff (1744, pt. I, chap. II). These versions of the doctrine of natural law deeply influenced the following theorists and contributed to the consideration of the natural law as a rigid discourse incompatible with the challenges of historical and cultural diversity (Fuchs 1996). For a contribution to overcoming this idea of natural law, see Hall 1994.

MacIntyrean aim to find a moral universality within the different traditions (MacIntyre 1990).

Second, this interpretation does justice to human nature's ordinate intertwining of transcendental and empiricity. The primacy of the transcendental dimension²⁴ within human nature justifies the enquiry into all human needs within the context of intersubjectivity. Moreover, the fact that the human being is entirely related to the transcendental horizon explains the concern about understanding the different cultural interpretations of human needs: in fact, being related to the transcendental horizon means these needs have a symbolic dimension. Given the intersubjective nature of the human being, it appears inevitable that that symbolic dimension flourishes within intersubjective practices.²⁵

Third, it seems this reading could easily connect natural law, virtues, and traditions, seen as a unified development, in that we can distinguish natural law as the condition of possibility (see ST, I–II 94.3) of a disposition to act morally—that is, virtue. In turn, the virtues underpin different practices, which give substance to each particular tradition (MacIntyre 2007).

I can conclude that the connection between natural law and the golden rule may open new possibilities in both the debates on intercultural ethics and on natural law. As we saw, this perspective is not devoid of weaknesses; nonetheless, it promises further fruitful results in the future.

²⁴ Within its history, the natural law has often been interpreted in an anthropological background in which the transcendental dimension of the human being is neglected. Upon neglecting this dimension, the cultural elaboration of human needs turns out to be forgotten and the natural law runs the risk of being considered as a mere focus on human biology.

²⁵ Jean Porter remarks correctly: 'While the scholastics hold that we can understand our fundamental inclinations by analogy with the inclinations exhibited by nonrational animals, they also recognize that even the most fundamental human inclinations are not experienced as the other animals would experience them. Normal adults experience these inclinations in and through the mediation of some kind of rational reflection, and this experience is further qualified and shaped by the cultural forms through which the inclination is expressed. In this way, even our most basic inclinations are inextricably bound up with the exigencies of our life as rational and social creatures, and we cannot adequately interpret them unless we see them within the context of human life considered as a whole' (2005, 75). See also Hall 1994.

References

Aquinas' works (Latin text of the Leonine edition)

DE=De Ente et Essentia

QDM=Quaestiones Disputatae de Malo

QDV=Quaestiones Disputatae de Veritate

ST=Summa Theologiae

Other works

Campodonico, A. (2013). 'How to Read Today Natural Law in Aquinas?' *New Blackfriars* 4: 716–32.

Catapano, G. (2005). 'La Regola d'Oro in Agostino'. In *La Regola d'Oro come etica universal*, edited by C. Vigna and S. Zanardo (pp. 103–38). Milano: Vita e Pensiero.

Cunningham, L. (ed.). (2009). *Intractable Disputes about the Natural Law. MacIntyre and His Critics*. Notre Dame, IN: Notre Dame University Press.

De Finance, J. (1993) *Cittadino di Due Mondi. Il Posto dell'Uomo nella Creazione*, translated by Monache Benedettine dell'isola di San Giulio. Città del Vaticano: Libreria Editrice Vaticana.

De Castro, A. (1568). *De Potestate Legis Poenalis Libri Duo*. Antverpiae: Stelsii.

Du Roy, O. (2008). 'The Golden Rule as the Law of Nature, from Origen to Martin Luther'. In *The Golden Rule. The Ethics of Reciprocity in World Religions*, edited by J. Neusner, & B. Chilton (pp. 88–98). London: Continuum.

Du Roy, O. (2012). *La Règle d'Or: Histoire d'Une Maxime Morale Universelle*. Paris: Du Cerf.

Fuchs, J. (1996). *Ricercando la Verità Morale*. Cinisello Balsamo: San Paolo.

Gensler, H. J. (2013). *Ethics and the Golden Rule*. New York: Routledge.

Hall, P. M. (1994). *Narrative and the Natural Law. An Interpretation of Thomistic Ethics*. Notre Dame, IN: Notre Dame University Press.

Hegel, G. W. F. (2000). *Fenomenologia dello Spirito*, translated by V. Cicero. Milano: Bompiani.

MacIntyre, A. (1988). *Whose Justice? Which Rationality?* Notre Dame, IN: University of Notre Dame Press.

MacIntyre, A. (1990). *Three Rival Versions of Moral Enquiry: Encyclopaedia, Genealogy, and Tradition*. Notre Dame, IN: Notre Dame University Press.

MacIntyre, A. (2000). 'Natural Law in Advanced Modernity'. In *Common Truths. New Perspectives on Natural Law*, edited by E. B. McLean (pp. 91–115). Wilmington, DE: ISI Books.

MacIntyre, A. (2006). *Ethics and Politics. Selected Essays (vol. 2)*. Cambridge: Cambridge University Press.

MacIntyre, A. (2007). *After Virtue. A Study in Moral Theory*. Notre Dame, IN: Notre Dame University Press.

MacIntyre, A. (2009). 'Intractable Moral Disagreements'. In *Intractable Disputes about the Natural Law. MacIntyre and His Critics*, edited by L. Cunningham (pp. 1–52), Notre Dame, IN: Notre Dame University Press.

Neusner, J., and B. Chilton, eds. (2008). *The Golden Rule. The Ethics of Reciprocity in World Religions*. London: Continuum.

Pagani, P. (2012). *Ricerche di Antropologia Filosofica*. Napoli-Salerno: Orthotes.

Pagani, P. (2014). 'Sull'Attualità del Concetto di "Anima"'. *Aquinas* 56: 425–39.

Porter, J. (2005). *Nature as Reason. A Thomistic Theory of the Natural Law*. Grand Rapids, MI: Eerdmans.

Puka, B. (n.d.). *The Golden Rule*. In *Internet Encyclopedia of Philosophy*. Retrieved from <http://www.iep.utm.edu/goldrule/>

Sciuto, I. (2005). *La Regola d'Oro nel Pensiero Medievale*. In *La Regola d'Oro come Etica Universale*, edited by C. Vigna, and S. Zanardo (pp. 85–102). Milano: Vita e Pensiero.

Suárez, F. (1872). *Tractatus de Legibus ac Deo Legislatore in Decem Libros Distributus*. Neapoli: ex Typis Fibrenianis.

Vigna, C. (2005). *Sulla Regola d'Oro*. In *La Regola d'Oro come Etica Universale*, edited by C. Vigna, and S. Zanardo (pp. xvii–xlviii). Milano: Vita e Pensiero.

Vigna, C. (2008). 'Human Universality and Natural Law'. In *Contemporary Perspectives on Natural Law. Natural Law as Limiting Concept*, edited by A. M. González (pp. 211–25). Aldenshot: Ashgate.

Vigna, C. (2015). *Etica del desiderio come etica del riconoscimento*. Napoli-Salerno: Orthotes.

Vigna, C. (2016). *Sostanza e relazione. Indagini di struttura sull'umano che ci è comune*. Napoli-Salerno: Orthotes.

Wattles, J. (1996). *The Golden Rule*. New York: Oxford University Press.

Wolff, Ch. (1744). *Philosophia Practica Universali Methodo Scientifica Pertractata*. Halae Magdeburgicae: in Officina Libraria Rengeriana.

Zanardo S. (2017). *Dentro e oltre Babele*, *Dialoghi*, 3, 40–47.

SYMPOSIUM: HAYEK-POLANYI



HAYEK AND POLANYI FOR THE 21ST CENTURY – AN INTRODUCTION

HAYEK E POLANYI PARA O SÉCULO 21 – UMA INTRODUÇÃO

António Baptista*
alsbaptista@gmail.com

Pedro A. Teixeira**
pedro.teixeira@fu-berlin.de

Why a special issue on these two authors and why now?

A simple – though inevitably incomplete – answer can be traced back to the academic activities of the editors. As members of CEPS (the Center for Ethics, Politics and Society of the University of Minho), the editors take part in monthly informal seminars and discussions on various authors and topics of either contemporary or everlasting relevance to the domain of ethical and political theory. During the first half of 2016, we focused our attention on two classic works of two authors whose insights we thought could help us grasp some of our society's present problems. These two works were Friedrich Hayek's *The Road to Serfdom* and Karl Polanyi's *The Great Transformation*, both published in 1944.

The fullest answer to these questions (why them and why now) lies, however, in the content of Polanyi's and Hayek's works, or better still, in the topics and enduring concerns of both authors and their connection with recent events and the spirit of our time.

* Integrated Researcher at CEPS (Center for Ethics, Politics and Society), University of Minho.

** Ph.D. Candidate at the Otto-Suhr Institute for Political Science, Free University of Berlin.

While debates on the merits of different forms of economic planning versus free markets took central stage in economic theory for the most part of the 20th century – even if the importance of these discussions dwindled in the past twenty years – both *The Road to Serfdom* and *The Great Transformation* remain particularly significant beyond the fields of political economy or “mainstream” economics. The innovative and sometimes counter-intuitive arguments they put forward and the fact that they represent two opposite views on almost every single topic they address – be it socialism and social democracy, the nature and intrinsic qualities of markets versus planning, the explanation of the origins of fascism, the fate of economic liberalism or the essence of human freedom, etc. – must surely account for a share of their intellectual and political impact. Another factor, which undoubtedly contributed to the popularity and relevance of both works, was the timing of their publication: they came to light at a time when, on the one hand, nazi-fascism had just been defeated by the allied forces and extensive economic planning was almost seen as a legitimate, indeed necessary, fact of life in almost all developed countries, socialist and capitalist; and when, on the other hand, the ideological and political tide was about to change against socialism – even if not yet in the direction Hayek would have preferred.

Despite the backlash against socialism and the political left in the years after Franklin Roosevelt’s death (climaxing in the McCarthy witch-hunt era), it seems that Polanyi was still naively optimistic in the late 1940s regarding the progressive superseding of what he called “our obsolete market mentality” (Polanyi, 1947). To be fair, it is true that, to a certain extent, Hayek’s perspective remained in the shadow in the 30 years following the end of WWII; but starting in the late 70s, as Keynesianism receded, neoliberal thought resurrected the interest in Hayek’s arguments. In the 80s, with the election of Reagan in the US and Margaret Thatcher in the UK, a window of opportunity opened for producing experiments in the economy that reversed what had been the usual policies in welfare capitalist states while these leaders were openly professing that the inspiration for them resided in the works of Hayek and other “libertarian” authors, who were now presented as respectable and much more insightful than the then unfashionable Keynesians. This window of opportunity was profoundly widened by the events that followed at the turn of the decade.

After the collapse of the Soviet Union, neoliberal triumphalism ensued throughout the 90's and 2000s, along with a euphoric expansion of capitalist globalization. Capitalism and “liberal democracy” became rhetorically entwined as if they were co-dependent phenomena or part of a single concept. Moreover, they were presented not only as “the only alternative” – since the “other” had “lost” – but also as ethically indisputable. Any notion of socialism or even economic and social planning became synonymous with backwardness, both in academic and political circles.¹ Even plain resistance to economic globalization, privatization and deregulation was regarded as both futile and morally wrong.

Anti-globalization movements and protests were often portrayed by media, politicians or academics as either the product of chaotic and angry crowds or of romantic rebels, who, much like the Luddites, were desperately trying to stop the destruction of obsolete institutions and realities that had to – and inevitably would – be eliminated by the force of historical necessity and the irresistible pressures of the (world) market.² Paradoxically, although Margaret Thatcher had explicitly identified Hayek's political philosophy as a source of inspiration and an ideological guideline,³ Hayek's triumph looked more complete than ever only when ‘third-way’ politicians gained the upper hand within socialist/social-democratic parties and subsequently rose to power. Former “opponents” had now adopted his ideas, just as he had hoped when he dedicated *The Road to Serfdom* “to socialists of all parties”.

Indeed, the modern socialist or labour politicians of the late 20th century, such as Tony Blair (Heath, Jowell, Curtice, 2001, pp. 2-7 and 101-110) or Gerhard Schröder, who would together publish a joint declaration or manifesto (*The Third Way/ Die Neue Mitte*) explaining their views and strategy for European socialist parties, came to accept the full legitimacy of markets and market forces and

¹ This was symbolized by the British Labour Party's abandoning of clause IV of its constitution.

² Judging by his treatment of the Speenhamland system, one suspects that, had Polanyi lived to witness these resistance movements of the 90's, he would have probably held a very different and sympathetic view towards them. Instead of perceiving these movements as a sort of “reactionary romanticism” of those wanting to “turn the wheel of history backwards”, but rather as valid and spontaneous attempts - with different degrees of success – of society to protect itself from yet another renewed offensive bent on realizing the destructive utopia of a self-regulated market society.

³ “Our inspiration was less Rab Butler's *Industrial Charter* than books like Colm Brogan's anti-socialist satire, *Our New Masters*, (...) and Hayek's powerful *Road to Serfdom*, dedicated to “the socialists of all parties”. Such books not only provided crisp, clear analytical arguments against socialism (...) but (...) they also gave us the feeling that the other side simply could not win in the end. That is a vital feeling in politics; it eradicates past defeats and builds future victories. It left a permanent mark on my own political character, making me a long-term optimist for free enterprise and liberty and sustaining me through the bleak years of socialist supremacy in the 1960s and '70s” (Thatcher, 1993, pp. 12.-13).

looked upon state interference in the economy as very often cumbersome, inefficient and, from the ethical point of view, unbearably “paternalistic”.⁴ According to the authors affiliated with the Third Way, the left – which for too long had regarded itself as the legitimate interpreter of the people’s needs – would, once in power, feel entitled to define and provide for each citizen as it saw fit.⁵ The “old” left, or so the story went, viewed itself as an authority that knew better than the beneficiary what was in his or her best interest (Le Grand, 2003, pp.6-7; Corrigan et al., 1988, p.12). A radical change in social policy and in the way the welfare state was understood and managed was imperative. As Blair himself put it: “the modern welfare state is not founded on a paternalistic government giving out more benefits but on an enabling government that through work and education helps people to help themselves” (Bevir, 2005, p. 90) Accordingly, the task at hand, then, was: 1) for the state to remove itself from the economic landscape, giving way to non-authoritarian, competitive providers of goods and services; 2) to free private economic actors from stifling regulations, excessive taxation and state bureaucracy in order for them to unimpeded pursue their entrepreneurial abilities and, in so doing, energize the economy and raise productivity; 3) to unburden the state from its own excessive bureaucracy and unsustainably growing expenses with social welfare; 4) and to stimulate individual responsibility.⁶ And it was in this context that the policies of

⁴ All these elements can be easily gleaned from the Blair/Schröder manifesto:“(…) we need to apply our politics within a new economic framework, modernised for today, where government does all it can to support enterprise but never believes it is a substitute for enterprise. The essential function of markets must be complemented and improved by political action, not hampered by it. We support a market economy, not a market society (...) Today we must develop realistic and feasible answers to new challenges confronting our societies and economies. This requires adherence to our values but also a willingness to change our old approaches and traditional policy instruments. In the past: The promotion of social justice was sometimes confused with the imposition of equality of outcome. The result was a neglect of the importance of rewarding effort and responsibility, and the association of social democracy with conformity and mediocrity rather than the celebration of creativity, diversity and excellence. (...) The belief that the state should address damaging market failures all too often led to a disproportionate expansion of the government’s reach and the bureaucracy that went with it. The balance between the individual and the collective was distorted. Values that are important to citizens, such as personal achievement and success, entrepreneurial spirit, individual responsibility and community spirit, were too often subordinated to universal social safeguards. Too often rights were elevated above responsibilities, but the responsibility of the individual to his or her family, neighbourhood and society cannot be offloaded on to the state (...)” (Cf. Blair and Schröder, 2000 [1998]).

⁵ In Britain, the more distinguished and articulate academic “ideologues” of this reformed perspective on the relations between market and society were arguably Anthony Giddens (1998) and Julien Le Grand (2003) – this latter author admitting that his “quasi-markets” had more or less been defended and implemented by the Thatcher government itself (pp. 9-10).

⁶ Much the same can be said of the transformation within the Democratic Party in the US with the arrival and election to office of Bill Clinton. As Joseph Stiglitz admits, the New Democrats were recentring the party to the “centre” (i.e. to the right): they wanted to show how they were “fiscally responsible” and spent most of Clinton’s term in office obsessed with balancing the budget, cutting federal (social) programs and reducing public jobs (Stiglitz, 2003 pp. 19 and 48 ff), making market-friendly reforms, deregulating and, to sum it up, yielding to the interests and ideology of the main actors of the financial system. A crucial instance of this is the repeal of the Glass-Steagall law, ushered in the aftermath of the Great Depression during FDR’s presidency and banning the fusion between investment and commercial banks. This repeal was the product of a consensus between the Clinton Administration and the republican majority in Congress and is considered to have been both a major concession to financial interests and a fundamental step in the deepening of the tendencies that led to the banking crisis of 2007. These new democrats, or a great deal of them, seemed to be convinced that financial markets – along with the then revered Alan Greenspan – were not only rational, but that they knew better than anyone else how to get the economy going (Stiglitz, 2003, pp. xiv and 32). Reagan and the republicans had been profligate with public spending, especially regarding military spending, according to Stiglitz, the idea was to “starve the beast” and force cuts on social state spending further ahead (Stiglitz, 2003, p. 47). And indeed, those cuts would take place in the Clinton administration, which

privatization (as in the case of the British Railway), of private-public partnerships, concessions, liberalization, and deregulation⁷ that have been the trademark of economic policy in recent decades (Crouch, 2009) came into being.

What emerged from this ideological and policy shift was a return to the radical boom-and-bust economic cycles that had become fairly moderate and increasingly rare in the preceding period but had been a usual feature of 19th century liberalism. Along with this came job precariousness, labor flexibility, looser labor protection laws and the further weakening of trade unions as measured in their dwindling numbers and eroding bargaining force. While in their manifesto Blair and Schröder seemed to pay a weak lip service to Polanyi's concerns by stating that "we support a market economy, not a market society" and through announcing that "the past two decades of neo-liberal laissez-faire are over" – despite actually entrenching and deepening it during their time in office –, it was clear that their goals and inspiration for their policies resided rather in Hayek's arguments and discourse. Polanyi would have arguably seen their project and actual policies as a thinly disguised renewal of the utopia of the "self-regulating" market system.⁸ The problems that extensive planning, regulation, protectionism and Keynesianism were meant to solve – e.g. mass unemployment and the destructive fluctuations of economic activity in the 30's – are now essentially back. Economic inequality, after 30 years of silent rise, is again at the spectacular record 1913 levels, and there seems to be no limit to that trend. The growth in inequality in income or wages, although common and very patent in all

also significantly diminished the workforce employed by federal government to unprecedented levels since the New Deal (Stiglitz, 2003, p.19). But the New Democrats acceptance of the ideas of their adversaries is also visible in the way Clinton put in place a severe and punitive attack on welfare recipients. In his candidacy speech to the democratic primaries (Clinton and Gore, 1992, pp. 187-198) he had appealed systematically to "the middle class" - impressively referred 11 times in an 11 page document - instead of "workers", and had focused very strongly on the topic of restoring individual responsibility in what he saw – just as the average republican did – as an unduly generous welfare state. He promised to "end the welfare state as we know it" (Clinton and Gore, 1992, p. 165 and 228) and this was a promise he fulfilled with the enactment of the Personal Responsibility and Work Opportunities Reconciliation Act (PRWORA) in 1996, that remarkably replicated the same punitive logic and arguments that had presided to the Poor Law reforms in the 1830's in England. In terms much similar to those employed by Malthus to condemn the Speenhamland system, the defenders of PRWORA sustained that the poor were guilty of their own situation and that the "generous" assistance they received was giving them an incentive to remain idle and poor. Cutting back the system would, so the argument went, help them help themselves. It was not cruelty, but actually a demonstration of concern for their welfare in the long run (Block and Sommers, 2012, p. 150-192).

⁷ Perhaps, in a Polanyian fashion one ought to speak not so much of "de-regulation", as markets always require some kind of state regulation in order to work, but rather of a "re-regulation" geared towards providing further advantage to different and privileged social groups, as Fred Block and Margaret Sommers argue: "By the term reregulation, (...) we aim to push back against the belief that the success of neoliberal ideology since the mid-1970s has been matched by markets being increasingly freed from regulations and government management. On the contrary, regulations did not go away; they simply changed. Those that had previously been written to protect employees or consumers were systematically rewritten to support business interests and reduce previous restrictions on business practices. Similarly, the tax code was rejiggered to shift the burden from high-income households to middle class and working class earners." (Block and Sommers, 2012, p. 20)

⁸ As the joint declaration of the Third Way point by way of a metaphor: "The state should not row, but steer: not so much control, as challenge. Solutions to problems must be joined up" (Blair and Schröder, 1998).

countries (Piketty, 2014, p. 271 ff), has been more extreme in the Anglo-Saxon world, particularly the US, apparently giving some credence to the Polanyian notion that specific cultural institutions and mentalities do have some autonomous and significant effect on the way the market is structured or regulated (Piketty, 2014, p. 315 ff). In any case, shifts in technology and educational differences, whatever their long run effects, cannot credibly account for the veritable explosion of income inequality, either in the US or in the world at large, at least not entirely. The main culprit for this is to be found elsewhere, possibly in the power imbalances between social groups that are associated with the rapid political transformations that took place in the late 20th century. The concentration of wealth, however, has been much more extreme than that of income, and this is observed in many countries (Piketty, p.336 ff). The hyperglobalized economy, promoted by the governments of the major economies and by international organizations enthralled by market fundamentalism such as the WTO, the IMF and World Bank, has taken away from nation-states a number of instruments for control and regulation of their domestic economies, as well as much of their bargaining strength vis-à-vis transnational (and national) capital, which may now flee and hide from fiscal authorities by moving to all sorts of tax havens.⁹ This has led Wolfgang Streeck (2014) and even mainstream economists such as Dani Rodrik (2011) to come out against free trade fundamentalism and denounce the incompatibility of this sort of hyperglobalization with the superior goals of national sovereignty and democracy. Some supra-national institutional arrangements that were presented as opportunities for economic and social progress and prosperity such as NAFTA, but even more so the European Union and its project of a Monetary Union seem to have aggravated some of these problems and tendencies, infused as they were by a more or less Hayekian ideology. The unrestricted freedom of circulation of goods and capital within the EU resulted in the de-industrialization of several national economies, increased the vulnerability to external economic shocks, incentivized crippling fiscal and social dumping practices between member-states, and led to dangerous trade imbalances within the EU and even beyond it. The adoption of the Euro, in

⁹ The lack of strong political action to address the recent though entirely unsurprising disclosures (e.g. the Panama Papers, the Luxembourg Leaks, and the recent Paradise Papers) about the staggering amounts of wealth stashed away in tax havens and other low-taxation jurisdictions (Cf. Zucman, 2015) is perhaps a damning testament of the dismal condition of our current capacity to act as members of a polity.

particular, has become a destructive straitjacket for peripheral economies in the EU (Stiglitz, 2016) that, in some respects, reminds one of the criticisms directed at the gold standard by Polanyi.

Political equality and popular participation too have now sunk to some of their lowest levels since universal suffrage was first achieved, as the decision-making process is now deeply distorted by the power of concentrated wealth, reinforcing the already existing oligarchical tendencies within capitalist polyarchies. This has been confirmed by some studies in the field of political science (Jacobs and Skocpol, 2005, p.11; Bartels et al, 2005, p.113). Among the general public, the perception is in accordance with these findings. In other words, the political process is seen as “broken” beyond repair, with the wealthy being able to manipulate it for their own benefit, be it through corruption, media manipulation, or economic coercion while the majority of the population, systematically impoverished and sacrificed in the name of an unequally applied austerity, feel the sting of their own political disenfranchisement. This has fueled a deep suspicion regarding politics and mainstream politicians (seen as part of an unresponsive and irresponsible elite) which may very well be an explanation for the upsurge of what are now termed “populist” movements and candidates.

If a renewed understanding, achieved in a somewhat comparative fashion, of the contributions of both authors to the fields of social and political theory, as well as to the discipline of political economy, does not in itself warrant this special issue, one may also remind the reader of the striking similarities between today’s events and those of the times in which Hayek and Polanyi lived. The problems we are currently witnessing, and which portend a certain breakdown of our societies as we know them, are serious enough to justify the objective lurking in the background: to find new solutions to our current predicaments by looking back to an intellectual dispute whose focal points retain their fundamental relevance in today’s complex societies.

In the last part of this introduction, we will briefly summarize the content of the contributions to this special issue. However, before we proceed to that final section, we believe that it is important to provide some historical background and contextual biographical information regarding the lives of both authors so as to

better understand the formation of their ideas, as well as their academic and political concerns.

A biographical introduction to Polanyi and Hayek

A - Polanyi

Karl Polanyi was born in 1886. He was a “child of late liberalism”, as Gareth Dale characterizes him, and this was a “civilization that (...) seemed to rest upon solid foundations” (Dale 2010, p.1). However, as the Hungarian author was soon to discover, as he entered adulthood and WWI approached, this comfortable stability was not to endure indefinitely. Born in Vienna into a wealthy (but later impoverished) and intellectual Hungarian Jewish family, Polanyi was raised in Budapest. He entered Budapest University to start a law degree and there he became the founder and first president of Club Gallilei. This club tried to unite progressive liberals and socialists in a joint cultural war against conservative clericalism, anti-Semitic reactionarism and ignorance. However, due to his involvement in a dispute between a liberal-minded teacher and the conservative university board, Polanyi was expelled from the university and only received a law doctorate in 1909 at Kolozsvar University (Maucorant, 2007, pp.37-38). Despite his friendship with figures on the radical left, including communists, such as a György Lukács, he was elected secretary general of the Radical Party (of Oszkár Jászi) and supported the liberal Hungarian government of Count Károlyi that came to power after the liberal democratic revolution of October 30th 1918, known as the “Aster Revolution” (Dale, 2010, p. 7). This government was later displaced by the revolutionary communist government of Béla Kun.¹⁰ The subsequent government came to be defeated by the reactionary forces of Horthy, and possibly anticipating this tragic reactionary end (Block and Sommers, 2012, p. 46) , Polanyi had moved shortly before to Vienna and found work there as a

¹⁰ Karl Polanyi always stood on the left side of the spectrum and could be broadly thought of as a “socialist”, although he was throughout most of his life politically quietist and fairly hesitant (Dale, 2010, p.11). One may also question if his notion of what socialism actually consisted of (or ought to consist of) did not vary significantly at different stages of his life. Another very controversial and possibly unsolvable question regards his theoretical affinities and acceptance of Marxism and Marxist concepts – which seems to have varied greatly through time. Gareth Dale (2014) explores some of these topics in a recent article.

journalist in *Der Österreichische Volkswirt*, writing a large number of articles mainly concerned with contemporary international affairs (Cangiani, 2011, p.3).

The politically and intellectually agitated Vienna of the 20's and early 30's was where Hayek and Polanyi coincided and where their political and economic views started to take shape: Polanyi as a more consistent and sympathetic adherent of socialism (particularly as it was advocated and practiced by the Austrian Social Democratic Party)¹¹ and Hayek as an anti-socialist. While it would be well beyond the scope of this introduction to attempt a portrayal of Austria and Vienna during these troubled and exciting years, it is useful at this juncture to make a brief characterization of this period in its social, political and intellectual aspects, as it is thought to have had an important impact on the authors to which this special issue is dedicated.

Throughout the second half of the 19th century and early 20th century, and much like other countries in Central Europe, Austrian society underwent a process of economic and political transformation. Industrial and financial capitalism was disturbing the equilibria of the past, concentrating wealth, industry and population in ever larger cities, while in the countryside feudal remains were being eliminated, along with the economic autonomy of many farmers who ended up inevitably in Vienna and other booming cities. Much of this transformation had its center in the intellectually flourishing capital of the Austro-Hungarian empire, Vienna. Along with these transformations, a number of tensions had been accumulating within the empire: those between the different nationalities that were comprised in it (with the Austrian and Hungarian elites dominating) and those between the radicalized working class, led by the increasingly powerful, both electorally and socially, Marxist social-democratic party and its trade-unions, on the one hand, and, on the other hand, the social and economic elites (bankers, businessmen and aristocrats) that resisted them and who felt increasingly under siege. While the issue of conflicting nationalities was to be solved after the end of WWI, with the declaration of independence of national minorities, the social conflict remained intact. To many, it felt as if the

¹¹ "(...) In Vienna Polanyi's old familiarity with socialist ideas turned to full adhesion; (...)The extremely stimulating reality of the *rote Wien* was as important for Polanyi's formation as the milieu of progressive-radical Budapest students before the war. The socialist municipality of Vienna became the seat of memorable intellectual and political advancements, recalled with never-ending enthusiasm in one of the 'Notes on the sources' added to *The great transformation*. Polanyi maintained a fruitful relationship with Austrian socialists, and was in particular influenced by their leader, Otto Bauer;(...)" (Cangiani, 2009, p. 3)

situation was fast approaching a deadlock. “Red Vienna” was where social democracy had its fortress. From 1919 onwards, Vienna was ruled by the Social Democratic Party that began applying a number of municipal social-democratic policies (mainly in the domains of housing, public health, educational reform and cultural utilities), geared to the improvement of the working class’ material conditions but also, and most importantly, of their “class consciousness” in the hopes of preparing it for a not so distant future in which socialism would be fully implemented after a decisive electoral triumph. Vienna was to be a model of the future socialist republic (Rabinbach, 1983, pp. 23-26). In fact, Otto Bauer, the leader of Austrian social-democracy, for whom Polanyi had shown enthusiasm and with whom he corresponded (Dale, 2014, p. 52), spoke of his strategy of a transition from capitalism to socialism as that of a “slow revolution”, which was a way of threading a sort of “middle path” between the reformism of Bernstein and the revolutionism of Rosa Luxembourg or Lenin’s communism (Rabinbach, 1983, p. 29). According to Bauer, the situation in Austria and Europe after the Bolshevik revolution was one of “balance” of forces (Gullick, 1948, pp.1374-1380), with neither bourgeoisie or proletariat being able to yet establish a clear “hegemony”: “(...) an idea that caught the attention of Polanyi (Dale, 2014, p.54). At the same time, Bauer was convinced that electoral victory for social-democracy was at hand given the growing triumphs outside of the capital (...)” – they would reach 42% of the national vote on April 1927 – and, for the most part, so were his enemies (Rabinbach, 1983, p. 32). As tensions rose during the 20’s, partisan militias were organized (the socialist Schutzbund on the one hand, and right wing paramilitary organizations, on the other). This political struggle expanded and “infected” all spheres of life, including intellectual and academic circles, in a very powerful way. It was impossible to keep a neutral stance and scientific discussions – in social sciences as well as in epistemology and philosophy – were, to a great extent, a continuation of these ideological battles.

While it is true that scientific disputes can very often be nothing but a cover for political disputes,¹² this had become poignantly patent in the case of the vibrant scientific scene of Vienna in the 20s. Even when theories were not

¹² As Benjamin Ginsberg puts it: “Even in the sciences, ostensibly the bastions of objectivity, new theories, procedures and findings are seldom accepted simply and immediately on their own merit (...) Indeed, at times what appear to be strictly scientific disagreements actually have political roots or are in fact political disputes couched in scientific language” (Ginsberg, 1986, pp.109-110).

partisan, nor could be said to be openly favoring one political side or another, their authors were. Karl Polanyi and Friedrich Hayek were obviously involved in these disputes and had quite openly taken sides, in many instances against their relatives' and friends' own positions.¹³

In 1933, the tensions within Austrian society would lead to what can be described as a preventive right-wing coup led by Engelbert Dolfuss and his conservative Christian government. With a repressive environment taking over the country and the press, Polanyi's co-workers at *Der Österreichische Volkswirt* invited him, given his standing as the most prominent left-winger of the newspaper and an open supporter of Bauer's municipal socialism, to leave while he could. Prudently, he did so. The new fascist-inspired regime would immediately suspend the parliament, ban the communist party (as well as their "rivals on the right", the Austrian Nazis) and soon all other parties, while instituting press censorship and prohibiting public meetings. In 1934, Austrian workers rebelled against the dictatorship and repression – an event Polanyi did not witness first-hand –, instigating a very short civil war which ended in the complete defeat of any remaining left-wing opposition to the regime (Gullick, 1948, pp.1266 ff). Things only worsened after the German Anschluss in 1938. Polanyi's wife, Ilona Duczynska, managed to escape and join him in London after taking part in this fleeting and desperate last episode of resistance.

Life in England did not prove as satisfying as he had once expected. Indeed, he languished in obscurity, not being able to find a suitable position at any university and being, for the most part, ignored within socialist intellectual circles. Despite his interest, intellectual admiration and sympathy for the Christian socialism of Tawney and the guild socialism of G.D. H. Cole – which served as a framework for his own version of what socialism ought to look like in a 1922 article (Bockman, Fischer and Woodruff, 2016 [1922]) – neither these nor other British socialists seemed to reciprocate his feelings (Rogan, 2013). He earned his meagre living by “teaching adult education courses to British workers through the Workers' Educational Association—the extramural outreach arm of

¹³ For instance, Karl's elder brother, Michael, was already expressing opposition to Marxism and socialism and would later become a founding member of the Mont Pellérin Society, along with Hayek and von Mises, whereas Wittgenstein – Hayek's cousin (Kresge and Wenar, 2005, p. 3) – is thought to have been rather sympathetic towards socialism and the Soviet Union (Gakis, 2015), even though he generally avoided engaging in politics and expressing (obvious or otherwise explicit) political content in his works.

the Universities of Oxford and London” (Block and Sommers, 2012, p. 5). While feeling somewhat isolated and under-appreciated by his peers, he used his time on the islands to develop extensive and proficuous research on the economic, social and political history of Great Britain, with much of this research material later becoming the basis for some of *The Great Transformation*’s best chapters.

In 1940 he was invited to deliver a series of lectures at Bennington College, in Vermont, US, and during this time he started the writing of his *magnum opus*. This would be, yet again, the continuation of the politically charged debates on economic planning and socialism which he had led in Vienna against Von Mises and Hayek. Beyond serving the minor goal of feeding academic research with a most valuable input, it was also written as a way for the civically engaged left-wing intellectual to continue, once again, the fundamental ideological and political battle of the day for the minds and hearts of the men who could usher a new civilization, one that would finally realize the liberal ideals of genuine freedom, if only they were freed from the destructive utopia of the self-regulating market that had been, in his view and in the last instance, the origin of all the evils of the 20th century. *The Great Transformation* was published in 1944, slightly after Hayek’s *The Road to Serfdom*. In 1947 Polanyi was invited to teach at Columbia University. Due to his wife’s past as a communist revolutionary, however, Polanyi decided to move to Canada, from where he drove to the university. In the US, and after the publication of the article “Our Obsolete Market Mentality” (1947), he avoided contemporary politics as an academic subject for the most part, very likely as a self-protecting measure against the tense intellectual environment that marked the beginning of the Cold War. For the rest of his life, he moved on to conduct anthropological research on primitive societies and economies of the present and past. He retired from Columbia University in 1953 and remained in Toronto, Canada, until his death in 1964 (Block and Sommers, 2012, p. 5).

B - Hayek

Friedrich Hayek was born in 1899 in Vienna into a non-practicing Catholic family, and was, therefore, 13 years younger than Polanyi. His father’s side of the family had always displayed an inclination for the natural sciences, and Friedrich

himself maintained a great deal of interest for these areas, although he would first enter the university to study law. Before ending the *Gymnasium* (secondary school) he fought in WWI for a year in an artillery regiment of the Italian front. Upon his return to Austria and to the University of Vienna he was struck by what he perceived to be a personal change: as the author himself recalls, the war experience, besides giving him “a severe infection of Malaria”, had instilled in him a keen interest in the social and human sciences (Kresge and Wenar, 2005, p.2). The immediate post-war Vienna left a durable impression on the young student: “The streets were filled with returning veterans, many of them unemployed, and because embargoes continued even after the war ended, near famine conditions prevailed in Vienna during the first postwar winter. (...)” (Caldwell, 2007, p. 17). The country and the whole of Central and Eastern Europe were in political turmoil.

At university, he finished a doctoral degree in Law in 1921 – with a strong emphasis on economics – and a second doctorate in political science in 1923. He became acquainted with the epistemological debates then unfolding in the Vienna Circle, even though he was never a member of it and later came to oppose most of their ideas on science.¹⁴ In 1922, Von Mises published “Socialism”, a thorough attack on the notion of economic planning that, in Hayek’s own opinion, would prove to be decisive in the evolution of his economic and political thinking, as it moved him away from his previous “Fabian” inclinations (Kresge and Wenar, 2005, p.5; Caldwell, 2007, p. 18). Indeed, Mises became his friend and mentor. In 1923-1924, with the help of Mises, Hayek developed postgraduate research at New York University regarding inflation and monetary policy (Kresge and Wenar, 2005, pp. 6-8) – a topic he would continue to study and that would render him two books: *Prices and Production* (1931) and *Monetary Theory and the Trade Cycle* (1933).

¹⁴ Again, this rejection was not purely “scientific” at a time when the debate in both science and politics were patently intertwined. Otto Neurath, a leading member of the Vienna Circle of positivists, was also a staunch socialist and had been the minister of the economy in the revolutionary government of the short-lived Bavarian council republic. In Vienna, as an author, he continued to elaborate on his radical ideas on the transition to a socialist economy, with the abolition of money and prices and its replacement with statistical data (*in natura*) to be centrally administered. Other positivists, while less radical in economics and politics, were nonetheless aligned with the SDAP. No wonder then, that right-wing intellectuals grew an almost instinctive opposition to it. “As recent scholarship emphasizes, the early days of the logical positivist movement had a distinctly political side, and Neurath played a central role in this. In advocating the unity of science, for example, he hoped to enlist all of the sciences to use them to refashion society along socialist lines (Reisch 2005). In any event, for the Austrian School economists, positivist philosophy of science was always aligned in their minds with socialist politics and economics.” (Caldwell, 2007, p. 27)

His friendship and collaboration with von Mises continued when they founded the “Austrian Institute for Business Cycle Research” in 1927. Hayek worked there for some time, making use of the expertise with statistics he had acquired while in the US. Unlike Polanyi, who did not enjoy much academic success before leaving for the US in 1929, Hayek had already become *Privatdozent* in Economics and Statistics at the University of Vienna. In 1932, after lecturing the previous year at the London School of Economics at the invitation of Lionel Robbins, he was offered a permanent position. He remained at the LSE for almost 20 years until he moved to the University of Chicago. Prior to this move, he acquired some notoriety in British academic circles by entering a dispute with John Maynard Keynes over the latter’s book *A Treatise on Money* (1930). However, with the publication and later tremendous success of Keynes’ *General Theory* (written in 1936), Hayek was naturally overshadowed (Caldwell, 2007, 22-23). In fact, Keynes book revolutionized economic theory, and marked the end of the era of the hegemony of *laissez faire* ideology, which in practice had already begun to be put aside by governments of all political colors in the capitalist world. In the meantime, Hayek was getting more and more engaged in the public battle of ideas and he became the editor of the book *Collectivist Economic Planning: Critical Studies on the Possibilities of Socialism*, a full assault on the socialist experiments in the Soviet Union and elsewhere. This book continued, extended and made available to the Anglo-Saxon public the attacks and criticisms made by von Mises and others against the possibility and desirability of socialism (the “socialist calculation” debates).¹⁵

The book that would grant him lasting fame, however, was *The Road to Serfdom*, published in 1944 – despite the initial difficulties that Hayek himself narrates in his introduction in finding a non-hostile publishing firm in the US due to the political climate at the time. Indeed, as mentioned above, the book was published when the idea of economic planning was at its zenith even in the US and at a time in which any association between the allied and victorious USSR and the common fascist foe would still seem to fall very short of heresy. It is also true, though, that things rapidly changed in the post-war scenario. Those who

¹⁵ Polanyi, too, had participated in the calculation debate against Mises in an article (Bockman, Fischer and Woodruff, 2016 [1922]) where he defends not a fully centrally planned and directed economy, but a sort of decentralized and democratic cooperative socialism inspired in G.D.H. Cole’s “guild socialism”.

were anxious to see some refraining of these “dangerous” socialist trends surely welcomed Hayek’s warnings. Among these, and even among other groups more sympathetic to socialism, the fundamental contention that extensive economic planning by the state would inevitably concentrate power (instead of limiting it) in the hands of political officials and that this – regardless of their good intentions or not – tended to change the moral character of citizens towards submissiveness and reduced their sphere of free choice and individual planning (the fundamental goal of liberalism), sounded compelling. Following this lead, the ideas expounded in “The Road to Serfdom” would eventually enjoy tremendous success in the US and even surpass academic circles, to a large extent due to a popular and very shortened version of his book that was sold with *Reader’s Digest*. The review by Keynes – the most respected economist of the period – also proved important to this success for, despite being critical in many respects, it praised Hayek’s moral commitments, and it was this part that was used for the promotion of the book.

Despite having obtained British nationality in 1938, and possibly due to a certain dissatisfaction with the intellectual and political climate of the immediate post-war years, during which the Labour Party won the general elections and was able to go forward with its nationalization policies (Meijer, p. 690), he decided to leave the UK in 1950 to go to the University of Chicago. As the reader is most likely aware, this institution would later acquire the reputation of being a bastion of neoliberal economic thought and a training ground for a number of Latin American economists and political leaders (the so-called “Chicago Boys”) that, on their return, effected a number of radical reforms in their countries, particularly in Chile after Pinochet’s coup (1973) and in Argentina.¹⁶ This generation of Chicago apprentices was also responsible for the “Washington consensus” that would only start to be abandoned with the emergence of a wave of left wing governments in the late 90’s and early 2000’s. This “American” period, thanks to the ideologically friendlier atmosphere Hayek would encounter there, allowed him to research and publish his more mature and praised works. First came *The Constitution of Liberty* in 1960 and later the trilogy *Law, Legislation and Liberty*, starting with *Volume I* in 1973 entitled *Rules and Order*, continuing in

¹⁶ Incidentally, Hayek would show his unequivocal approval of the economic policies of Pinochet’s government in an interview to the Chilean rightwing newspaper, *El Mercurio*. (Cf. Caldwell and Montes, 2015)

1976 with *Volume II. The Mirage of Social Justice* and finishing with *Volume III. The Political Order of a Free People* in 1979. In these later works, he rejects the more moderate position he had taken in an earlier period, namely in *The Road to Serfdom*. In 1974 he was awarded the Nobel Prize along with Gunnar Myrdal (ironically, a staunch and traditional Scandinavian social democrat). Later in his life, Hayek moved to Freiburg, Germany, where he remained until his death in 1992.

A bird's eye view of this special issue

The contribution of José Colen, António Baião and Pedro Moreira Góis (“The liberal creed and its critics: pauperism and the safety net”) in this special issue explores the views of both authors on the issue of poverty and economic security under a market society. They compare and contrast Hayek’s and Polanyi’s distinctive and opposed perspectives on a series of intellectual debates. The emergence of market society, the historical and conceptual relationship between the market and freedom, the emergence of fascism and totalitarianism in the 20’s and 30’s and, finally, the hegemony of economic liberalism in the preceding decades were all topics in which both men left their mark. Both authors shared the goal of rescuing the “hearts and minds” of citizens from dangerous ideas, which allows one to say that, in a sense, the “enemy” was for them internal rather than external. It was the product of an intellectual error or deceit that thrived in the political and economic climate of Western societies, more than a hostile nation, be it Nazi Germany or the Soviet Union. While Polanyi forcefully argued that the self-regulating market was not only a utopia but a dangerous one, on the basis that the attempts to establish it had so dire and unbearable consequences that they would inevitably result in an unplanned, spontaneous reaction of society in the form of a backlash, Hayek in turn maintained that liberalism and the market economy were not to be blamed for the emergence of fascism. On the contrary, Hayek contended that nazi-fascism was the twin-brother of socialism, both resulting from a similar collectivist prejudice and impatience towards the market and its “unplanned” results, while feeding off a misinterpretation of the liberal individualism that at once supports and is fostered by the market institutions. Furthermore, it seemed to Hayek that both

ideological outlooks demonstrated a lack of understanding or undue disregard for the market's unparalleled ability to rationally convey and synthesize the best available information about economic preferences of agents. Finally, in spite of the critical view of *laissez faire* ideas – unduly conflated with liberalism according to Hayek – which he still held in 1944, he believed that with proper policies and moderate state action, based on the rule of law (general laws instead of direct, arbitrary orders or instructions by officials) and the proper working of competition, the market (or capitalism) was able to provide for the needs of people better than bureaucratic institutions – no matter how well intentioned its leaders – without simultaneously encroaching on their fundamental freedoms or planting the seeds for their mental enslavement *vis-à-vis* concentrated state power. This optimism was in deep contrast with the bleak picture Polanyi drew of the history of unfettered capitalism, namely as exemplified in Great Britain after the end of the Speenhamland System and the reform of the Poor Laws in 1834.

A contentious point is how to interpret Hayek's later radical rejection of the "safety net" which he had thought both reasonable and important in his earlier writings. While he himself would see his early defense of the need for that safety net and related criticisms of classic or *laissez faire* capitalism of the 19th century as the mere result of concessions he felt pressured to acquiesce to under the somewhat hostile ideological environment of the time, the authors of this article claim that the issue is probably more complex than he himself thought. On the one hand, it does look like a concession, as Hayek was perfectly aware that, after going through a period of massive unemployment and structural insecurity, it would have been tactically unwise to attempt a defense of "unplanned" capitalism based on the mere hope that the supposedly higher productivity of the free market would eventually, in the long run, provide more prosperity to all (the "a rising tide lifts all boats" argument). On the other hand, defending the legitimacy of the state in promoting a minimal safety net as a "middle path" was not entirely inconsistent with his long standing distinction between legitimate state action, in which the state acts and regulates through law in a way that may enhance people's abilities to develop the best plans for themselves and act in accordance with them, and illegitimate ("totalitarian") state action in which the state's global plans for society replace or impede people's pursuit of their own individual or local plans.

Given that, in Hayek's perspective, "global" (or society-wide) planning requires either an impossible consensus on values and goals (beyond mere rhetorical commitment to a vague notion of equality), or coercion against the dissident minority and the restriction of their capacity for individual planning, it was necessary for any person genuinely concerned with freedom to abandon any sort of state action that consisted of "global planning" in this sense. While one may question if a safety net as suggested by Hayek in his earlier writings belongs to this category of "global planning" or not, a safety net must, of necessity, according to these authors, distort the workings of the market and its fundamental role in revealing unknown preferences. This, in their mind, could be what ultimately led Hayek to give up altogether the task of solving the problem of pauperism in his later days.

Regardless of the clear ideological and philosophical differences between Hayek and Polanyi, there is also evidence of some methodological similarities from a metatheoretical standpoint. In his contribution, Filipe Faria ("Double Movement in Polanyi and Hayek: Towards the Continuation of Life") argues that the crucial notion of "double movement" in Polanyi finds its counterpart in similar considerations regarding the existence of opposing and contradictory social tendencies in Hayek's evolutionary theory of market society. But while both thinkers seem to share, in broad terms, remarkably similar views about the processes through which a society adapts itself to the challenge of the nascent market sphere, their point of departure is given by the normative assessment of those movements: what Polanyi sees as the self-protective tendency of a society facing the perils of the invasive and heteronomous marketplace, Hayek identifies a mere "tribal" attitude toward what should be otherwise seen as a benevolent market force.

At the bottom of this fault line lie not only their differences regarding the proper characterization of the effects of an expanded market sphere, but also opposing perspectives on what constitutes a sufficient, stable and thriving social life. On this level the split between Hayek and Polanyi runs deep: even if Polanyi seems at times to presuppose a certain moral superiority of the small scale community over the enlarged society, which allows him to add a normative charge to his general criticism of the encroachment of market interactions in the life of

the community, Hayek's comparisons of the "tribal" society with the enlarged society entailed by market liberalism are grounded on a seemingly larger and more demanding conception of "social evolution", whereby certain social forms forced other, more archaic social complexes, into extinction. However, in this respect, Faria makes a compelling case for the somewhat paradoxical presence, in both thinkers, of something akin to a "double movement" with an emphasis on the evolutionary dynamics of social change, while simultaneously exposing the shortcomings of Hayek's criticisms of the insufficiency of the small-scale community life and a certain over-reliance on the superiority of individualistic (or atomistic) forms of social life.

Part of the reason why works like *The Great Transformation* and *The Road to Serfdom* still garner so much attention stems from the perception that the specific social issues and debates they tried to tackle still exist today, albeit in slightly different ways. As many observers have stressed, some of the economic and political problems which affected most advanced and integrated capitalist economies over the previous decade seem remarkably similar to the path of progresses and setbacks that so gravely afflicted Western societies in the inter-war period: comparisons between the effects of the gold standard with the constraints imposed by the European monetary union, or the existence of coincidental aspects between the rise of today's populist movements with past political movements, are just two examples that lend credibility to attempts to actualize the message of these two classic works. Patrícia Fernandes's article ("Myth, utopia, and democracy in Karl Polanyi's *The Great Transformation*") fits neatly into the contemporary task of reviving the insights of past classics, by making a strong case for future-oriented political reforms that take into account the democratic deficits of current, overly market-based, societies.

While not hiding a certain allegiance to a Polanyian reading of our current malaise, Fernandes argues that an informed critique of the latest social misdevelopments – e.g. the subjection of Portugal to a severe form of economic austerity – must rediscover the importance of democratic control over heteronomous and illegitimate market interests, as well as the intrinsic worth of a democratic reflection rooted in an expanded and multidisciplinary take on the contributions produced by a variety of fields of knowledge. Against those who

implicitly or unconsciously assert that market interactions and their ancillary institutions must be regarded as a natural feature of our world, whereby its results necessarily attain the status of 'realistic' as opposed to 'unrealistic' or utopian perspectives, Fernandes makes a case for a re-energized notion of democracy which strips the market sphere from its pretension of naturality and reclaims for it a holistic reflection on the means to achieve democratic control over nature proper, man-made spheres of economic interaction and the permanent limitation of politics as a succession of contingent political projects.

How far the democratic control over the realm of economic interactions should go, and which type of legal and political instruments should be used in order to exert that control, depends on the shape that the political should assume. Here one is forced to venture into the histories of several traditions of thought so as to elicit the different responses to this difficult question. Bru Laín's contribution in this special issue ("Polanyi's Economic Embeddedness, Countermovement, and Republican Political Economy") focuses on the concurrence between central Polanyian notions such as the "double movement" or "economic embeddedness" and the solutions envisaged within the republican tradition of thought so as to rein in the outcomes of the marketplace. According to Laín, although Polanyi does not seem willing to endorse republican views on a number of issues, both Polanyi's main criticisms of market expansion and the republican attempts to keep "self-regulating" markets under control draw from a common concern for the "the material conditions for freedom". Here, the specific manner in which the republican tradition (abstracted from its varied historical and intellectual manifestations) faced the problem of the wildly asymmetric distribution of material goods plays an important role.

In fact, regardless of whether the realization of this shared concern entails that Polanyi's remedies for the dismaying state of wealth and income distribution would necessarily assume the specific contour of, for example, a "property-owning democracy", and therefore presuppose that the solution relies on a reinterpretation of a specific type of property as prerequisite for real freedom and political participation, both the Polanyian type of social analysis and the guiding thread of the republican tradition reject a watered-down conception of a merely "formal" economic realm. On the contrary, the dangers that such a 'free-standing'

economic sphere impose on the life of the community, on the one hand, and on the self-rule of freely associated and materially independent citizens, on the other, can only be adequately countered if these citizens finally come to terms with the underlying political features of the economic institutions and with their “substantive” structure, as opposed to their merely “formal” character. For contemporary observers, however, the corollary of this captivating discussion is the underlying call for a return of political economy as a self-standing field of investigation and debate, wherein a range of alternative “political economies” – e.g. of socialism and republicanism – is again thrust into the limelight, thereby replacing an ossified mainstream economic science left to its own devices for far too long.

As we have evidenced in our cursory glance at the lives of Hayek and Polanyi, and in keeping with the intellectual and academic practices of the day, both thinkers worked on topics that intersected several autonomous disciplines and often put forward complex arguments which some today would hesitate to fit into the dominating internal practices and language of economics or political theory. If anything, many of these arguments and debates would today fit into the field of political economy, the return of which the authors of this introduction would naturally stand for.¹⁷ Now, this political economy finds its highest expression when it unfolds under an internal logic that does not make too many concessions on the supposed autonomy of either purely “economic” considerations – e.g. what economic “efficiency” or “productivity” mean – or squarely “political” conceptions – e.g. “autonomy” or “freedom”.

An aspect that stands out in this special issue is the surprisingly small amount of direct engagement between Polanyi and Hayek, who refer to each other’s work only a handful of times. Against the backdrop of this seeming lack of dialogue, João Rodrigues (“The debate in the great transformation: Ludwig von Mises, Friedrich Hayek e Karl Polanyi on capitalism, socialism and markets”) critically assesses their theoretical development in the shadow cast by Von Mises. Taking as a point of departure a revisitation of the debate regarding the nature and embodiment of the markets, in which Polanyi and Hayek took opposite

¹⁷ We do not intend to claim here that the field of political economy is wholly absent; that would amount to an irresponsible and unwarranted position, given the many great political economists who labour to revive this field. Our use of the word ‘return’ here is meant as an appeal to the restatement of the primacy of political economy vis-à-vis ‘mainstream’ economics.

stances partly as a consequence of their own reading of the criticisms laid out by Ludwig von Mises in his influential take on the shortcomings of the marketless form of socialism (1920), Rodrigues presents in great detail the evolution of Hayek and Polanyi's positions not only on the issue of economic calculation within a socialist economy – which, with respect to the existence of markets, could assume different forms –, but also on the most accurate description of liberalism as a political-economic project and on the prospects of a reconceptualised idea of socialism, freed from ailing forms of economic centralization.

Beneath these theoretical positions lies, however, the fact that for both men a proper description of the inner workings of each type of social organization must rely on the identification of the political underpinnings of such organization – for example, whether individuals regard others in an instrumental way in the political or economic spheres, or whether the interaction between them could be made dependent on a specific form of mutual recognition, and therefore constrained by what that recognition could allow. At this level, the debate had already moved beyond the thorny issue of economic calculation and coordination. Rodrigues depicts this manifold space of argumentation which required both Hayek and Polanyi to subscribe specific positions regarding the role and necessity of social institutions, the drivers of social change and, particularly for Polanyi, the problem of transition from a capitalist regime.

Despite the intricacies of these characterizations, one is advised not to forget that each man's political leanings – taken here holistically, and not merely on the level of real politics – form the basis for their theoretical clashes: Polanyi's attack on the supposed 'spontaneous' nature of the markets, or Hayek's critique of socialism based on its inadequate means to convey information about individual preferences, could result from each author's incapacity to analyse the best and most coherent version of their theoretical enemy. On this, however, Rodrigues portrays Polanyi in a more favourable light, as he offers us a nuanced discussion of Polanyi's perspectives on the topic of the economic organization of socialism: as early as the mid-20s, Von Mises' indictment against centralized forms of economic planning seemingly had found an echo in Polanyi, whose rejection of a planned economy in favour of associational or guild-inspired forms of socialism

put him at odds with all socialists for whom centralized economic planning was the only feasible and worked out socialist political economy. Here, as in other places, one struggles to find much cause for disagreement between Hayek and Polanyi, despite what superficial readings might suggest.

Ultimately, what set both Austrian-born thinkers apart is what still divides neoliberalism and forms of socialism, which accept some form of markets, today: in which form and under which type of democratic control or oversight should the latter exist. And there is no better way to show our recognition for Hayek and Polanyi's work than by reinstating the ongoing, open-ended nature of these debates.

References

Bartels, Larry M., Hecl, Hugh, Hero, Rodney E. & Jacobs, Lawrence R. (2005). *Inequality and American Governance*. In Lawrence R. Jacobs & Theda Skocpol (eds.) *Inequality and American Democracy, What We Know and What We Need to Know* (pp. 88-155). New York, NY: Russell Sage Foundation.

Bevir, Mark (2005), *New Labour, a Critique*. London, UK: Routledge.

Blair, T., & Schröder, G. (2000 [1998]). *Europe: The Third Way/Die Neue Mitte*. In B. Hombach (ed.), *The Politics of the New Centre* (pp. 157-177). Cambridge, M.A: Polity Press.

Block, Fred & Somers, Margaret (2014), *The Power of Market Fundamentalism, Karl Polanyi's Critique*. Cambridge, Mass: Harvard University Press.

Bockman, J., Fischer, A. & Woodruff, D. (2016), "Socialist Accounting" by Karl Polanyi: with preface "Socialism and the embedded economy", *Theory and Society*, Vol. 45, issue 5, pp. 385-427.

Caldwell, Bruce. (2007). Hayek and the Austrian School. In Edward Feser (ed.) *The Cambridge Companion to Hayek* (pp. 13-33), Cambridge, UK: Cambridge University Press.

Caldwell, Bruce & Montes, Leonidas (2015). Friedrich Hayek and his visits to Chile. *Review of Austrian Economics*, 28 (3), 261-309.

Cangiani, Michele. (2009). The unknown Karl Polanyi. *International Review of Sociology: Revue Internationale de Sociologie*, Vol 19, number 2, pp. 367-375.

Cangiani, Michele. (2011). "A Staggering World": Karl Polanyi's Contribution to Der Österreichische Volkswirt. in Jürgen Backhaus (ed.), *The Beginnings of Scholarly Economic Journalism, The Austrian Economist and The German Economist* (pp. 1-14). New York, NY: Springer.

Clinton, Bill & Gore, Al. (1992). *Putting People First*. New York, NY: Times Books.

Corrigan, Paul, Jones, Trevor, Loyd, John & Young, Jock. (1988). *Socialism, Merit and Efficiency. Fabian Tract No 580*, London, UK: The Fabian Society.

Crouch, Colin. (2009). Marketisation. In Matthew Flinders et al. (eds.) *Oxford Handbook of British Politics* (pp.879-895). Oxford, UK: Oxford University Press.

Dale, Gareth (2010), *Karl Polany: the Limits of the Market*, Cambridge, UK: Polity Press.

Dale, Gareth. (2014). The iron law of democratic socialism: British and Austrian influences on the young Karl Polanyi. *Economy and Society*, Vol. 43, number 4, 650-667.

Gakis, Dimitris. (2015). Wittgenstein, Marx, and Marxism: Some Historical Connections. *Humanities*, vol. 4, 924-937.

Ginsberg, Benjamin. (1986). *The Captive Public, How Mass Opinion Promotes State Power*, New York, NY: Basic Books.

Gullick, Charles A. (1948). *Austria, From Habsburg to Hitler. Vol. II*, Berkeley, CA: University of California Press.

Heath, Anthony; Jowell, Roger, & Curtice, John. (2001). *The Rise of New Labour, Party Policies and Voter Choices*. Oxford, UK. Oxford University Press.

Jacobs, Lawrence R. and Skocpol, Theda (2005). *American Democracy in an Era of Rising Inequality*. In Lawrence R. Jacobs & Theda Skocpol (eds.) *Inequality and American Democracy, What We Know and What We Need to Know* (pp.1-18). New York, NY: Russell Sage Foundation.

Kresge, Stephen & Wenar, Leif. (2005). *Hayek on Hayek, an Autobiographical Dialogue*. London, UK: Routledge.

Le Grand, Julian. (2003). *Motivation, Agency, and Public Policy. Of Knights and Knaves, Pawns and Queens*. Oxford, UK: Oxford University Press.

Maucorant, Jérôme. (2007). Karl Polanyi, une biographie intellectuelle. *la Revue du Mauss*, 1, vol. 29, 35-62.

Meijer, Gerrit. (2012). Friedrich August Hayek (1899–1992). In Jürgen Backhaus (ed.) *Handbook of the History of Economic Thought* (pp. 689-711). New York, NY: Springer.

Mises, L. von (1935). Economic Calculation in the Socialist Commonwealth. In F. Hayek (Ed.), *Collectivist Economic Planning* (pp. 87-130). New York, NY: Augustus M. Kelley.

Piketty, Thomas. (2014). *Capital in the Twenty-First Century*. Cambridge, Mass: The Belknap Press.

Polanyi, Karl. (1947). Our Obsolete Market Mentality. *Commentary*, 3, 109-117.

Rabinbach, Anson. (1983). *The Crisis of Austrian Socialism, From Red Vienna to Civil War*. Chicago, Il: The University of Chicago Press.

Reisch, George. (2005). *How the Cold War Transformed Philosophy of Science: to the Icy Slopes of Logic*. Cambridge, UK: Cambridge University Press.

Rodrik, Dani. (2011). *The Globalization Paradox, Why Global Markets, States, and Democracy Can't Coexist*. Oxford, UK: Oxford University Press.

Rogan, Tim. (2013). Polanyi at the Margins of English Socialism, 1934-1947. *Modern Intellectual History*, Volume 10, issue 2, 317-346.

Stiglitz, Joseph. (2002). *Globalization and its Discontents*. New York, NY: WW. Norton & Company.

Stiglitz, Joseph. (2003). *The Roaring Nineties, Seeds of Destruction*. London, UK: Allen Lane.

Stiglitz, Joseph (2016). *The Euro: And its Threat to the Future of Europe*. London, UK: Allen Lane

Streeck, Wolfgang. (2014). *Buying Time, The Delayed Crisis of Democratic Capitalism*. London, UK: Verso.

Thatcher, Margaret. (1993). *The Downing Street Years*. London, UK: Harper Collins Publishers.

Zucman, Gabriel. (2015). *The Hidden Wealth of Nations*. Chicago, Il: The University of Chicago Press.



O CREDO LIBERAL E OS SEUS CRÍTICOS: PAUPERISMO E REDE DE SEGURANÇA

LIBERAL CREED AND ITS CRITICS: PAUPERISM AND SAFETY NET

António Baião*
antoniosaraivabaião@gmail.com

J. A. Colen**
jacolen64@gmail.com

Pedro Moreira***
pedrogoismoreira@gmail.com

Sumário. *Em 1944, duas obras marcantes da teoria política, A Grande Transformação e O Caminho da Servidão, apresentam visões quase opostas da relação entre o nascimento do “mercado” e os fascismos: no primeiro caso, o mercado provoca o fascismo como reacção de defesa da sociedade, condenada ao pauperismo, contra o credo liberal; no outro, o mercado é a melhor, senão única defesa da liberdade contra o regresso das tendências totalitárias que asfixiam a liberdade, agora sob a forma da planificação. O texto examina os argumentos dos dois autores, procurando elucidar e discutir as concessões de Hayek à intervenção do estado e explicar porque, posteriormente, ele se inclina a abandonar a proposta de uma rede de segurança que defenderia as vítimas das falhas do mercado.*

Palavras-chave: *Karl Polanyi, Friedrich Hayek, rede de segurança, pauperismo, liberalismo.*

Abstract. *In 1944, two remarkable books on theories of political economy, The Great Transformation and The Road to Serfdom, present opposite visions of the relationship between the birth of the “market” and the origin of fascisms: according to the former, the market generates fascism as society’s (unplanned and unforeseeable) self-defense response to increasing pauperism and against the liberal creed; the latter holds that the market is the best or even the necessary means to protect liberty against the return of totalitarianisms, now in the guise of “planning”. This paper examines the arguments of both authors, seeking to elucidate and debate Hayek’s (supposed) concessions to state intervention. It also*

* Doutorando no CEPS, Universidade do Minho.

** Investigador Integrado no CEPS, Universidade do Minho e professor convidado no IEP (UCP).

*** Doutorando no IEP, UCP.

throws light on the question of why, later on, he felt inclined to abandon entirely the proposal of a safety net to defend the victims of market failures.

Keywords: Karl Polanyi, Friedrich Hayek, safety net, pauperism, liberal creed.

o. Encontros e desencontros em 1944

Duas obras marcantes da teoria política, *The Great Transformation* e *The Road to Serfdom*, foram publicadas em 1944, no dealbar, incerto ainda, do fim da guerra. Curiosamente, para além do momento histórico, os autores destas duas obras hoje clássicas partilham mais de um traço biográfico. Ambos os pensadores nasceram no Império Austro-Húngaro em vias de se liberalizar, viveram em Viena ainda em ambiente de *fin de siècle*¹ e até trabalharam ligados à Escola Austríaca de Economia, em lados opostos da rua.

Com uma diferença de idade de quase quinze anos, Karl Polanyi nascido em 1886 e Friedrich August von Hayek em 1899, tiveram naturalmente percursos pessoais e intelectuais diferentes, que decorrem em parte dos dilemas que atravessaram o contexto da Viena daquela altura. Já em 1908 e com apenas 22 anos, Polanyi foi muito activo durante a sua vida estudantil em Budapeste e fundou o Círculo Galileu, um grupo marcadamente progressista que chegou a ter dois mil membros e incluiu figuras como Georg Lukács e Karl Mannheim.² Também Hayek, com a mesma idade em 1921, mas anos depois, tinha acabado o seu primeiro curso de direito em Viena e fundou um pequeno círculo de leitura, o *Geistkreis* (algo como “os irmãos em alma”), que chegará a incluir figuras como Alfred Schütz, Eric Voegelin e Felix Kaufmann.³

O desboroar do Império e a violência da Grande Guerra afecta ambos, ainda que de modo diverso. Hayek, mais jovem, só foi alistado em 1917, mas mesmo assim volta transformado pela sua participação na guerra. Polanyi regressa da guerra a Viena em 1919 com uma depressão profunda que o marcará para o resto

¹ Sobre Viena no início do século, ver Schorske, C. E. (1992). *Fin-de-siècle Vienna: politics and culture*. Cambridge: Cambridge University Press, e Janik, A., & Toulmin, S. E. (1996). *Wittgensteins Vienna*. Chicago: Dee.

² Sobre Polanyi, ver Polanyi-Levitt, K., & Mendell, M. (1987). Karl Polanyi: His Life and Times. *Studies in Political Economy*, 22(1), 7-39, e Block, F. (2001), ‘Introduction,’ em Polanyi, K. (2001). *The Great Transformation: The Political and Economic Origins of Our Time*, Boston: Beacon Press, xviii-xxxviii.

³ Sobre Hayek, ver: Ebenstein, A. O. (2003). *Friedrich Hayek: A Biography*. Chicago: University of Chicago Press.

da vida. Um escritor prolífico, Polanyi escreveu entre 1924 e 1938 centenas de artigos (jornalísticos, académicos, notas) sobre tópicos diversos (política externa e interna, relações internacionais, economia...), chegando mesmo a publicar, em 1922, um artigo que se destacou, contrariando o livro *Socialism* (1922) de Ludwig von Mises e a participar no famoso debate sobre o “cálculo económico” no socialismo.

Por contraste, Hayek, depois de ter acabado um segundo curso em ciência política em 1923, decidiu passar o ano de 1923-24 nos Estados Unidos a trabalhar na Universidade de Nova Iorque. Apesar de ter sido muito recomendado pelo famoso economista austríaco Friedrich von Wieser, quando regressou, Hayek trabalhou em condições de relativa pobreza. De maior relevância na sua vida intelectual do que von Wieser, porém, revelar-se-ia a obra de von Mises, cuja descoberta, após a leitura de *Socialism*, Hayek dirá ter sido um dos eventos que mais influenciou a sua vida intelectual. Na verdade, esta obra só encontraria rival, segundo Hayek, nos *Principles of Economics* de Karl Menger. Se tinha sido vagamente socialista até então, o livro de Mises fez dele um liberal convicto e, a partir daí, Hayek abandonou as suas raízes fabianas. Nos anos seguintes, Hayek esteve sempre muito mais interessado na área da economia.

Enfim, o encontro de Polanyi com Mises, nos seminários deste último sobretudo, levaram o húngaro a montar um ataque contra o seu “liberalismo utópico”, um esforço intelectual que culminará mais tarde no *The Great Transformation*. Hayek, pelo contrário, mudou radicalmente de sinal e a partir da leitura de *Socialism* nunca deixou de se considerar um liberal.

1. Resgatar os “corações e mentes” do Ocidente

Os anos da Segunda Guerra e mesmo os anos entre as guerras foram enegrecidos não só pela sombra de Hitler, mas também pelos fascismos em geral. Durante este período ambos se opuseram ao que se veio a designar como “totalitarismo” nas suas diversas formas, incluindo a do comunismo soviético. Os autores de ambas as obras foram sobretudo mobilizados pelo combate ao que consideravam o principal mal do século. Polanyi e Hayek estão de acordo em ligar

estas formas novas e assustadoras de tirania política com a economia ou, melhor, a crise da economia.

Esta ligação pareceu-lhes talvez obscura e necessitada de elucidação, exigindo de ambos um esforço intelectual imenso, mas nem por isso menos certa, com reflexos sobre toda a vida social e a cultura. Como Polanyi afirma já desde 1942 e desenvolve em todo o capítulo 20 de *The Great Transformation*, não só a ligação é clara como se estende à arte, religião, filosofia, etc.

Tal como o socialismo, o fascismo deitava raízes numa sociedade de mercado que se recusava a funcionar. Daí que fosse mundial pelo horizonte, universal e internacional na sua acção: as suas consequências iam além da esfera económica, dando lugar a uma transformação geral com traços sociais próprios. Afirmou-se em quase todos os domínios da actividade humana – políticos e económicos, culturais, filosóficos, artísticos ou religiosos. E, até certo ponto, fundia-se com as tendências e os temas de cada um dos seus terrenos de acção. É impossível compreendermos a história da época se não distinguirmos entre o movimento fascista subjacente e as tendências efémeras com que esse movimento se fundiu em diferentes países.⁴

O fim da guerra não acalma os receios de nenhum destes. Hayek, na introdução de *The Road to Serfdom*, vê mesmo na situação do mundo anglo-saxónico do pós-guerra sinais preocupantes do pensamento alemão que originou a tendência totalitária:

Não é com a Alemanha de Hitler, com a Alemanha desta guerra, que este país se parece. Mas os estudiosos das ideias atuais não podem deixar de reparar que há mais do que uma mera semelhança superficial entre a tendência do pensamento na Alemanha durante e após a última guerra e as ideias atuais. Há hoje neste país certamente a mesma ideia segundo a qual a organização da nação que conseguimos para efeitos de defesa deve ser mantida para fins de criação.⁵

Esta preocupação comum com o que ambos sentem como uma profunda crise da sociedade e até da civilização, que parecia ter raízes na história das ideias, convoca os esforços tanto de Polanyi como de Hayek numa tentativa “desesperada” de resgatar os “corações e mentes” do Ocidente através de ambiciosos empreendimentos teóricos que contrariassem as tendências dominantes.

Em 1944, o espectro do Nazismo estava à beira de ser substituído pelo do comunismo. Apenas o horror da solução final impedia o esquecimento absoluto

⁴ Polanyi, K. (1944), *The Great Transformation*, 248, ver todo o cap. 20: History in the Gear of Social Change. Para todas as citações recorremos à tradução portuguesa: Polanyi, K. (2016). *A Grande Transformação: As Origens Políticas e Económicas do nosso Tempo*, Lisboa: Edições 70.

⁵ Hayek, F. A., & Caldwell, B. (2014). *The Road to Serfdom: Text and Documents: The Definitive Edition*. New York: Routledge, 3, 27 da tradução portuguesa. Todas as citações portuguesas são de: Hayek, F. A. (2008). *O Caminho para a Servidão*. Lisboa: Edições 70, embora pontualmente corrigidas, quando a tradução se afasta do original.

do fantasma dos fascismos, já aniquilados na Europa, horror que durante muito tempo ainda ensombrou o horizonte político. Com frequência crescente, depois do fim da guerra, pensadores liberais e não liberais por igual voltaram-se contra Marx, argumentando que a “ordem social perfeita” tentada pelo comunismo estava votada a revelar-se uma ilusão, apesar dos sucessos militares e económicos visíveis, apontado as semelhanças entre os “extremos.” A palavra “totalitarismo,” doravante popularizada, cobria ambos.

A União Soviética, que antes fora um aliado e contribuíra para o sucesso ocidental na guerra, está presa nas teias e cega pela teoria marxista que leva os soviéticos no pós-guerra a procurar assemelhar-se ao Ocidente, com qual partilha a meta final, diferente só nos meios, como um irmão mais impaciente.

Não se tem habitualmente em conta que os bolcheviques, embora fossem socialistas fervorosos, se recusavam obstinadamente a «construir o socialismo na Rússia». Bastavam as suas convicções marxistas para que pusessem de lado semelhante tentativa num país de base agrária e atrasado. Mas, deixando de parte o episódio excepcional do ‘comunismo de guerra’ (1920), os dirigentes adoptavam a posição de que a revolução mundial teria de começar a partir da Europa Ocidental. O socialismo num só país surgia-lhes como uma contradição nos termos, e quando se tornou uma realidade os velhos bolcheviques rejeitaram-na quase como um só homem.⁶

Hayek e Polanyi, que escrevem no despontar deste novo contexto intelectual, descobrem raízes muito diferentes para esta vaga. Mais que a dureza política das “tirantias” ou dos “sistemas”, vêem a raiz comum do mal num “erro intelectual” formidável que é preciso denunciar. Com efeito, em ambos os casos, a preocupação parece ser menos com os bárbaros às portas do império, o Exército Vermelho, que com o inimigo já no interior. No caso de Hayek, o feitiço pensamento alemão e, no caso de Polanyi, o feitiço da economia de mercado inglesa.

2. Uma Sociedade de mercado que se recusou a funcionar

O diagnóstico sobre a crise presente era, num sentido limitado, algo semelhante, apontando as culpas ao mau funcionamento do Estado e aos mercados, realidades que lhes pareciam indissociáveis. Hayek não desdenharia a sentença de Polanyi que culpava “uma sociedade de mercado que se recusou a

⁶ Polanyi, *The Great Transformation*, 255, cap. 20, p. da XXX da tradução portuguesa.)

funcionar”. Mas o tratamento do paciente que era recomendado pelos dois autores não poderia ser mais diferente.

Em *The Road to Serfdom*, Hayek aconselha os seus leitores, parafraseando Robert Frost, a seguir a via menos percorrida do risco da liberdade e a recusar os excessos do estado e recuperar o mercado. Em *The Great Transformation*, Polanyi, desde o início, adverte os seus leitores contra a crença de que o mercado não está ligado ao estado na gestação das tiranias modernas: “A nossa tese é que a ideia de um mercado capaz de se auto-ajustar era uma mera utopia. Semelhante instituição não poderia existir duradouramente sem aniquilar a substância natural e humana da sociedade; destruiria fisicamente o homem e transformaria o seu meio ambiente num deserto”.⁷ Os dois livros são escritos sob este pano de fundo de uma sociedade em risco, muito para além dos perigos da guerra.

O livro de Hayek torna-se rapidamente um sucesso comercial que torna o autor famoso,⁸ embora geralmente ostracizado como exemplo de liberalismo extremo. A versão mais conhecida é a versão abreviada de vinte páginas que a *Reader's Digest* vende nos Estados Unidos e que o autor apresenta em *tournee*. Hayek ganha fama de autor académico com veia popular porque foi esta última versão que o tornou decisivamente conhecido do público (o *Reader's Digest* distribuiu 600.000 cópias). Não há dúvidas que as circunstâncias contribuíram muito para o sucesso do livro: a guerra estava quase a acabar e a pergunta “*What's next?*” começou a despertar na mente de todos. Churchill leu *Road to Serfdom* e as suas ideias sobre a incompatibilidade do socialismo e de um regime livre saíram sem dúvidas reforçadas. Attlee contra-atacou afirmando que a posição de Churchill era “uma versão em segunda mão das posições académicas de um professor austríaco, Friedrich August von Hayek.”⁹

O livro de Polanyi, entretanto conferencista das “delegações externas” das Universidades de Oxford e Londres, nasce como um conjunto de palestras preparadas no essencial em 1939-40, e reduzidas a escrito depois, graças a uma bolsa da Fundação Rockefeller, e pronunciadas como um todo na sua versão

⁷ Polanyi, *The Great Transformation*, 3, cap. 1 da tradução portuguesa.

⁸ Kresge, S. (1994). *Hayek on Hayek: An Autobiographical Dialogue*, London: Routledge, p. 91. Hayek diz ao entrevistador que “Acho que já lhe tinha dito que não recebi um tostão pela versão abreviada do *Reader's Digest*.”

⁹ Ver cap. 16 e 17 de Ebenstein, *Friedrich Hayek*, mas também a introdução de Caldwell da *Definitive Edition*, 18-22, e o cap. 3 da entrevista de Hayek com Kresge, *Hayek on Hayek*.

quase final durante dois anos passados em Bennington College.¹⁰ O período sabático na América que lhe permite passar o livro a escrito viria a assegurar-lhe mais tarde um lugar em Columbia University (1947-1953). Se bem que influenciado na juventude por Karl Mannheim, o autor escreve, todavia, mais ao estilo da sociologia histórica de Max Weber e baptiza a sua própria teoria de “economia sociológica”.

Hayek sugere a necessidade de acarinhar o credo liberal, que, segundo Polanyi, tanto sob a forma do liberalismo keynesiano como de Von Mises, não sobreviverá à sua geração.

Apesar da clara oposição, ambos sempre rejeitaram explicitamente qualquer forma de doutrinário, radicalismo e revolucionarismo. Tanto Polanyi como Hayek, mesmo quando este último estava ainda na sua fase “fabiana”, mantiveram saudáveis dúvidas em relação ao marxismo.¹¹ E por mais entusiasmado que Hayek tivesse ficado com a obra de Mises, nunca acharam que o *laissez-faire* radical deste último fosse uma boa maneira de salvar o mercado, nem os fascismos a boa maneira de salvar a sociedade de mercado.

3. O teorema das cabras e cães na ilha de Robinson Crusóe: naturalismo e economia

Ambos partilham, todavia, por um breve momento, uma crítica ao “mercado” como resultado lógico das leis da natureza, autorregulável e só perturbado pela falta de liberdade.

A emergência gradual de uma classe sem um “superior feudal”¹² no séc. XVI torna-se um problema sério. Os Quakers, diz Polanyi, seriam os primeiros a reconhecer que “o desemprego involuntário era necessariamente o resultado de uma deficiente organização do trabalho”¹³ e a apontar soluções. John Bellers é

¹⁰ “Karl Polanyi: Five Lectures on The Present Age of Transformation-Lecture Series Listing of Topics”. Bennington College. Bennington College. <https://crossetlibrary.dspace.org/handle/11209/8502>

Consultado em 11 de Junho de 2017.

¹¹ Polanyi-Levitt e Mendell, *Karl Polanyi*, 18; e Ebenstein, *Friedrich Hayek*, 23.

¹² Polanyi, *The Great Transformation*, 109.

¹³ Polanyi, *The Great Transformation*, 110.

exemplar ao sugerir a organização da força de trabalho em corporações ou “Colleges”. Polanyi chega a dizer que esta proposta está no coração do socialismo utópico — das aldeias de Owen até às comunas dos planos quinquenais de Estaline.

No seu retrato do problema da pobreza, Polanyi passeia-se entre os séculos, desde Thomas More até à época de Jeremy Bentham, para compreender a evolução da noção de pauperismo. Um panfleto de 1704 de Daniel Defoe e a fábula das abelhas de Mandeville convencem os mais avessos do mal necessário da pobreza, mas o ponto de viragem crítica dá-se com o sistema de Speenhamland, que modela como nenhum outro “o destino de uma civilização”.¹⁴

Expliquemos com mais detalhe a releitura alternativa que Polanyi faz da revolução industrial e da miséria *dickensiana* que a acompanhou à medida que o mundo tradicional ruía. Com efeito, em especial nos dois capítulos que encerram a secção “*Satanic Mill*” da segunda parte da obra *The Great Transformation*, o autor conclui a tarefa “arqueológica” de compreender qual a origem da economia de mercado e a sua relação com a noção de pauperismo. Os textos “*Pauperism and Utopia*” e “*Political Economy and the Discovery of Society*” avançam com a tese de que, durante a era moderna, surgiram um conjunto de equívocos sobre o fenómeno da pobreza e da miséria, da sua relação com a moral, com o funcionamento da comunidade e a subordinação a leis que escapavam ao controlo do homem e da criação secular do Estado. O que revela o exame da Inglaterra do século XVIII, onde o Iluminismo se propaga, era a evidência quase-paradoxal de que pauperismo e progresso eram conceitos inseparáveis. Se a religião do progresso se inscrevia no próprio tecido ideológico de um novo regime, quais as razões para que houvesse mais pobres na civilizada Inglaterra que nas nações bárbaras?

A pergunta de partida não é de todo despropositada, principalmente quando o conceito de progresso influenciou a maior parte dos enquadramentos filosóficos anti-traditionalistas que se tornaram hegemónicos no pensamento filosófico ocidental após a Revolução Francesa. *Do Esquisse d'un tableau historique des progrès de l'esprit humain*, de Condorcet, ao *Code de la Nature*, de Morelly, e aos

¹⁴ Polanyi, *The Great Transformation*, XXX e todo o cap. 8.

socialismos utópicos oitocentistas, liberais e socialistas concordavam com a tese de que o desenvolvimento tecnológico e económico produziria um crescimento dos níveis de bem-estar geral. A realidade, contudo, demonstrou que a situação não seria assim tão facilmente resolvida. Por exemplo, Adam Smith declarou que nem sempre os salários mais altos se encontravam nos países mais ricos e não pode ser ignorado o facto do século XVII ter fomentado a tensão (quando não a agravou) entre máquina e Homem. A evolução industrial, ao substituir a força humana pela máquina, abriu as portas para a pauperização em larga escala de uma sociedade que estava já na transição para o capitalismo, mas que não tinha ainda lugar para um mercado laboral.

Nesta comunidade, os pobres já não eram os invasores bárbaros do século XVI, forças exteriores à organização social, exércitos hostis que corrompem e degeneram. O pauperismo tornava-se uma dimensão social a ter em conta, até porque foi compreendido que este fenómeno era provocado por uma deficiência na própria organização social do trabalho. Se assim era, qual o papel que a autoridade pública devia desempenhar para corrigir esta falha e minimizar os efeitos deste problema? Esta pergunta, aparentemente inofensiva e quase demasiado evidente, não o era nos séculos XVII e XVIII, pois –segundo Polanyi– foram geralmente reconhecidas “vantagens” económicas deste estado de miséria social, desde logo através do trabalho barato e da facilidade em recrutar indigentes para os campos de guerra. Mesmo os cidadãos que beneficiavam do sistema proto-assistencialista das *Poor Laws* podiam ter algo a oferecer à comunidade. Na concepção de *laissez-faire* de Edmund Burke, por exemplo, o trabalho tornava-se uma mercadoria que só podia encontrar o seu preço no mercado e, necessariamente, os pobres deviam ser retirados da chancela assistencialista do Estado para fazer lucrar os interesses privados. Para Jeremy Bentham, tratava-se de promover a utilidade pública, como é facilmente entendido através do plano de engenharia social utilitária que era o *Panopticon*, no qual os pobres assistidos seriam utilizados para os trabalhos mais duros, compensando a sociedade pela ajuda que lhes era prestada. Mas Bentham tinha perante si uma realidade, na qual se revela a tendência para o agravamento do

pauperismo, como ficaria provado no final das primeiras décadas do século XIX.¹⁵ O dramático aumento da miséria social impossibilitava que se negligenciasse a reflexão em torno da génese desse problema. Qual seria, afinal, a razão do pauperismo e como poderia este ser condicionado?

Não seria certamente pelo sistema Speenhamland, aplicado a partir do final do século XVIII, que era obsoleto desde a sua origem e não passava de uma organização da pobreza universal. Os trabalhadores continuavam a receber salários baixos e mantinham-se no limiar da sobrevivência graças ao auxílio garantido pelas paróquias. Além disso, a economia de mercado apresentava os seus primeiros desenvolvimentos modernos, numa combinação de capitalismo sem mercado laboral. O problema continuava a existir de forma estrutural e não era atacado na sua raiz: quebrar os mecanismos assistencialistas existentes e edificar um mercado do trabalho. De facto, David Ricardo terá compreendido que o problema da pobreza se devia ao facto de a organização económica estar subjugada a leis que não são humanas e que, por essa razão, não podiam ser aplicadas para a restrição da pobreza.¹⁶

¹⁵ Polanyi, *The Great Transformation*, 115 (cap. 9). Os planos de Bentham visavam em primeiro lugar a reforma do sistema prisional, mas seguimos aqui a polémica de leitura Polanyi.

A questão é secundária para o propósito deste texto, mas veja-se sobre o assunto, por exemplo, Semple, J., *Bentham's Prison: A Study of the Panopticon Penitentiary*, (Oxford: Clarendon Press 1993), bem como Bahmueller, C. F., *The National Charity Company: Jeremy Bentham's Silent Revolution* (Berkeley, CA: University of California Press, 1981).

¹⁶ Mais tarde, Ricardo evoluirá num sentido diferente, contemplando a possibilidade de não só a Natureza, como as próprias decisões sociais interferirem na definição do valor do trabalho e, por essa razão, serem leis económicas de validade similar. Polanyi tem certamente em consideração a peculiar evolução do pensamento de Ricardo, quando refere que o economista se afastava de um naturalismo rígido. Na verdade, o que está em causa é um distanciamento entre a perspectiva expressa nas duas primeiras edições de *On the Principles of Political Economy and Taxation* (1817 e 1819), que se distancia da revelada na terceira edição (1821), devido à incorporação do capítulo "On Machinery". Nas duas primeiras edições, Ricardo suporta a sua teoria na "(i) lei biológica da população de Malthus, (ii) lei físico-química da diminuição da fertilidade do solo e (iii) na concepção Smithiana da economia como um processo naturalmente harmonioso" (John P. Henderson e John B. Davis (aut.), Warren J. Samuels e Gilbert B. Davis (ed.), *The Life and Economics of David Ricardo*, New York: Springer Science and Business Media, 1997, p. 587). Na terceira edição, Ricardo incorpora o capítulo "On Machinery", que contribui para uma reflexão e questionamento em torno das teses naturalistas expressas anteriormente. Ao argumentar que o processo de mecanização pode ser prejudicial para a classe trabalhadora (aumento do desemprego e diminuição dos salários), Ricardo põe em causa as teses naturalistas, mas não as afasta totalmente: enquanto que a segunda pode ser revertida através do contributo do desenvolvimento tecnológico, através da aplicação de processos tecnológicos que revertam os problemas da fertilidade dos solos, a primeira tese revela uma evolução no pensamento de Ricardo bem mais significativa. O carácter biológico da tese de Malthus, que implicava que a divergência entre mercado e preço natural do trabalho seria corrigida pela alteração dos ritmos de morte e reprodução (ibidem), é questionado pela adição de uma nova variável que torna redundante a existência de população: a maquinaria também iria contribuir para a flutuação do mercado laboral. Durante este período de reelaboração das teses definidas nas duas primeiras edições de *Principles*, Ricardo inicia um processo de reflexão em torno da possibilidade de decisões sociais poderem interferir com a oferta de trabalho, particularmente através da utilização de receita poupada (*saved revenue*) para a minimização do

Essas leis não-humanas estariam na base da economia naturalista desenvolvida por Joseph Townsend no tratado *A Dissertation on the Poor Laws* (1786), na qual é desenvolvida uma tese proto-darwinista social e competitiva e anti-assistencialista, pela simples razão de que o Homem estaria a agir contras as normas impostas pela Natureza.

A mais crua apresentação do “naturalismo” da teoria económica que Polanyi combate, aparece na *The Great Transformation* sob a forma do equilíbrio entre cabras e cães deixados numa ilha deserta:

A Dissertation de Townsend, dez anos mais tarde, centrava-se no teorema das cabras e dos cães. O cenário é a ilha de Robinson Crusoe no oceano Pacífico, ao largo da costa do Chile. Nessa ilha, Juan Fernandez desembarcou algumas cabras que lhe forneceriam carne se lhe acontecesse tornar a visitar a ilha no futuro. As cabras multiplicaram-se com uma fertilidade bíblica, transformando-se numa reserva alimentar vantajosa para os corsários, sobretudo ingleses, que assediavam os navios mercantes espanhóis. Para combater as cabras, as autoridades espanholas desembarcaram na ilha um cão e uma cadela que, também eles, se multiplicaram profusamente, fazendo diminuir o número das cabras, alimento dos cães. ‘Restabeleceu-se então uma nova forma de equilíbrio’, escrevia Townsend. ‘Os animais mais fracos de ambas as espécies foram os primeiros a pagar a sua dívida à natureza; os mais ativos e vigorosos preservaram a sua vida’. E acrescentava: ‘É a quantidade de alimento que regula o número da espécie humana’.¹⁷

O próprio governo, aliás, não seria necessário para o funcionamento da economia, como fica demonstrado pela sua alegoria do equilíbrio garantido pelas dificuldades dos cães em comerem as cabras que se refugiaram nas zonas montanhosas para alcançar a segurança. Projectando esta narrativa para a comunidade dos homens, demonstrar-se-ia que o equilíbrio económico era garantido pela escassez e pela fome. Assim, tal como o equilíbrio entre cães e cabras era garantido pela escassez de comida, se só a fome poderia domar os animais selvagens, também só a fome iria impelir os pobres a trabalhar. O determinismo zoológico de Townsend definia, portanto, que o assistencialismo devia ser abolido. A “fome” teria sido implacavelmente colocada entre as

desemprego (idem, p. 588). Neste sentido, a “sociedade juntou-se à Natureza para determinar a forma como as leis da economia política poderiam operar, acabando com o seu *status* de combinações puramente naturais” (ibidem). Inevitavelmente, a terceira tese foi gradualmente revista: a adição do desenvolvimento tecnológico a um mercado livre dominado por capitalistas seria desvantajosa para os interesses da classe trabalhadora e a harmonia social ficaria em risco. Esta harmonia só poderia ser assegurada por uma combinação da natureza com a sociedade, através da intervenção desta na diminuição do antagonismo social que emergia (Ibidem, p. 590). A posição de Ricardo é complexa e a leitura que Polanyi faz é algo ambígua: mas a afirmação que melhor a sintetiza é talvez esta: “no caso de Ricardo, a própria teoria incluía um elemento que contrabalançava o rígido naturalismo”. Agradecemos as observações de António Luís Silva Baptista em geral e em especial sobre este ponto.

¹⁷ Polanyi, *The Great Transformation*, 118 (cap. 10).

motivações do recém-criado (e largamente fictício) *homo oeconomicus*. O “pauperismo” passa a andar de braços dados com a Utopia.

Segundo Polanyi, o naturalismo de Townsend influenciaria de forma determinante a lei da população de Thomas Malthus e a sua célebre analogia entre a fertilidade da terra e a fertilidade das pessoas. Também seria a fome a desempenhar um papel de legislador significativo, ao determinar o limite natural após o qual os homens não se poderiam multiplicar. Malthus recebe todos na selva e determina que as suas leis são a autoridade necessária para as relações económicas que se estabelecem. O naturalismo servia, portanto, para encontrar uma causa para o crescente problema da miséria e do pauperismo.

Pelo contrário, Ricardo opõe-se a uma concepção naturalista rígida e avança com uma fractura epistemológica relevante ao determinar que o trabalho é o verdadeiro constituinte do valor, desenvolvendo uma teoria do valor lockeana¹⁸. O que isto significa é que as leis da sociedade emergente, do sistema do mercado, não seriam já fundadas nas leis morais da Natureza, mas na lei dos homens. O Homem, ao criar as leis, é encarregado de transformar a sociedade a partir dos princípios da justiça. Isso pode significar um aumento do papel interventivo do Estado, que não passará pelo papel assistencialista que remete os pobres para o limiar da sobrevivência, mas para um contexto em que poderão florescer e maximizar as suas potencialidades. De certa forma, era isso que Robert Owen advogava quando remetia para o Estado o dever de minimizar os danos da comunidade. Se a justiça passava a reger a sociedade, o paternalismo paroquial ou estatal caía e abria a sociedade ao seu propósito de autonomização do indivíduo através da atribuição de responsabilidades que permitiriam alcançar um propósito emancipatório e auto-protector.

¹⁸ Polanyi, *The Great Transformation*, 132. Na perspectiva de Polanyi, a teoria económica de Ricardo assenta num equilíbrio entre elementos naturalistas e "humanistas".

4. Adam Smith, mão invisível e utopia

Quando Adam Smith, no fim do séc. XVIII, defendeu a libertação da economia dos controlos mercantilistas, não se limitou a formular uns quantos objectivos de política económica, ao estilo pragmático ou incremental, propôs um novo princípio de ordem social. Esse “princípio era o da anarquia ordenada: um regime descrito por direitos individuais bem definidos e pela liberdade e garantia dos contratos voluntários”,¹⁹ que dispensava a necessidade de um decisor centralizado para além de um “estado estritamente protector”, às vezes descrito pelos seus críticos como mero “guarda-nocturno,” para utilizar a expressão tornada famosa por um Thomas Carlyle descoroçoado com a estreiteza da *dismal science*.

Este ideal, que vigorou no essencial ao longo do século XIX, teria sofrido o impacto negativo da crítica socialista. “Os críticos socialistas tiveram êxito na identificação de particulares falhas na ordem conceptualmente ideal do *laissez-faire*”,²⁰ bem como nas suas aplicações práticas, mas não em descrever o que seria o estado ideal depois da sua substituição. Mas mesmo Marx é parco na descrição da ordem social “depois da revolução” e o que surgiu foram “monstruosidades burocráticas.”²¹ Mas por causa dessa crítica ao princípio do *laissez-faire*, a erosão pragmática do princípio do “governo limitado” adquiriu respeitabilidade intelectual e académica. “Seguindo isto foram propostas correcções, correcções que tomaram quase sempre a forma institucional de acção governamental” e a controvérsia e o debate político deslocou-se dos princípios alternativos de organização social para “escolhas de política específica num contexto situacional.”²²

“A ciência económica do *welfare*, no seu brilho de século XX, tornou-se uma teoria das falhas do mercado”. Não deve surpreender-nos, pois, que isso conduzisse a um crescimento rápido da dimensão e escopo do sector público. “Correctivos governamentais a presumidas falhas particulares da operação do

¹⁹ Buchanan, J. *The limits of liberty: between anarchy and Leviathan*, Chicago: The Univ. of Chicago Press, pp.170-173, (1975, reed. 1984). Este retrato segue de perto o apresentado nesta obra.

²⁰ Buchanan, op. cit., p. 171.

²¹ Ibidem.

²² Ibidem.

mercado foram consideradas incrementalmente e independentemente uns dos outros”,²³ julgando-se que funcionariam perfeitamente uma vez introduzidos, pois na ausência de uma visão global do estado, presumia-se ingenuamente que intenções equivaliam a resultados.

Hayek endossa essencialmente este retrato esquemático, enquanto Polanyi propõe uma leitura muito diferente dos acontecimentos. A este retrato histórico opõe o seguinte retrato alternativo, que Fred Block sintetiza assim

[Polanyi] [d]enuncia explicitamente os adeptos liberais do mercado que acusavam uma ‘conspiração colectivista’ de erigir barreiras de protecção contra os efeitos do funcionamento dos mercados globais. Sustenta que, pelo contrário, a criação dessas barreiras foi uma resposta espontânea e não planeada por parte de todos os grupos da sociedade perante as exigências impossíveis do sistema do mercado autorregulado. O contramovimento de protecção *era necessário* para impedir o desastre de uma economia desincrustada. Polanyi sugere que o movimento orientado para uma economia do *laissez-faire* torna necessário um contramovimento destinado a criar estabilidade.²⁴

Esta leitura resume-se com força e elegância na mensagem central do livro, formulada por Polanyi como um paradoxo: “O *laissez-faire* foi planeado; a planificação não o foi.” Quando Polanyi declara que o mercado “foi planeado,” quer dizer que a imposição da lógica do mercado, que veio substituir a rede de protecção da sociedade pré-industrial, implicou legislação e repressão política, na qual as teorias de Speenhamland tiveram um papel de charneira.

O original e polémico estudo histórico que a obra contém visa apresentar o contexto da emergência do liberalismo clássico: a invenção da economia política, com a criação deliberada da ideia de mercados autorreguláveis, leva a sociedade a defender-se de uma mudança para a qual não estava preparada através da planificação, essa sim, espontânea. O manifesto económico-político de Polanyi faz-se acompanhar de finos estudos antropológicos das economias não mercantis, algumas pré-industriais (incluindo uma recuperação da autarquia grega em Aristóteles), outras em civilizações não ocidentais, objecto de estudos etnológicos. Estes estudos prosseguem em obras individuais e colectivas posteriores.²⁵

²³ Ibidem, p. 172; ver nota 5 da mesma página.

²⁴ Ver na mesma linha, Williams, K. (1981). *From Pauperism to Poverty*, London: Routledge.

²⁵ Em 1957 volume publica com outros *Trade and Markets in the Early Empires* e mais tarde na revista que funda, “Coexistence”. Morre no Canadá em 1964. Cf. “Polanyi” (2003) in *Encyclopædia Britannica*, Chicago: Encyclopædia Britannica Inc., vol 9. p.554.

Até que ponto o argumento de Polanyi colhe? Muita da investigação histórica está inevitavelmente datada e, em qualquer caso, a história revisionista que apresenta oscila entre o estudo minucioso de certos textos e uma panorâmica várias vezes secular, que não pode deixar de seleccionar os eventos ou legislação que confirmam a interpretação que propõe.²⁶

Mas o caso que apresenta contra o nascimento do mercado é válido apenas contra aqueles que o defendem como uma inevitabilidade histórica. Hayek não tem ilusões sobre a inevitabilidade: o mercado, ou o estado de direito, etc. são conquistas civilizacionais frágeis, que têm que ser protegidas, porque são em si mesmas a melhor protecção da liberdade.

5. Falhas do Mercado e pobreza.

Ainda assim, o que Polanyi descreve como a “ilegalização da pobreza” e a criminalização do ócio contrasta com o optimismo de Hayek, no *The Road to Serfdom*, quanto à capacidade do mercado para resolver (imperfeitamente, reconhece) o problema da pobreza.

Embora Hayek só use a expressão “*safety net*” em obras posteriores e finalmente pareça repudiar a ideia,²⁷ a necessidade de resolver o problema está claramente presente no *The Road to Serfdom*:

Não há para que, numa sociedade que atingiu um nível geral de riqueza como o nosso, o primeiro tipo de segurança não seja garantido a todos sem pôr em perigo a liberdade geral. Há questões difíceis quanto ao nível específico que deve ser garantido; em especial, a questão importante de se saber se os que contam com a comunidade deverão usufruir indefinidamente das mesmas liberdades dos restantes. A abordagem irrefletida destas questões pode muito bem causar problemas políticos sérios e, até, perigosos. Mas não pode haver dúvidas de que deve ser assegurado a todos um mínimo de alimentação, abrigo e roupas, o suficiente para cada pessoa se manter sã e poder trabalhar.²⁸

²⁶ Mas acerca de Speenhamland, vários estudos da Old Poor Law em Snell, K. D. M. (1985). *Annals of the Labouring Poor: Social Change and Agrarian England, 1660-1900*, Cambridge: Cambridge University Press; e Boyer, G. (1990). *An Economic History of the English Poor Law, 1450-1850*, Cambridge: Cambridge University Press, questionam a interpretação de Polanyi.

²⁷ “Ainda não me tinha libertado por inteiro das superstições intervencionistas da atualidade e, conseqüentemente, fiz várias concessões que hoje considero injustificadas.” Prefácio de 1976 de Hayek, *Road to Serfdom*, 54, a tradução portuguesa é nossa.

²⁸ Hayek, *Road to Serfdom*, 148, (156-157 da tradução portuguesa).

Pode surpreender os que o lêem à luz da sua obra tardia que Hayek, em *The Road to Serfdom*, favorecesse uma série de intervenções governamentais bastante pronunciadas. Por um lado, advoga arranjos que permitissem que uma concorrência “harmoniosa” pudesse surgir. Na prática, isto significava não simplesmente a conservação de certas instituições, mas uma mão activa do governo em melhorá-las: “(...) há toda a diferença entre criar intencionalmente um sistema dentro do qual a competição será tão benéfica quanto possível, e aceitar passivamente as instituições tal como são”.²⁹

Hayek também acreditava em alguns pressupostos que fariam com que a concorrência funcionasse: o estabelecimento de uma moeda estável (mas, ao contrário de Mises e Polanyi, não dá demasiada importância ao padrão-ouro), vias de comunicação, mercados, prevenção de fraude, etc. Achava também que: “(...) o Estado pode fazer muito para ajudar a divulgar o conhecimento e a informação, e ajudar à mobilidade,”³⁰ ou mesmo combater falhas de mercado, construir estradas e sinais indicadores, desenhar regulamentos e leis laborais e, enfim, regular companhias privadas detidas pelo Estado — o que pressupõe a sua legitimidade.

Mais importante, reconhece a necessidade do que depois chamou “rede de segurança”, cujo escopo é bastante amplo e inclui seguros contra terremotos de terra e outros desastres naturais e ainda ajudas sociais para momentos de flutuação económica — não excluindo sequer políticas monetárias e obras públicas.³¹

À primeira vista, Hayek parece estar a criar uma “terceira via” entre o mercado entregue a si mesmo e a planificação socialista. Devemos ter presente que *The Road to Serfdom* consagra uma firme negação de todas as formas extremas de *laissez-faire*.

²⁹ Hayek, *Road to Serfdom*, 71, (43 da tradução portuguesa).

³⁰ Hayek, *Road to Serfdom*, 129, (128 na tradução portuguesa).

³¹ Aqui estamos a seguir a cuidadosa lista de Shearmur, J. (1997). Hayek, Keynes and the State. *History of Economics Review*. 26 (1), : “Hayek, Keynes and the State,” 71-72.

Ver referências à famosa “safety net” em Tebble, A. J. (2010). *F. A. Hayek (Major Conservative and Libertarian Thinkers)*. Continuum International Publishing Group Ltd, nota de rodapé 62, página 75.:

Algumas referências às intervenções em *Road to Serfdom* na nota de rodapé 63, página 66 de: Jackson, B. (2012). Freedom, the Common Good, and the Rule of Law: Lippmann and Hayek on Economic Planning. *Journal of the History of Ideas*, 73(1), 47-68.

É importante não confundir a oposição a este tipo de planeamento com uma atitude *laissez-faire* dogmática. O argumento liberal defende que se faça o melhor uso possível das forças da concorrência como forma de coordenar os esforços humanos, e não como argumento para se deixar tudo como está. Baseia-se na convicção de que, nos casos em que se pode criar verdadeira concorrência, esta será uma maneira melhor de orientar os esforços individuais que qualquer outra.³²

Esta posição “intermédia” não deve ser lida, como alguns o fizeram,³³ incluindo o próprio Hayek algumas décadas mais tarde,³⁴ como uma espécie de “concessão.” Com efeito, só depois do famoso Colóquio Walter Lippmann que, para alguns, marca o nascimento do “neo-liberalismo” como tal, Hayek tentou genuinamente construir uma alternativa que evitasse o Caríbdis do *laissez-faire* e o Cila do planeamento centralizado. (Para não dizer nada dos “ciclopes” do Socialismo).³⁵ A “rede de segurança” fazia, portanto, sentido neste contexto de tentativa de encontrar uma verdadeira “via intermédia” (e não simplesmente de “fazer uma concessão.”)

Isto, no entanto, deixa por explicar as posições que Hayek acabará por ter sobre a falta de significado do termo “justiça social” e a sua crítica mordaz do Estado Social. Como é que reconciliamos o que Hayek disse antes e depois? Não há dúvidas que Hayek se opôs cada vez mais às intervenções estatais e, certamente, houve aqui uma mudança gradual de posição.³⁶ Mas a explicação não pode parar numa eventual alteração do seu ideário, visto que já no seu ensaio “*Freedom and the Economic System*,”³⁷ de 1938, Hayek já tinha feito uma distinção clara entre o Estado *incentivar* escolhas e *impor* escolhas, entre planear a concorrência e planear o socialismo. Tais distinções eram ainda vagas e Hayek só definiria o seu pensamento em *The Constitution of Liberty*, mas mesmo assim

³² Dificilmente se lê o livro sem sentirmos a influência, e a reação contra, Ludwig von Mises: Hayek, *Road to Serfdom*, 85-86, 63 da tradução portuguesa. “Tudo o que um bom governo pode fazer para melhorar o bem-estar material das massas é estabelecer e preservar o enquadramento institucional no qual não existem obstáculos à acumulação progressiva de novo capital e à sua utilização para melhorar as tecnologias de produção.” Esta resposta de Mises, se não diretamente contra Hayek, é claramente dirigida contra quem tenha a mesma posição. Vem do ensaio de 1945 com o nome sugestivo ‘Planning for Freedom.’ Mises, L. V., Greaves, B. B., & Mises, L. V. (2008). *Planning for Freedom: Let the Market System Work: A Collection of Essays and Addresses*. Indianapolis: Liberty Fund., 6.

³³ Tebble, *Hayek*, 70-72.

³⁴ Ver a primeira nota de rodapé. Não é surpreendente que, olhando retrospectivamente, Hayek sentisse que as intervenções governamentais que sugere em *Road to Serfdom* eram “concessões.” Hayek tornou-se cada vez mais cético de qualquer intervenção estatal o que, no final da sua vida, culminou na sua posição mais extrema em *The Fatal Conceit*. Ver também: Ebenstein, A. O. (2015). *Chicagonomics: The Evolution of Chicago Free Market Economics*. New York: St. Martins Press.

³⁵ Sobre isso ver, Jackson, “Freedom the Common Good and the Rule of Law

³⁶ Cf. nota anterior.

³⁷ Hayek, F. A. (2012). *Freedom and the Economic System*. Mansfield Centre: Martino Publishing. REF INCOMPLETA

entendia já naquela altura que o Estado podia, e até devia, intervir para que os indivíduos possam efectivar os seus planos.³⁸

A sua evolução posterior decorre desta distinção.

6. Concessão ou incoerência?

Aos nossos olhos a “rede de segurança” parece ser ora uma “concessão” ora uma “incoerência”, mas, na verdade, esta posição está fundamentalmente em acordo com a posição de Hayek na altura em que este escreveu *The Road to Serfdom*. O Estado deveria ter um papel activo para aumentar a capacidade (*local*) dos indivíduos ao planearem as suas vidas. O que o Estado deve evitar é endossar um plano (*global*) que acabaria inevitavelmente em totalitarismo.

Esta insistência em capacitar os indivíduos para que sejam “agentes” do mercado parece estar a obscurecer a grande questão que os “planeadores” queriam resolver: o que se faz com a pobreza? Tanto Hayek como os seus críticos concordam que as formas mais *absolutas* de pobreza devem ser evitadas. Hayek até era a favor de medidas que reduzissem a desigualdade de oportunidades (mas só na medida em que o “plano” de um indivíduo não se sobrepusesse ao dos outros.)³⁹ Hayek sempre disse também que as intervenções deveriam prevenir problemas que impedissem o agente de se tornar “bom” concorrente, bom “micro” planeador, etc.

Hayek estava consciente que defender que o capitalismo podia reduzir a pobreza (a longo prazo) à época parecia um argumento bastante fraco, principalmente no momento em que escreve *The Road to Serfdom*: no fim de uma década onde tudo indicava que o capitalismo tinha falhado na luta contra a pobreza e o desemprego e onde o “planeamento” que Hayek detestava tinha sido aparentemente bem-sucedido. Pior ainda, nas palavras de Caldwell “Os sacrifícios comuns da guerra criaram o sentimento de que todos deveriam

³⁸ Algo que, para alguém como o Mises, era anátema: no máximo, o governo pode *retirar* obstáculos, mas não pode *incentivar* nada.

³⁹ Há motivos para reduzir esta desigualdade de oportunidades, na medida em que as diferenças congénitas o permitam e que se o possa fazer sem destruir o carácter pessoal do processo, no qual todos têm de arriscar e em que nenhuma opinião quanto ao que é legítimo e desejável se sobrepõe à dos outros. ”Hayek, *Road to Serfdom*, 134 (136 da tradução portuguesa).

partilhar tudo de forma mais igual quando a reconstrução viesse.”⁴⁰ Perante a fraqueza conjuntural do capitalismo para resolver os problemas que os planeadores tentavam resolver, Hayek teve que criticar a superioridade produtiva (presumível) do sistema planeado. Há várias passagens de *The Road to Serfdom* onde Hayek critica o “argumento irresponsável” dos planeadores de que centralizar a economia seria uma espécie de corno de abundância à espera de ser libertado:

(...) [E]mbora este embuste tenha servido a propaganda socialista a coberto de vários nomes desde que o socialismo existe, é ainda tão manifestamente falso quanto era há mais de cem anos, quando foi pela primeira vez utilizado. Em todo este tempo, nenhuma das muitas pessoas que o usaram, apresentou um plano exequível de como se poderia aumentar a produção por forma a abolir o que consideramos ser a pobreza, mesmo no Ocidente – já para não falar do resto do mundo. Creia o leitor que, quem quer que lhe fale de uma abundância potencial, ou é desonesto ou não sabe do que fala.⁴¹

Enfim, Hayek estava condenado a defender o capitalismo com uma abordagem “deontológica”. Numa das poucas passagens onde Hayek aborda a questão da pobreza, aponta claramente nesta direcção:

O facto de, numa sociedade em regime de concorrência, as oportunidades disponíveis aos pobres serem muito mais restritas do que aquelas à disposição dos ricos, não faz com que seja menos verdadeiro que, numa sociedade baseada na concorrência, os pobres sejam muito mais livres do que uma pessoa com maior conforto material num outro tipo de sociedade. Embora num regime de concorrência a probabilidade de alguém que nasceu pobre enriquecer ser menor do que para alguém que tenha herdado propriedade, ainda assim é possível, e só no regime de concorrência é que essa pessoa depende apenas de si próprio e não de favores dos poderosos, e ninguém o pode impedir. Só o facto de nos termos esquecido o que significa a falta de liberdade é que faz com descuremos [*sic*] o facto evidente de, neste país, um trabalhador não qualificado e mal pago ter mais liberdade para decidir a sua vida do que muitos pequenos empresários na Alemanha ou engenheiros ou directores bem pagos na Rússia.⁴²

Um pouco cruamente, podemos dizer que a resposta global de Hayek em relação à questão da pobreza era “não se pode ter tudo”: desejar mais do que dar aos indivíduos capacidades para que estes formulem e sigam os seus próprios planos implica necessariamente um plano global e isto já é um passo a mais na direcção do totalitarismo. Nesta linha, Hayek critica o facto de que valorizamos

⁴⁰ Caldwell, ‘Introduction,’ em Hayek, *Road to Serfdom*, 14.

⁴¹ Hayek, *Road to Serfdom*, 131 (131-132 da tradução portuguesa).

⁴² Hayek, *Road to Serfdom*, 135 (136 da tradução portuguesa).

demasiado os nossos desejos “locais” à custa de levantar obstáculos à capacidade de todos os outros seguirem as suas próprias preferências:

(...) [T]odos cremos que a nossa ordem de valores pessoal não é apenas pessoa, e que numa discussão livre entre pessoas racionais conseguiríamos convencer os outros da justeza da nossa. O apreciador do campo que pretende, acima de tudo, que a aparência tradicional deste mesmo campo seja preservada e que a mácula que a indústria já infligiu à sua bela face seja removida, tal como o entusiasta da saúde que quer que todas as casas do campo pitorescas, mas pouco sanitárias, sejam eliminadas, ou um motorista que pretende que se rasgue o campo com amplas estradas, o fanático eficiente que deseja o máximo de especialização e mecanização e o idealista que, por causa do desenvolvimento da personalidade, pretende preservar o maior número possível de artesãos. Todos sabem que o seu objetivo só pode ser plenamente alcançado pelo planeamento – e, por essa razão, todos querem o planeamento.⁴³

Melhor dito, e como Hayek já tinha afirmado em 1938 em “Freedom and the Economic System”, para passarmos de planos individuais para planos globais, exigir-se-ia uma unanimidade de valores que, na prática, nunca existe. E isso por mais que os planeadores achem que tal consenso existe em torno da “igualdade.”⁴⁴

Que a vida e a morte, a beleza e a virtude, a honra e a paz de espírito, muitas vezes só podem ser preservadas com considerável sacrifício material, e que alguém tem de fazer a escolha, é tão inquestionável como o facto de por vezes não estarmos preparados para fazer os sacrifícios materiais necessários para proteger esses valores mais nobres de qualquer ataque.

Um exemplo: podíamos, evidentemente, acabar com as mortes em acidentes rodoviários se estivéssemos dispostos a suportar os custos da abolição do automóvel. O mesmo é válido para milhares de outros casos em que arriscamos constantemente a vida e a morte e os mais nobres valores de espírito, nossos e do nosso semelhante, para melhorar aquilo a que, ao mesmo tempo, nos referimos desdenhosamente como o nosso conforto material. Nem poderia ser de outra forma, pois todos os nossos fins competem pelos mesmos meios. E só poderíamos aspirar a estes valores caso eles não estivessem em risco.⁴⁵

Retrospectivamente, tal parece ser demasiado pouco no que toca ao pauperismo que aflige Polanyi. A resposta é, em parte, tática: em 1944, Hayek não podia montar uma defesa do capitalismo em torno de promessas de prosperidade geral. Podia apenas, por um lado, tranquilizar os seus leitores ao garantir que a pobreza absoluta seria eventualmente resolvida com a sua

⁴³ Hayek, *Road to Serfdom*, 99, (83-84 da tradução portuguesa).

⁴⁴ “Concordar sobre um plano particular requer muito mais do que um acordar de algumas regras éticas gerais; requer muito mais do que uma adesão geral a qualquer código ético que tenha alguma vez existido; requer um determinado tipo de escala de valores, quantitativa e exaustiva, que emerge nas decisões concretas de cada indivíduo, mas sobre a qual, numa sociedade individualista, não há acordo necessário ou existente.” Caldwell, B. (2014). *Socialism and War: Essays, Documents, Reviews*. London: Routledge, 183.

⁴⁵ Hayek, *Road to Serfdom*, 130 (131 da tradução portuguesa).

proposta. Podia ainda, por outro lado, insistir nos benefícios *inerentes* ao capitalismo e nos defeitos inerentes ou decorrentes das suas alternativas.

7. Porque desiste Hayek de resolver o problema do pauperismo?

O problema é, cremos, que a rede de segurança, ou a garantia do mínimo decente, não podia e não pode deixar de distorcer o sistema de recompensas do mercado e o seu papel na revelação de preferências. A intervenção de Hayek na querela do cálculo em regime socialista não se baseia essencialmente em dificuldades de computação que hoje seriam facilmente resolvidas, mas na capacidade única do mercado para satisfazer necessidades ou preferências que ninguém conhece *a priori* e que só o mercado revela, isto é, que são resultado de escolhas livres e de planos pessoais não organizados dos indivíduos, que ninguém pode prever.

Em suma, a alocação de bens produzidos numa sociedade faz-se através de escolhas individuais e colectivas. Esta alternativa é muitas vezes apresentada de forma algo simplificada⁴⁶ como equivalente à escolha entre dois mecanismos, ambos imperfeitos, *mercados* e *governo*, “que é complexa e não é usualmente binária.”⁴⁷ A escolha individual seria expressa através de trocas voluntárias no mercado. Constata-se, todavia, que estas se afastam, por vezes de maneira imprevista, de referenciais sociais como a eficiência, liberdade ou equidade: são o que se chama as *falhas do mercado*.

A escolha colectiva, exercida através das estruturas governamentais, proporciona a possibilidade de corrigir algumas dessas deficiências. Mas a acção colectiva através do governo também se desvia dos objectivos que a justificam: são as *falhas do estado*. Algumas falhas do mercado são demasiado custosas de corrigir, algumas soluções distributivas implicam um empobrecimento geral. Às

⁴⁶ Por exemplo Buchanan, J. M., & Tullock, G. (1961). *The Calculus of Consent; A Preliminary Analysis of Individual Constitutional Choice*, Ann Arbor: The University of Michigan, cap. V "A organização da atividade humana" distinguem seis categorias em função de parâmetros relativos ao custo da decisão privada, da organização voluntária e da decisão coletiva.

⁴⁷ Wolf, C. (1988). *Markets or Governments: Choosing between Imperfect Alternatives*, Cambridge: The Rand Corp e MIT Press, 151

vezes os custos são superiores aos benefícios. A intervenção pública deve, pois, ser decidida tendo em conta o saldo final de ambas as falhas.⁴⁸

A história que Polanyi conta não é só a das falhas do mercado autorregulado, mas de uma autodefesa da sociedade à custa da eliminação da liberdade. A história que Hayek conta não é só a das falhas do estado, mas da impossibilidade dos mercados resolverem o problema da equidade social—e mesmo, como mais tarde se dará conta, de assegurarem os mínimos decentes. O balanço que os dois fazem é distinto, e será provavelmente diferente do que outros farão em diversos momentos da história.⁴⁹

O que Polanyi mostra de um modo mais claro que Hayek é que a avaliação desse resultado tem também uma dimensão histórica. Houve momentos de abandono à esfera privada de áreas que hoje justificam a intervenção pública e vice-versa. Não devemos, pois, esquecer as falhas ou limites da abordagem da própria ciência económica na definição das políticas públicas, mas, apesar dos receios que os movimentos populistas podem levantar, não estamos certamente na situação em que Hayek e Polanyi apresentaram as suas propostas em que a liberdade, senão a civilização, parecia estar em jogo.

Alguns autores que se reclamam agora das teorias de um ou outro autor e confundem a crise financeira recente com uma prova do fracasso do mercado esquecem convenientemente que a crise teve o seu epicentro nos sectores mais regulados (financeiro e seguros).⁵⁰ Graças ao estado social, o desemprego não teve consequências à escala do passado, nem nunca esteve realmente em causa o regresso dos fascismos. Por outro lado, se a regulação ainda assim fracassou, a rede de protecção, esta sim, funcionou, e, se funcionou, foi porque o estado social, e não apenas os chamados “estabilizadores automáticos”, impediu as consequências mais extremas da crise.

⁴⁸ Weimar, D. L. e Vining, A. R. (2005). *Policy Analysis: Concepts and Practice*, New Jersey: Pearson Prentice Hall, 156-7

⁴⁹ Um moderado como Raymond Aron dizia que Marx parecia ter razão nos anos 30 e que Tocqueville parecia ter razão no fim do pequeno século XX.

⁵⁰ Não pretendemos obviamente discutir aqui o problema, demasiado complexo, das causas da crise financeira, nem da importância da desregulação na crise, apenas notamos que o sector financeiro, devido à sua óbvia importância para a chamada “economia real”, foi sempre historicamente um dos sectores económicos mais regulados, antes durante e depois da “desregulação,” o que não exclui uma mudança de atitude para uma perigosa complacência em relação à capacidade do mercado para se autorregular.

Não estamos a assistir ao choque de duas vagas políticas com dimensão de quasi-crise civilizacional, que se combatem entre si, como quando Hayek e Polanyi escreviam. Quanto muito, podemos falar de uma curva apertada no caminho, devida mais ao excesso de “poesia” democrática que ao fracasso da democracia representativa.

Bibliografia

Boyer, G. (1990). *An Economic History of the English Poor Law, 1450-1850*. Cambridge, UK: Cambridge University Press.

Buchanan, J. M., & Tullock, G. (1961). *The Calculus of Consent; A Preliminary Analysis of Individual Constitutional Choice*. Ann Arbor, MI: The University of Michigan.

Caldwell, B. (2014). *Socialism and War: Essays, Documents, Reviews*. London, UK: Routledge.

Ebenstein, A. O. (2003). *Friedrich Hayek: A Biography*. Chicago, IL: University of Chicago Press.

Ebenstein, A. O. (2015). *Chicagonomics: The Evolution of Chicago Free Market Economics*. New York, NY: St. Martins Press.

Hayek, F. A. (2012). *Freedom and the Economic System*. Mansfield Centre, Conn: Martino Publishing.

Hayek, F. A., & Caldwell, B. (2014). *The Road to Serfdom: Text and Documents: The Definitive Edition*. New York, NY: Routledge.

Hayek, F. A. (2008). *O Caminho para a Servidão*. Lisboa, Portugal: Edições 70.

Jackson, B. (2012). Freedom, the Common Good, and the Rule of Law: Lippmann and Hayek on Economic Planning. *Journal of the History of Ideas*, 73(1), 47-68.

Janik, A., & Toulmin, S. E. (1996). *Wittgensteins Vienna*. Chicago: DeeKresge, S. (1994). *Hayek on Hayek: An Autobiographical Dialogue*, London, UK: Routledge.

Mises, L. V., Greaves, B. B., & Mises, L. V. (2008). *Planning for Freedom: Let the Market System Work: A Collection of Essays and Addresses*. Indianapolis: Liberty Fund.

Polanyi, K. (2001). *The Great Transformation: The Political and Economic Origins of Our Time*, Boston, MA: Beacon Press.

Polanyi, K. (2016). *A Grande Transformação: As Origens Políticas e Económicas do nosso Tempo*. Lisboa, Portugal: Edições 70.

Polanyi-Levitt, K., & Mendell, M. (1987). Karl Polanyi: His Life and Times. *Studies in Political Economy*, 22(1), 7-39.

Schorske, C. E. (1992). *Fin-de-siècle Vienna: politics and culture*. Cambridge, UK: Cambridge University Press.

Shearmur, J. (1997). Hayek, Keynes and the State. *History of Economics Review*. 26 (1), 68-82.

Snell, K. D. M. (1985). *Annals of the Labouring Poor: Social Change and Agrarian England, 1660-1900*, Cambridge, UK: Cambridge University Press.

Tebble, A. J. (2010). *F. A. Hayek (Major Conservative and Libertarian Thinkers)*. London, UK: Continuum International Publishing Group Ltd.

Weimar, D. L. e Vining, A. R. (2005). *Policy Analysis: Concepts and Practice*. New Jersey, NJ: Pearson Prentice Hall.

Williams, K. (1981). *From Pauperism to Poverty*, London, UK: Routledge.

Wolf, C. (1988). *Markets or Governments: Choosing between Imperfect Alternatives*. Cambridge, Mass: The Rand Corp e MIT Press.



THE DOUBLE MOVEMENT IN POLANYI AND HAYEK: TOWARDS THE CONTINUATION OF LIFE

O DUPLO MOVIMENTO EM POLANYI E HAYEK: EM DIREÇÃO À CONTINUIDADE DA VIDA

Filipe Nobre Faria*
filipefaria@gmail.com

Abstract. *Karl Polanyi's double movement is a dialectical process characterized by a continuous tension between a movement towards social marketization and a movement towards social protectionism. Notably, Polanyi condemns the former movement while defending the latter. Without using the term "double movement", F.A. Hayek's theory of social evolution acknowledges the same phenomenon but reaches different normative conclusions. While for Polanyi the marketization of society is a utopia with dystopian consequences, Hayek's evolutionary explanation of this dialectical process asserts that there is no alternative to a market oriented society. Both authors defend that their favoured movement is the one that truly supports the continuity of life. This article compares the authors' normative readings of the double movement and concludes that, from an evolutionary perspective, Polanyi's conclusion possesses a robustness that Hayek's postulate lacks.*

Keywords: *Karl Polanyi, F.A. Hayek, double movement, group selection, cultural evolution, market liberalism.*

Sumário. *O duplo movimento de Karl Polanyi é um processo dialéctico caracterizado por uma tensão contínua entre o movimento a favor da comercialização social e o movimento a favor do protecionismo social. Notavelmente, Polanyi condena o primeiro movimento enquanto faz a defesa do segundo. Sem usar o termo "duplo movimento", a teoria da evolução social de F.A. Hayek reconhece o mesmo fenómeno mas chega a conclusões normativas distintas. Enquanto que para Polanyi a comercialização da sociedade é uma utopia com consequências distópicas, a explicação evolucionista de Hayek conclui que não há alternativa a uma sociedade de mercado. Ambos os autores defendem que o seu movimento preferido é aquele que de facto suporta a continuidade da vida. Este artigo compara as leituras normativas que os dois autores fazem do duplo movimento e conclui que, de uma perspectiva evolutiva, a conclusão de Polanyi possui uma robustez que falta ao postulado de Hayek.*

* Universidade Nova de Lisboa (FCSH-NOVA), Nova Institute of Philosophy (IFILNOVA), Lisbon, Portugal.

Palavras chave: *Karl Polanyi, F.A. Hayek, duplo movimento, seleção de grupo, evolução cultural, liberalismo de mercado.*

1. Introduction

Karl Polanyi and F.A Hayek are two of the most important social thinkers of the 20th century. They are iconic representatives of two contrasting positions: while Polanyi (2001, p. 136) affirms that market societies are self-destructive, Hayek (1944) defends the virtues of market liberalism. As Peter Lindsay notes, “while both authors covered much similar ground and employed similar arguments, they somehow arrived at diametrically opposed conclusions” (Lindsay, 2015, p. 377). It is particularly interesting that both authors developed social theories about the phenomenon known as “the double movement”.¹ This is a term coined by Polanyi (2001, p. 136) that refers to the continuous dialectical tension between a movement favouring the marketization of society and a movement aiming at social protection. Yet, without using the term “double movement”, Hayek (1979, 1988) also identifies the same phenomenon and theorizes it from an evolutionary perspective. Most importantly, when analysing the double movement, both thinkers believe that their favoured movement supports the continuation of life. While Hayek (1979, p. 168) defends that market liberalism expands life, Polanyi (2001, pp. 3, 74) believes that life can only flourish when markets are properly embedded within social relations and subjugated to non-economic norms. Hence, these two authors independently acknowledge the same phenomenon and postulate opposite normative positions.

The aim of this paper is to show how Polanyi and Hayek theorize the phenomenon of double movement and to reveal how they come to their distinct normative positions. Furthermore, the article aims at critically assessing their normative interpretation of this phenomenon, especially focusing on their arguments in favour of the continuation of life. This critical assessment makes use of evolutionary theorization in order to appraise the normative claims of both authors. Specifically, it makes use of the evolutionary models of group selection/multilevel selection (Sober & Wilson, 1998; Wilson & Wilson, 2007)

¹ Although the existence of a tension between pro-market and pro-protectionist forces in society is here being described as a “double movement”, it is worth noticing that both Hayek and Polanyi’s postulates are theoretical interpretations of the same development. Hence, it is not implied that Polanyi’s double movement is a scientific fact while Hayek’s insight is a mere theory.

and gene-culture co-evolution (Boyd & Richerson, 2005).² These well known contemporary models have the same primary evolutionary understanding as Hayek's social theory and provide an important framework with which to evaluate not only Hayek's claims but also Polanyi's. Given that both authors justify their normative positions with claims about the continuation of life, this critical assessment will focus on the evolutionary equivalent of this continuation: reproductive fitness.³

The paper is structured as follows. After this introduction, the second section describes Polanyi's understanding of the double movement. Subsequently, the third section presents Hayek's evolutionary explanation of the same phenomenon. The fourth section critically discusses and compares the robustness of the authors' visions. Ultimately, Hayek's normative understanding of the double movement is found to be less robust than Polanyi's. The conclusion summarizes the main insights.

2. Polanyi's double movement

The double movement is a concept coined and developed by Polanyi. He asserts that the forces that aim at expanding the marketization of society will inevitably be met by a countermovement that tries to protect social life from the negative effects of this marketization. In his own words: "For a century, modern society was governed by a double movement: the market expanded continuously but this movement was met by a countermovement checking the expansion in definite directions" (Polanyi, 2001, p. 136). Polanyi concludes that this countermovement is a reaction against the "dislocation which attacked the fabric of society" and that such dislocation would have ultimately "destroyed the very organization that the market called into being" (Polanyi, 2001, p. 136). Put differently, the destructive social forces created by the self-regulating market would not only undermine the social fabric but also the market system itself. In

² Both these models share a basic common understanding with Hayek's cultural group selection theory, therefore providing a common evaluative standard. Like Hayek's framework, the group selection/multilevel selection model regards selection at the level of social groups as an important evolutionary force. Furthermore, similarly to Hayek's theory, the gene-culture co-evolution model understands that the cultural practices and institutions of groups are under selection in the same way as genes. Hence, due to mutual interaction, cultural selection affects genetic selection and vice-versa.

³ This paper understands reproductive fitness as it is understood in standard evolutionary biology, that is, as reproductive success (of genotypes and phenotypes).

the end, the unregulated market system is self-destructive. For Polanyi, “the idea of a self-adjusting market implied a stark utopia” (Polanyi, 2001, p. 3) with serious dystopian consequences. Ultimately, the final consequence of market liberalism is the annihilation of the “human and natural substance of society” (Polanyi, 2001, p. 3), which means the physical destruction of Man and his surroundings.

Polanyi sees this countermovement against the marketization of society as vital to fight the negative social consequences coming from market forces; specifically:

the dangers involved in the exploitation of the physical strength of the worker, the destruction of family life, the devastation of neighbourhoods, the denudation of forests, the pollution of rivers (...), the disruption of folkways, and the general degradation of existence including housing and arts (Polanyi, 2001, p. 139).

He adds that market forces are so destructive that even the “capitalist production itself had to be sheltered from the devastating effects of a self-regulating market” (Polanyi, 2001, p. 138) via this countermovement.

The attempt to protect society from the negative consequences of social marketization can come in many different forms. As Fred Block and Margaret Somers assert, these counter movements “are just as likely to be conservative, even populist and fascist” as to be socialist or communist, because “market destabilizations will mobilize the right no less than the left” (Block & Somers, 2014, p. 10). Most relevantly, Polanyi’s far-reaching hypothesis in his book *The Great Transformation* (2001) is that it was the liberal market system of the nineteenth-century that led to the two world wars of the twentieth century (Patomaki, 2014, p. 736). In order to prove it, he tries to historically show that the nineteenth-century market system was artificial, alien to people and a fairly recent invention. Therefore, because this market system was politically created from the top down, the twentieth century reaction against market liberalism was to be expected. With this reaction came the attempt to restrict market forces and to politically control the economy. As Polanyi famously puts it: “Laissez-faire was planned; planning was not” (Polanyi, 2001, p. 147).

For Polanyi, the main reason why market liberalism is unsustainable is that it attempts to transform vital elements of social life and of human dignity into commodities. Namely, the market system promotes the commodification of people by transforming them into priced labour. Furthermore, marketization removes the social significance of land and nature by allowing them to be valued through market prices. Even money is commodified “as a token of purchasing power” (Lindsay, 2015, p. 382). As a result, Polanyi sees market liberalism as a process that will ultimately “denaturalize the individual” (Polanyi, 2001, p. 245), adding that “the commodity fiction disregarded the fact that leaving the fate of soil and people to the market would be tantamount to annihilating them” (Polanyi, 2001, p. 137).

2.1 Double movement: disembedding and re-embedding

The phenomenon of double movement can also be conceptualized as the disembedding and the re-embedding of the economy. An economy is embedded when it is integrated in broader social relations of a non-economic nature, that is, “the economic system will be run on noneconomic motives” (Polanyi, 2001, p. 48). Although Polanyi understands that all societies have systems of material production, he asserts that this “does not imply the existence of separate economic institutions” (Polanyi, 2001, p. 74). His historical research reveals that “normally, the economic order is merely a function of the social order. Neither under tribal nor under feudal nor under mercantile conditions was there, as we saw, a separate economic system in society” (Polanyi, 2001, p. 74). Hence, the movement towards the marketization of society attempts to disembed the economy from the wider social relations, while the protective countermovement tries to re-embed it.

When exactly an economy is disembedded is a matter of contention. Some academics like Fred Block claim that an economy can never be really disembedded from social and political relations. In the words of Block:

One might say that disembedding the market is similar to stretching a giant elastic band. Efforts to bring about greater autonomy of the market increase the tension level. With further stretching, either the band will snap – representing social disintegration – or the economy will revert to a more embedded position (Polanyi, 2001, p. xxv).

According to this perspective, an economy is always embedded and only a complete social collapse would disembed it. Nonetheless, the alternative perspective is that the economy can be disembedded even before reaching total social disintegration. Nuno Cardoso Machado (2011) asserts that the latter perspective was the one Polanyi had. And indeed, Polanyi is quite explicit about this. He writes that:

The disembedded economy of the nineteenth century stood apart from the rest of society, more especially from the political and governmental system. (...) It is motivationally distinct, for it receives its impulse from the urge of monetary gain. It is institutionally separated from the political and governmental center. It attains to an autonomy that invests it with laws of its own. In it we possess that extreme case of a disembedded economy which takes its start from the widespread use of money as a means of exchange (Polanyi, 1957, pp. 67-68).

Thus, it seems clear that Polanyi thinks that contemporary market economies can be disembedded.⁴

The double movement can therefore be regarded as a dialectical tension between the disembedding and the re-embedding of the economy in social relations (i.e. in social goals of a non-economic nature). More specifically, the double movement is constituted by market forces aiming at an independent economy and by a protective countermovement aiming at re-embedding the economic sphere.

3. Hayek's evolutionary double movement

Hayek's theory of social evolution identifies the phenomenon of double movement without calling it as such. Namely, he gives an evolutionary account of this continuous tension between market liberalism and protective tribal countermovements. Furthermore, by using a model of cultural group selection, Hayek concludes that groups that embrace a market morality will bio-culturally expand in relation to groups with tribal moralities. Hence, market liberal groups

⁴ For further discussion on the question of embeddedness see Dale (2010) and Machado (2011).

will be favoured over tribal ones in the process of evolutionary inter-group competition (Hayek, 1979, 1988).

Hayek's double movement is defined by the continuous tension between two conflicting moral systems: the tribal morality and the "Great Society" market morality (Miller, 1989, p. 313). The tribal morality is a moral system underpinned by instincts of altruism towards the tribe. These instincts are the result of humans having lived most of their evolutionary past in small tribes. Consequently, such tribal traits were naturally selected. Hence, "mankind had hundreds of thousands of years to acquire and genetically to embody the responses needed for (...) the preservation of a small band of hunters and gatherers" (Hayek, 1979, p. 164). Conversely, the "Great Society" market morality is characterized by a high interdependence of a very large number of individuals. It is characterized by market institutions such as property rights, contracts and rule of law, which allow for an extended cooperation between many individuals and across vast tracts of land (Miller, 1989, p. 313).

The permanent tension between these two systems of morality is analogous to what Polanyi sees as the double movement. Indeed, Hayek (1988, p. 134) understands that individuals have been civilized in the market order largely against their wishes and their natural (tribal) instincts. For him, the reason why market morality developed is that those groups that practiced it gained reproductive/evolutionary advantages over groups operating under moral tribalism. Thus, market morality prevailed not because of our explicit wishes, but because it was an evolutionary winner. As Hayek notes, "we may not like the fact that our rules were shaped mainly by their suitability for increasing our numbers, but we have little choice in the matter" (Hayek, 1988, p. 134).

Hayek surely thought that this process of population growth was a key explanation of how market society evolved, but he also thought that this process is of high relevance for the present and future. After all, evolution simply cannot come to a halt. As Naomi Beck notices, despite being critical of the power of science and reason to understand morality, Hayek ended up "making growth or reproductive success the main ethical value" (Beck, 2011, p. 421). He clearly stated that the most essential effect of the evolutionary process is the maximization of the "prospective stream of future lives", because "life has no

purpose but itself: life exists only so long as it provides for its own continuance” (Hayek, 1988, pp. 132-133). At the moral level, Hayek concludes that “as with every other organism, the main purpose to which man’s physical make up as well as his traditions are adapted is to produce other human beings” (Hayek, 1988, p. 133). Hence, although he claims that there is no point in asking if fitness maximising actions are good in themselves (Hayek, 1988, p. 133), what is positive becomes tacitly tied with fitness maximization in the group selection process, and, according to his normative judgement, tied to market morality.⁵

For Hayek, the “Great Society” market morality allows for an exponential increase in the division of labour, which expands the capacity to produce more goods and services. Consequently, as Beck explains, due “to the increase in output, more mouths can be fed and population can grow” (Beck, 2015, p. 89). Yet, individuals are constantly rebelling against the market order because the individualistic market morality is opposed to our natural collectivist instincts, which are ever present due to our evolutionary tribal past. For Hayek, these instincts that favour the tribe have to be repressed because:

it was necessary for the rise (...) (of the open society) that he (mankind) not only learned to acquire new rules, but that some of the new rules served precisely to repress the instinctive reactions no longer appropriate to the Great Society (Hayek, 1979, p. 164).

Hence, due to being the drivers of the countermovement against the market order, tribal instincts should be repressed in the process of double movement. If not, Hayek (1988, p. 120) claims, millions would be sentenced to death by starvation due to the destruction of the material foundation of the “Great Society”.

3.1 Towards spontaneous individualism, against tribal goals

In his famous book *The Road To Serfdom* (1944), Hayek postulates that a relatively unencumbered market economy is the most efficient way to reach high

⁵ Hayek (1988, p. 133) remains ambiguously amoral when emphasizing that evolution will continue to select for the continuation of life regardless of our reasoned moral judgements on the matter. But his endorsement of an alleged fitness maximizing market morality reveals a clear normative preference for the maximization of life in relation to its minimization.

material prosperity while at the same time protecting individual liberties from considerable coercion. For him, the price system is uniquely capable of allocating resources in the most efficient way possible and of fomenting innovation through cooperation. To drastically interfere with this price system would cause considerable social and economic damage, eventually leading to the rise of coercive regimes aiming at social conformity. Hayek (1945) understands the market as a discovery process, where the law of supply and demand allows for a more efficient satisfaction of individually subjective preferences. Hence, to conserve the price system is to conserve the precious tacit information contained in it (Hayek, 2009), which allows for the cooperation between high numbers of individuals.

For Hayek (1988), a market morality is likely to be selected in the process of cultural group selection due to the survival benefits that it confers. This general idea leads him to assert that tribal goals are not suitable to survive in the evolutionary system. A tribal reaction against the market, which is analogous to the counter-reaction of the double movement, can therefore be regarded as maladaptive. Hayek claims that “the abstract society rests on learnt rules and not on pursuing perceived desirable common objects” (Hayek, 1979, p. 167). In other words, he considers that social groups should not pursue common goals or any specific notions of common good. Societies should instead rely on an individualistic spontaneous market order and on a decentralized transmission of knowledge.

In summary, Hayek (1988, p. 120) understands that our tribal tendencies exist as a legacy of our evolutionary past, but he believes that modern prosperity was possible precisely due to the suppression of these tribal instincts. For Hayek (1988, p. 74), any claims of justice based on a tribal morality are regarded as simply “inappropriate”, because, ultimately, evolution is not about justice. Hence, the countermovement of Polanyi’s double movement can be disregarded as inadequate to a naturalistic evolutionary process.

4. Polanyi vs Hayek: the affirmation of life and the double movement

Polanyi and Hayek ultimately understand the double movement as a process that leads to the affirmation and flourishing of life or to its negation and destruction. Hayek (1988) considers that market liberalism is a force for the expansion of life and that tribal counter-movements, with their sense of an objective common good, can only be destructive. Conversely, Polanyi (2001, p. 139) asserts that a market society inherently destroys life, family, environment and prosperous continuity. For him, it is precisely the counter-movements that can re-establish social sustainability by instituting a political system underpinned by non-economic goals, that is, based on social relations (Polanyi, 2001, p. 74).

It can be inferred that both authors affirm the importance of life, although prescribing different paths to affirm it. As the marketization of society is the first movement of the double movement, it is important to scrutinize if Hayek's market liberal postulate truly achieves the positive life expansion that he asserts. It is also important to understand if the tribal counter-movements are intrinsically destroyers of life expansion (as Hayek claims) or if they are vital for the continuation of life (as Polanyi asserts).

4.1 Hayek: is market liberalism an evolutionary winner?

Hayek's social theory of evolution is challenged by contemporary demographic evidence that reveals the negative effect that market liberalism can have on the expansion of life. As noted by several authors (Miller, 1989; Witt, 1994), the social groups who have the best market institutions are the ones with the lowest fertility and growth rates (e.g. western societies), while the areas of the world with rudimentary market institutions present high fertility and growth. As a result, less-developed economies are expanding their native populations while developed economies are contracting their own. The observed world population growth is almost exclusively a phenomenon of the developing world. Together, these countries "accounted for 97 per cent of this growth because of the dual effects of high birth rates and young populations" (Haub, 2012).

The evidence also seems to show that the contact with market effects slows down population growth even in the developing world (Plumer, 2013). This can

be mainly attributed to the contact with liberal culture propagated by soap operas, movies and other technologically cultural vehicles (Chong, Duryea, & Ferrara, 2012; Jensen & Oster, 2009). Even the Soviet Union's population "has more than doubled" (Miller, 1989, p. 315) during its existence, while Russian fertility rates drastically decreased after the introduction (in 1991) of a market oriented economy (Demoscope.ru, 2016).

Against Hayek's prediction, it does look like the institutionalization of market liberalism minimizes reproductive fitness instead of maximizing it. The market society seems to change habits and customs in ways that do not expand the life of social groups. Hence, Hayek's assertion that a market morality confers evolutionary advantages to groups in the process of evolutionary inter-group competition is unverified. Actually, non-liberal, religious and collectively oriented groups reveal much better results in terms of population expansion (Inglehart & Norris, 2011).⁶ As a consequence, his underestimation of the importance of tribal moral goals in the evolutionary process undermines his normative opposition to the protective social counter-movements.

4.2 Hayek: misunderstanding the tribal countermovement

As it is present in Hayek's model of cultural group selection, cultural practices and beliefs can be adaptive or maladaptive for groups. The effect of culture on biological evolution is currently studied under the logic of gene-culture co-evolution (Boyd & Richerson, 2005). Relevantly, evolutionary scientists Robert Boyd and Peter Richerson (2005, Chap. V) identifies current liberal western societies as being strikingly maladaptive. Although having an unprecedented amount of general wealth available, individuals from these societies do not seem to use these resources to have more children. Instead, liberal individualism means that other non-adaptive cultural traits are expanding, like the pursuit of professional careers and hobbies at the cost of having large families. Boyd and Richerson identify wealthy and successful individuals in liberal societies as promoters of liberal culture, leading the conformist masses to imitate maladaptive practices.

⁶ Islam, Catholicism, Hinduism and Protestantism are the main growing religions.

Hayek assumes that the drive for reproduction will always remain strong and that market liberalism's wealth production will allow groups to expand their numbers. Nonetheless, he neglects the importance of moral and cultural framing in promoting fertility. That is, preferences for high reproduction are not strictly innate but also influenced by a group's culture. The liberal satisfaction of preferences not only weakens the prestige of large families as a social goal, it also atomizes preferences in numerous directions. The Hayekian liberal discovery process means the discovery of preferences that are beyond reproductive fitness, making child bearing a mere preference that is not necessarily highly ranked in relation to others (Faria, 2017, pp. 316-317). Moreover, the pressures of market liberalism on (non-economic) social norms are substantial and the introduction of money in social relations has a strong capacity to crowd out, erase and replace traditional (fitness enhancing) social norms (Ariely, 2008, Chap. IV).

The research paradigm in experimental economics shows the importance that context and elite framers have in shaping individual preferences, especially because preferences are not well defined *a priori* (Kahneman & Tversky, 1981, 1984). Hayek's disregard for the active role of tribal altruism and human agency in shaping mass preferences underpins his failure to acknowledge the erosion that the liberal satisfaction of preferences can cause on the reproductive fitness of groups.

By aiming at embedding the economy in non-economic social relations and goals, the counter-movement (of the double movement) can be seen as a process that attempts to re-establish a moral framing that values life and social continuity above indeterminate economic goals. Put differently, the counter-movement is a protective attempt to realign preferences in ways that do not minimize group fitness. Hence, Hayek's rejection of the counter-movement ignores the importance that this force can have in the evolutionary process of inter-group competition. Although not coming from an evolutionary perspective, Polanyi's acknowledgement of the need for the counter-movement reveals that he understands the relevance of this counter-movement for the continuation of life.

4.3 Polanyi and the importance of re-embedding markets

If the counter-movement is an attempt to restore the sustainable balance of life, it is expected to happen on a regular basis. Hence, the re-embedding of markets in social relations becomes an inescapable life force. Moreover, both Hayek and Polanyi recognize this historical trend towards social protection (Lindsay, 2015, p. 385). Hayek acknowledges that the disturbance of the market on our social life “undoubtedly offends our sense of justice” and that demands for social protection “are certain to receive popular sympathy and support” (Hayek, 1944, p. 127). Nonetheless, although favouring a temporary governmental safety net (Hayek, 2011, p. 424), Hayek asserts that it is necessary for societies to endure hardship for the sake of individual liberty and of a better economic future (Hayek, 1944, p. 137). Discordantly, Polanyi sees Hayek’s claim not only as destructive of the social fabric but also outright unrealistic. In his own words:

to expect that a community would remain indifferent to the scourge of unemployment, the shifting of industries and occupations and to the moral psychological torture accompanying them, merely because economic effects, in the long run, might be negligible, was to assume absurdity (Polanyi, 2001, p. 224).

Most importantly, Polanyi’s main opposition to free markets does not come from doubting their economic efficiency but from understanding their incompatibility with the sustainable life of social groups. As Peter Lindsay notes, “groups respond fearfully to markets (...), free markets became humanly (rather than economically) unsustainable, irrespective of all they might offer” (Lindsay, 2015, p. 385). Polanyi recognizes that whatever amount of wealth market liberalism brings, it cannot do it without eroding folk norms and ways of being that underpin the perpetuation of life. By commodifying land and labour, market liberalism erodes the sacredness of norms that aim at protecting stability and collective survival. Relevantly, Polanyi’s understanding of (bio) stability emphasizes the role of land, kinship and collective identity. As he explains it:

Land is tied up with the organization of kinship, neighbourhood, craft and creed – with tribe and temple, village, guild, and church (...) It invests a man’s life with stability; it is the site of his habitation; it is a condition of his physical safety; it is the landscape and the seasons (Polanyi, 2001, p. 187).

Polanyi's focus on the importance of kinship and of a tribal moral framework that subjugates material elements (like land) to the sustainable continuity of collective life is evolutionarily significant. Especially from a group selection perspective (Sober & Wilson, 1998), the one Hayek takes. Conversely, Hayek deflates the evolutionary relevance of an extended tribal kinship in modern market society. Yet, he claims that practices that defend the family and private property are universally present in all lasting religions of the world and that natural selection selects these cultural practices due to their fitness maximizing potential (Hayek, 1982, p. 4). Nonetheless, Polanyi (2001, p. 139) is correct when he points out that market liberalism has a devastating effect on the family. Hayek does not entirely perceive the clear tension between an individualistic market order and the family structure.

The most advanced market economies of the western world (with sub-replacement fertility levels) have unsurprisingly seen an erosion of the institution of the family. Noticeably, Europe reveals increasing divorce rates and decreasing marriage rates (Eurostat, 2016). Polanyi has predicted some of the main reasons for this phenomenon. Due to the commodification of labour and land, individuals deal with permanent job uncertainty, unemployment and mass dislocations provoked by the search for career opportunities. Furthermore, children often hurt career prospects that determine survival in a competitive market. Child-care can be costly and generations (grandparents, sons and grandsons) are often separated from the same geographical place due to market pressures on mobility. But most importantly, market liberalism is not simply an apology of a market system, it is a moral framework for the empowerment of the individual. What is good for the individual is not always good for the family, in the same way that what is good for the family is not always good for society or tribe. There are logical conflicts of interests between these levels.

Moreover, by appreciating the family as a reproductive engine while disregarding the importance of a tribal wider collective, Hayek underestimates the extent to which the family structure is an organizational construction of the tribe/society. Given that the family is not simply an autonomous institution detached from collective organization, a rejection of moral tribalism can have a detrimental impact on the structure of the family. Hence, by emphasising the

importance of the social whole for the sustainable continuation of life, Polanyi's historical and sociological approach also attains robustness from an evolutionary perspective.

4.4 Polanyi's evolutionary double movement

In order for a social group to be protected from the negative effects of market liberalism, markets need to be embedded in fitness maximizing social norms. These norms have the function of preventing the atomizing market effects while allowing for the existence of a system of production and distribution that serves social interests and the continuation of collective life. That is, as Polanyi upholds, non-economic social norms must be strong and vigilant in order to make sure that society is not run as "an adjunct to the market" (Polanyi, 2001, p. 60).

Although not taking an explicitly evolutionary approach as Hayek does, Polanyi's postulates have a strong evolutionary resonance. For instance, his focus on the importance of organizational kinship and tribe highlights the importance of inclusive fitness (Hamilton, 1963). Inclusive fitness explains that altruism evolves when individuals are altruistic towards closely related others because these others are likely to share altruistic genes (which fosters adaptive social cooperation). This phenomenon can also maximize fitness because it helps to pass on a higher degree of one's own genes. Polanyi seems to understand how extended kinship is relevant for the evolution of altruistic and cooperative groups. This contrasts with Hayek's narrow defence of the family, who, by rejecting tribal and social goals, regards the family as the last level of collective deliberation in social organization.

Polanyi's insights are also relevant from a group selection evolutionary perspective, which is the model that Hayek uses as an explanatory framework. The central idea of group selection (Sober & Wilson, 1998) is that natural selection does not simply operate at the individual (or gene) level but also at the group level. Hence, groups are understood as units of selection in competition with other groups. In this process, the groups that are more altruistic and cooperative beat groups that are less cohesive due to the latter being plagued with widespread egoist individualism. In order for groups to achieve high levels of

cooperation and for altruism to evolve, they are required to have systems of punishment against free riders and egoists. That is, against individuals that will pursue their self-interest regardless of the interests of the collective.

For Polanyi, politics and non-economic collective goals are indispensable for the organization of society. As Fred Block and Margaret Somers note, Polanyi's overall point is "that in a complex society we cannot escape the necessity of politics and governmental coordination of economic and social life" (Block & Somers, 2014, p. 11). Unlike Hayek and his scepticism of politics, Polanyi understands that an individualistic market system run on the basis of individual interests is ultimately dysfunctional for social groups. From a group selectionist perspective, only the political can enforce legal and cultural systems that punish free riding and egoism. While the market process is a form of cooperation that relies mostly on the virtue of self-reliance, it is groupless and morally individualistic, which leaves the door open to a liberal satisfaction of preferences that neglects the social whole, potentially fostering the evolution of egoism.

Because the evolution of egoism is detrimental for the adaptiveness of groups (Wilson & Wilson, 2007), it contributes decisively to the erasure of a group. This is in line with what Polanyi (2001, p. 3) concludes: a market system based on individual self-interest creates disruptions in the social order that ultimately leads to the collapse of that social collective. Due to culture and morality being as much under natural selection as genes (Boyd & Richerson, 2005), cultural group selection selects cultural practices and beliefs that are evolutionarily adaptive, that is, that have survival value for groups. In this case, if free markets facilitate the spread of egoistic anti-group elements and provoke disruptions in the social fabric that negatively affect the physical and cultural continuity of groups (e.g. low reproductive rates), it is to be expected that market liberalism will be naturally selected against. Thus, the counter-movements of Polanyi's double movement seem to be the inevitable attempt to rescue societies from their existential decline.

Nonetheless, the creation of wealth via the market system can be a positive evolutionary force. It is indeed the case that the European population exponentially increased its numbers after the event of the Industrial Revolution (Piketty, 2014, p. 79). Yet, Polanyi's insights on embeddedness can help to explain

the process behind this demographic phenomenon. While the market created enough wealth to allow for a higher number of children to be raised, culture and morality were still traditional. That is, the markets were still embedded in traditional social relations. Since then, market liberalism became internalised within the culture and morality of the west and the result was a relative demographic decline and the erosion of folkways. In Polanyi's (2001, p. 60) language, western societies are now managed as adjuncts to the market. Put differently, the "Hayekian" market morality rose to be the meta-morality, therefore disembedding markets from their traditionally subaltern and utilitarian social position.

Polanyi's support for the embedment of markets in non-economic social relations and in collectively deliberated political goals emphasises the importance of stability, regularity and meaning of land and folkways. This stability and regularity are of high relevance to the evolution of altruism in a given moral framework and to the creation of cohesive and meaningful social groups. Moreover, as Charles Darwin (1871, p. 166) postulates, these cohesive and altruistic groups also have evolutionary advantages when in direct competition with other groups via warfare.⁷

Nevertheless, unlike Hayek, Polanyi is not a thinker known for emphasizing competition. On the contrary, his most famous book *The Great Transformation* (2001) was written with the intention of showing that market liberalism was the primary cause of the devastating world wars of the 20th century. Still, from an evolutionary perspective, these negative consequences of free markets are at best a proximate cause, not the ultimate one. Group conflict has deep evolutionary roots that predate market liberalism (Shaw & Wong, 1989; Turchin, 2015).

⁷ Charles Darwin described the process now known as group selection in the following way:

"There can be no doubt that a tribe including many members who, from possessing in a high degree the spirit of patriotism, fidelity, obedience, courage, and sympathy, were always ready to aid one another, and to sacrifice themselves for the common good would be victorious over most other tribes; and this would be natural selection" (Darwin, 1871, p. 166).

5. Conclusion

Polanyi and Hayek both theorize about the phenomenon of double movement and come on opposite sides of the equation. They both affirm that the continuation of social life would be better served through the support of distinct movements. Hayek supports the movement towards market liberalism while Polanyi favours the counter-movement that aims at re-embedding markets in stable social relations.

Nonetheless, Hayek's evolutionary normative reading of the double movement reveals weaknesses concerning its main claim: the claim that free markets expand the life of social groups, while protective tribalism does not. Polanyi's normative interpretation of the double movement is more robust because, unlike Hayek, he recognizes the need for life affirming non-economic collective social norms. Polanyi also understands the negative effects that market liberalism can have on the reproduction and continuity of social life. Namely, that market liberalism erodes social norms of non-economic cooperation and institutional norms of reproduction that are important for the continuation of life in social groups.

Hayek's rejection of tribal goals as maladaptive for groups overlooks the importance that these goals have in protecting societies from the atomization of preferences and social disintegration. Because preferences are not well defined *a priori*, common social and political group norms set the standard that shape desirable behaviour. Most importantly for the continuation of life, even preferences for reproduction are shaped by culture and politics. These social and political group norms are essential to shape vital preferences, such as the general preference for the reproductive continuity of life. Polanyi accurately perceives that non-economic social and political norms have the important function of preventing that market goals prevail over social goals, while understanding the relevant role of markets in producing and allocating commodities. From an evolutionary perspective, this means that it is important to prevent market liberalism from becoming society's meta-morality in order to counter the market driven erosion of fitness maximizing norms.

Polanyi's assertion of the importance of environmental stability and of non-economic common goals in society also has an evolutionary relevance, especially from a "Hayekian" group selection perspective. Namely, environmental stability and common goals allow for the evolution of altruism within groups, conferring cohesion to a given collective in inter-group competition through the selection of pro-group elements and the rejection of rational egoists (e.g. free riders). Although Polanyi is aware that the exaltation of the political and of the counter-movements can lead to harsh conflicts, he nevertheless correctly perceives that the political realm cannot be abolished because it serves important functions of social coordination. Conclusively, even though Polanyi is not an explicit evolutionary thinker, he understands that political goals have an important role in promoting the continuation of social life and that, in the end, cooperation has to go well beyond the market.

Unlike Polanyi who focuses on cooperative stability, Hayek is an author that normatively highlights the importance of economic competition and evolutionary selection in social life. But paradoxically, from a group selection perspective, Polanyi's interpretation of the double movement is more robust than Hayek's evolutionary interpretation. This is not to claim that Polanyi's whole theory and work is evolutionarily solid, but that the normative understanding of the double movement is more robust than Hayek's one, even when analysed from a "Hayekian" group selection framework.

References

Ariely, D. (2008). *Predictably Irrational: The Hidden Forces That Shape Our Decisions*. New York: HarperCollins Publishers.

Beck, N. (2011). Be Fruitful and Multiply: Growth, Reason, and Cultural Group Selection in Hayek and Darwin. *Biological Theory*, 6(4), 413-423.

Beck, N. (2015). The Garden of Orderly Polity: F. A. Hayek and T. H. Huxley's Views on Social Evolution. *Journal of Bioeconomics*, 17(1), 83-96.

Block, F., & Somers, M. R. (2014). *The Power of Market Fundamentalism: Karl Polanyi's Critique*. Cambridge: Harvard University Press.

Boyd, R., & Richerson, P. (2005). *Not by Genes Alone: How Culture Transformed Human Evolution*. Chicago: The University of Chicago Press.

Chong, A., Duryea, S., & Ferrara, E. (2012). Soap operas and fertility: Evidence from Brazil. *American Economic Journal: Applied Economics*, 4(4), 1-31.

Dale, G. (2010). *Karl Polanyi: The Limits of the Market*. Cambridge: Polity Press.

Darwin, C. (1871). *The Descent of Man and Selection in Relation to Sex*. New York: Appleton.

Demoscope.ru. (2016). Russian Birth Rates 1950-2013. Retrieved from <http://demoscope.ru/weekly/ias/ias05.php?tim=0&cou=26&terr=1&ind=26&Submit=OK>, Accessed 24 October 2016.

Eurostat. (2016). Marriage and Divorce Statistics. Retrieved from http://ec.europa.eu/eurostat/statistics-explained/index.php/Marriage_and_divorce_statistics, Accessed 26 January 2017.

Faria, F. N. (2017). Is Market Liberalism Adaptive? Rethinking F. A. Hayek on Moral Evolution. *Journal of Bioeconomics*, 19(3), 307-326. doi:10.1007/s10818-017-9251-1

Hamilton, W. D. (1963). The Evolution of Altruistic Behavior. *American Naturalist*, 97, 354-356.

Haub, C. (2012). Fact Sheet: World Population Trends 2012. Retrieved from <http://www.prb.org/Publications/Datasheets/2012/world-population-data-sheet/fact-sheet-world-population.aspx>, Accessed 24 October 2016.

Hayek, F. A. (1944). *The Road to Serfdom*. London: Routledge.

Hayek, F. A. (1945). The use of knowledge in society. *The American Economic Review*, 35(4), 519-530.

Hayek, F. A. (1979). *Law, Legislation and Liberty; Volume 3: The Political Order of a Free People*. Chicago: The University of Chicago Press.

Hayek, F. A. (1982). *Our Moral Heritage*. Washington: The Heritage Foundation.

Hayek, F. A. (1988). *The Fatal Conceit: The Errors of Socialism*. Chicago: University of Chicago.

Hayek, F. A. (2009). *The Pure Theory of Capital*. Auburn: The Ludwig von Mises Institute.

Hayek, F. A. (2011). *The Constitution of Liberty: The Definitive Edition* (Vol. 17). Oxford: Routledge.

Inglehart, R., & Norris, P. (2011). *Sacred and Secular: Religion and Politics Worldwide*. Cambridge: Cambridge University Press.

Jensen, R., & Oster, E. (2009). The power of TV: Cable television and women's status in India. *The Quarterly Journal of Economics*, 124(3), 1057-1094.

Kahneman, D., & Tversky, A. (1981). The framing of decisions and the psychology of choice. *Science*, 211(4481), 453-458.

Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *American Psychologist*, 39(4), 341-350.

Lindsay, P. (2015). Polanyi, Hayek, and the Impossibility of Libertarian Ideal Theory. *Polity*, 47(376-396).

Machado, N. M. C. (2011). Karl Polanyi and the New Economic Sociology: Notes on the Concept of (Dis)embeddedness. *RCCS Annual Review [Online]*(3).

Miller, D. (1989). The fatalistic conceit. *Critical Review: A Journal of Politics and Society*, 3(2), 310-323.

Patomaki, H. (2014). On the Dialectics of Global Governance in the Twenty-first Century: A Polanyian Double Movement? *Globalizations*, 11(5), 733-750.

Piketty, T. (2014). *Capital in the Twenty-First Century*. Cambridge, Massachusetts: The Belknap Press of Harvard University Press.

Plumer, B. (2013). Why Are Birthrates Falling Around the World? Blame Television. Retrieved from

<http://www.washingtonpost.com/news/wonkblog/wp/2013/05/13/why-are-birthrates-falling-around-the-world-in-a-word-television/>, Accessed 24 October 2016.

Polanyi, K. (1957). Aristotle Discovers the Economy. In C. Arensberg, H. Pearson, & K. Polanyi (Eds.), *Trade and Market in the Early Empires* (pp. 64-94). Chicago: Henry Regnery.

Polanyi, K. (2001). *The Great Transformation*. Boston: Beacon Press Books.

Shaw, P., & Wong, Y. (1989). *Genetic Seeds of Warfare: Evolution, Nationalism, and Patriotism*. Boston: Unwin Hyman.

Sober, E., & Wilson, D. S. (1998). *Unto Others: The Evolution and Psychology of Unselfish Behavior*. Cambridge, Mass.: Harvard University Press.

Turchin, P. (2015). *Ultrasociety: How 10,000 Years of War Made Humans the Greatest Cooperators on Earth*. Connecticut Beresta Books.

Wilson, D. S., & Wilson, E. O. (2007). Rethinking the theoretical foundation of sociobiology. *The Quarterly Review of Biology*, 82(4), 327-348.

Witt, U. (1994). The Theory of Societal Evolution. Hayek's Unfinished Legacy. In J. Birner & R. Zipp (Eds.), *Hayek, Co-ordination and Evolution. His Legacy in Philosophy, Politics, Economics and the History of Ideas* (pp. 178-189). London: Routledge.



MITO, UTOPIA E DEMOCRACIA EM A GRANDE TRANSFORMAÇÃO DE KARL POLANYI

MYTH, UTOPIA, AND DEMOCRACY IN KARL POLANYI'S *THE GREAT TRANSFORMATION*

Patrícia Fernandes*
patriciafernandes@protonmail.com

Sumário. Partindo do tripé conceptual mito, utopia e democracia, o nosso texto propõe-se considerar os contributos de Karl Polanyi em três linhas de reflexão. A primeira incide sobre o seu trabalho de desconstrução do mito da naturalidade do mercado e da natureza económica do ser humano; a segunda concerne a sua defesa do projeto de uma sociedade de mercados livres como utópico; e a terceira considera o seu princípio do duplo movimento como uma expressão da democracia europeia. Estas três linhas de reflexão constituem contributos relevantes para o processo, ainda em curso, de análise, interpretação e ponderação de soluções para a crise económica e financeira iniciada em 2007/8 e que, tendo-se transmutado em crise das dívidas soberanas, se traduziu na adoção do austeritarismo como projeto político.

Palavras-chave: mito, utopia, duplo movimento, democracia, políticas de austeridade.

Abstract. Considering the concepts of myth, utopia and democracy, I shall center my analysis on three of Karl Polanyi's contributions: the first of them focuses on the deconstruction of the myth of natural markets and of the economic nature of human beings; the second one centers on Polanyi's defense of a market society as utopic; and the last one considers his double movement formulation as an expression of the European democracy. Ultimately I shall use these contributions to reflect on the financial and economic crisis of 2007/8 and its transformation in the sovereign debts crisis that led to the adoption of Austerity as a political project.

Keywords: myth, utopia, double movement, democracy, austerity politics.

* Investigadora associada do CEPS (UM); Professora auxiliar Convidada na UBI.

o. Introdução

O maior elogio que podemos dirigir a uma obra, que é necessariamente resultado de condições sociais e históricas específicas, parece ser o reconhecimento de que ela nos proporciona novas ou repetidas reflexões em contextos posteriores e distintos daqueles que a motivaram, ajudando-nos a compreender o presente e a futurar alternativas para os desafios que enfrentamos. O livro de Karl Polanyi, *A Grande Transformação*, publicado em 1944, é precisamente uma dessas obras, tendo sido especialmente retomada depois da recente crise económico-financeira que gerou graves consequências um pouco por todo o mundo. Recuperar não só *A Grande Transformação* como também *O Caminho para a Servidão*,¹ publicado por Friedrich Hayek no mesmo ano, constitui, nesse sentido, um valioso contributo para a tarefa de refletir sobre os desafios dos nossos dias.²

A virtude desta dupla abordagem de reflexão prende-se com o facto de estes dois autores promoverem uma leitura contrastante sobre os factos que antecederam a crise política de 1933-45. Oferecendo essa leitura distinta, complementam-se na apreciação dos factos, permitindo-nos examinar a complexa realidade de forma mais abrangente. Podem, dessa forma, ser usados como instrumentos de reflexão recíproca entre aqueles que, por um lado, analisam a crise atual como resultado de um processo de des-regulação e diminuição da intervenção estatal e aqueles que, por outro, a interpretam como resultado da pesada herança keynesiana que impossibilita uma prossecução adequada dos objetivos do projeto recuperado do liberalismo clássico.

Escrever sobre tudo isto em 2017, isto é, precisamente dez anos depois dos primeiros sinais que indicaram a grave crise que se viria a revelar plenamente em 2008, é escrever com o reconhecimento de dois distintos fatores: se os dez anos nos permitem já algum distanciamento em relação aos acontecimentos, e, por isso, uma reflexão mais amadurecida, continuamos simultaneamente presos às consequências daqueles factos e por isso refletindo sobre eles a partir da sua vivência e sem sabermos como, no fundo, termina a história. Estes dois fatores

¹ Hayek, Friedrich (2014). *O Caminho para a Servidão*. trad. port. Marcelino Amaral. Lisboa: Edições 70.

² Viriato Soromenho-Marques coloca igualmente em diálogo estes dois autores no livro em que analisa as condições do projeto europeu após a crise económico-financeira, publicado em 2014: *Portugal na queda da Europa*. Lisboa: Temas e Debates – Círculo de Leitores.

tornam ainda mais relevante uma obra como a de Polanyi: considerando os pontos de coincidência entre a crise do início do século XX e a crise atual, as reflexões cruciais de Polanyi detêm a pertinência analítica de nos permitir apreciar o momento presente e, ao mesmo tempo, considerar os seus perigos e divisar saídas possíveis. É isso que procuraremos fazer, partindo do seguinte tripé conceptual: mito, utopia e democracia.

1. Mito e utopia

O primeiro contributo da obra de Karl Polanyi que queremos recuperar prende-se com as observações que resultaram do seu minucioso estudo histórico e jurídico sobre o surgimento da lógica dos mercados e do *homo economicus*. Abordando o surgimento e a consolidação da doutrina do liberalismo económico, Polanyi diz-nos: “A despeito de toda a insistência do coro de fórmulas académicas hipnóticas do século XIX, o ganho e o lucro obtidos através da troca nunca desempenharam, no passado, um papel importante na economia humana.” (Polanyi, 2012: 175) Na realidade, os estudos históricos e etnográficos apresentados por Polanyi revelam-nos que não há nada *a que nos possamos agarrar* para defender a ideia de mercado como condição natural. No mesmo sentido, os elementos disponíveis apontam para a impossibilidade de se afirmar uma natureza humana de cariz económico:

Um pensador da envergadura de Adam Smith sugeriu que a divisão do trabalho na sociedade dependia da existência de mercados, ou, nos seus termos, da “propensão [do homem] para negociar, permutar ou trocar umas coisas por outras.” Estas palavras serviriam para fundamentar mais tarde a conceção do Homem Económico. (Polanyi, 2012: 176)

No entanto,

[e]m retrospectiva, poderia dizer que nenhuma outra falsa interpretação do passado se revelaria mais profética a respeito do futuro. Porque, se até à época de Adam Smith essa propensão não se manifestara de modo muito considerável na vida de qualquer comunidade observável, e se mantivera, quando muito, como um traço subordinado da vida económica, cem anos mais tarde funcionava já em pleno na maior parte do planeta um sistema industrial, implicando, em teoria e na prática, que a espécie humana passara a ser dirigida em todas as suas atividades económicas – se é que não, também, em todos os seus esforços políticos, intelectuais e espirituais – por essa propensão particular. (Polanyi, 2012: 176)

Polanyi caracteriza estas pretensões do pensamento liberal económico como *o mito da naturalidade dos mercados e da natureza económica do ser*

humano. O nosso entendimento é o de que essas pretensões têm as suas raízes no liberalismo filosófico que lançou as bases da sociedade atual. Aquilo que os primeiros autores liberais, contratualistas, tentaram foi encontrar as condições naturais do ser humano e da sociedade, por forma a apresentarem uma nova articulação política para a sociedade moderna: recordemos como esse pensamento inicial se centrava nas noções de estado de natureza, direitos naturais e, a partir daí, na ideia de contrato e sociedade civil. Trata-se de uma estratégia discursiva particularmente eficaz porque repete, como Thomas Hobbes procurou fazer à imagem da geometria euclidiana, o discurso científico que dava na altura os seus primeiros passos. Era a forma de encontrar os fundamentos últimos, objetivos e incontestáveis, capazes de assegurar uma argumentação racionalmente inatacável. Afinal, como contestar política ou economicamente aqueles que apenas procuram estabelecer *o modo natural como as sociedades se devem organizar*? Se esse é o *modo natural* é porque se trata do *modo correto* e por isso as políticas só podem ter como objetivo promovê-lo.

Esta poderosa estratégia discursiva pode, contudo, ser atacada recorrendo às armas usadas por Polanyi, isto é, chamando à colação os contributos das áreas de estudo que introduzem o fator da *contingência*, como a história, a antropologia ou a etnografia, o que permite convocar a ideia de que a naturalidade dos mercados ou a natureza económica do ser humano não passa de um *mito* – no sentido de uma construção fictícia e discursiva. E é o que Polanyi faz no capítulo IV da segunda parte do seu livro, recorrendo a contributos daquelas áreas de estudo para mostrar que nada na informação de que dispomos nos permite afirmar o mercado ou a natureza económica como condição natural da organização social humana. E este é o primeiro nível de atuação do argumentário de Polanyi.

O objetivo político desta construção discursiva consistia em legitimar políticas de restrição do Estado enquanto agente económico. Em última instância, o mercado, como condição natural, seria autorregulado: regulado pelas suas leis naturais e não carecendo da intervenção estatal que, pelo contrário, perturbaria aquelas leis. Ora, a desconstrução mitológica de Polanyi atua também neste segundo nível: ao longo do capítulo V, a narrativa do nosso autor procura expor de que forma os tais mercados naturais são, em oposição à própria

coerência ideológica dos seus defensores, resultado de intervenção estatal. Na verdade, o Estado funcionou “como instrumento da ‘nacionalização’ do mercado e criador do mercado interno.” (Polanyi, 2012: 205) Longe de se tratar de uma condição natural, o sistema de mercado tal como viria a ser desenvolvido pelos economistas liberais resultou de um projeto político do Estado moderno. A ideia de mercado natural, livre e autorregulado perde, assim, sentido. Como diz Polanyi,

O *laissez-faire* nada tinha de natural; os mercados livres nunca teriam chegado a existir por simples efeito da força das coisas. Tal como as manufaturas do algodão – a principal indústria associada à liberdade de comércio – tinham sido criadas mediante o auxílio de tarifas protecionistas, de incentivos à exportação e de subsídios salariais indiretos, também o próprio *laissez-faire* foi imposto pelo Estado. (Polanyi, 2012: 308-9)

Contra a ideia de naturalidade dos mercados, contra a ideia de que bastaria suprimir a intervenção estatal para encontrar, por debaixo dessa crosta, condições naturais de mercado, Polanyi constata:

[A] introdução dos mercados livres, longe de pôr fim às exigências de controle, regulação e intervenção, expandiu imenso o seu domínio. Os administradores tinham de se manter a todo o momento vigilantes para assegurar o livre funcionamento do sistema. Assim, até mesmo aqueles que desejam mais ardentemente libertar o Estado de todas as obrigações desnecessárias e cuja filosofia reclamava em todos os aspetos a limitação das atividades do Estado, não puderam fazer outra coisa que não fosse atribuir a esse mesmo Estado os novos poderes, órgãos e instrumentos requeridos pela instauração do *laissez-faire*. (Polanyi, 2012: 310-1)

Paradoxalmente, o próprio valor de *liberdade* advogado pelos liberais³ carece desse *intervencionismo*:

A via do mercado livre foi aberta e manteve-se aberta graças ao aumento constante de um intervencionismo organizado e controlado centralmente. Tornar a “simples e natural liberdade” de Adam Smith compatível com as exigências de uma sociedade humana revelava-se uma tarefa extremamente intrincada. Testemunham-no a complexidade das disposições das inúmeras leis sobre os *enclosures*; o enorme controle burocrático associado à administração das Novas Leis dos Pobres, que, pela primeira vez desde o reinado da rainha Isabel, eram efetivamente supervisionadas pela autoridade central, ou o crescimento da administração governamental suscitado pela meritória iniciativa de uma reforma municipal. E contudo, todas estas praças-fortes da interferência governamental eram construídas com o propósito de organizar esta ou aquela simples liberdade – a da terra, do trabalho ou da administração municipal. (Polanyi, 2012: 310)

³ No texto, de cariz privado, escrito por Murray Rothbard sobre o livro de Polanyi em análise, «Down With Primitivism: A Thorough Critique of Polanyi» (1961), o economista afirma: “Referi que a sociedade livre permitiria a Polanyi ou a qualquer outra pessoa que concorde com ele abandonar o mercado e encontrar qualquer outro modelo que lhes sirva. Mas *uma e única coisa* a sociedade livre *não* permitiria a Polanyi fazer: usar de coerção em relação ao resto de nós.” (cf. <https://mises.org/library/down-primitivism-thorough-critique-polanyi> - último acesso: 16/07/2017). Esta é a contradição permanente daqueles que defendem o mercado livre esquecendo que ele resulta, e só resulta, de decisão estatal, que é sempre uma forma de coerção.

Avançemos agora para o segundo dos nossos conceitos de análise: a *utopia*. A par desta desconstrução mitológica, Polanyi considera a dimensão utópica do projeto liberal: “A nossa tese é que a ideia de um mercado capaz de se auto-ajustar era uma mera utopia.” (Polanyi, 2012: 120) Podemos ter em conta este aspeto considerando duas perspetivas. Por um lado, Polanyi chama a atenção para as ideias-chave dos pensadores liberais que destacam a sociedade de mercado como projeto político a prosseguir: “o mercado autorregulado resultava das leis inexoráveis da natureza” pelo que “a libertação sem freio do mercado era uma necessidade inelutável.” (Polanyi, 2012: 290) Destaca, desta forma, a sociedade de mercado como projeto utópico que deve orientar a atividade e as decisões políticas, juntando-se à série de utopias que marcariam a modernidade com os seus resultados calamitosos.⁴

Reconhecendo este aspeto, a análise de Polanyi pretende, por outro lado, considerar a sociedade de mercado como *u-topia*, isto é, o seu não lugar, a sua impossibilidade de realização. Recorre para esse efeito a Robert Owen para notar que “a economia de mercado, se a deixassem evoluir segundo as suas próprias leis, criaria grandes males permanentes.” (Polanyi, 2012: 298) A argumentação do nosso autor centra-se, em especial, na tese das mercadorias fictícias, apresentada no capítulo VI: é inerente ao sistema de mercado alargar progressivamente a sua lógica de mercantilização, e isso implica transformar em mercadorias todos os fatores de produção, incluindo aqueles que não são mercadorias – afinal, “uma economia de mercado só pode existir numa sociedade de mercado” (Polanyi, 2012: 214). Isto levou à criação de uma categoria, que Polanyi designa como *mercadorias fictícias*, por forma a incluir no processo aqueles elementos essenciais da indústria que não são mercadorias: o trabalho, a terra e a moeda. O ponto de Polanyi é o de que esta ficção não atendeu ao facto de que deixar os seres humanos e a terra nas mãos do mercado equivale a aniquilá-los.⁵

⁴ Jens Beckert, que contribuiu para o campo da sociologia económica com uma poderosa crítica à ideia de eficiência económica (em *Grenzen des Marktes: die sozialen Grundlagen wirtschaftlicher Effizienz*, 1997), tem desenvolvido o seu trabalho mais recente em torno do papel fulcral das ficções no domínio económico e de como os agentes económicos são motivados por um futuro imaginado nas decisões em situação de risco (*Imagined Futures: fictional expectations and capitalist dynamics*, 2016).

⁵ Zygmunt Bauman, em *Liquid Modernity* (2000), faz corresponder a esta transformação a ideia de “melting the solids”, pela destruição dos vínculos comunitários, dos costumes e regras consuetudinários e dos poderes intermédios, que seriam substituídos por *realidades derretidas e fluídas* colocadas em novos moldes de acordo com a lógica do “novo começo” que marca a modernidade (p. 143).

São os movimentos de reação e proteção às terríveis perturbações que resultaram da tentativa de imposição da lógica de mercado a essas mercadorias fictícias que pararam a concretização do sistema de mercado. E é neste sentido que Polanyi destaca a dimensão utópica do projeto liberal: apesar de, a partir da década de 1830, “o liberalismo explodir numa cruzada apaixonada e o *laissez-faire* tornar-se um credo militante” (Polanyi, 2012: 305), as consequências devastadoras daquelas políticas económicas provocaram reações de proteção da sociedade que tornaram impossível a concretização de uma sociedade de mercado, que a tornaram *u-tópica*. Mas para que tal ideia se torne mais clara, teremos de avançar com o princípio do duplo movimento.

2. O princípio do duplo movimento

A terceira parte do livro abre com a seguinte ideia: “Durante um século, a dinâmica da sociedade moderna foi governada por um duplo movimento: o mercado expandiu-se continuamente, mas esse movimento era contrabalançado por um contramovimento a controlar a expansão em determinadas direções.” (Polanyi, 2012: 297) A descrição empírica que Karl Polanyi faz destes princípios organizadores é particularmente relevante para uma apreciação da dinâmica social: o primeiro desses movimentos foi por nós descrito nas páginas precedentes e corresponde ao liberalismo económico, “visando a instauração de um mercado autorregulado, contando com o apoio das classes ligadas ao comércio e utilizando em larga medida como métodos o *laissez-faire* e o livre-cambismo”; o segundo “era o princípio da proteção social visando a conservação do homem e da natureza bem como da organização da produção, contando com o apoio variável dos mais imediatamente atingidos pela ação nociva do mercado.” (Polanyi, 2012: 300) Como tomava forma essa proteção social? Através de legislação protetora, associações restritivas e outros instrumentos de intervenção.

A análise que ocupa toda a terceira parte do livro centra-se num ponto fulcral. O autor apresenta-nos este segundo movimento, de proteção, como tendo sido espontâneo: “A ação legislativa de primeira linha do contramovimento perante a autorregulação do mercado, tal como se desenvolveu no meio século que se seguiu a 1860, revelava-se como tendo sido espontânea, não dirigida pela

opinião e adotada segundo uma inspiração puramente pragmática.” (Polanyi, 2012: 311) Ou na sua célebre formulação, “[o] *laissez-faire* fora planeado – a intervenção do plano, não”. (Polanyi, 2012: 311) Importa considerar a razão para esta posição.

A espontaneidade constitui um elemento fundamental na medida em que Polanyi pretende recusar a ideia de uma conspiração antiliberal. Na verdade, a argumentação liberal face ao fracasso económico do seu projeto assenta na ideia de que ele resultou

da incapacidade por parte do homem de se manter fiel à inspiração dos primeiros liberais; que a generosa iniciativa dos nossos antepassados foi frustrada pelas paixões do nacionalismo e da guerra de classes, dos interesses estabelecidos e dos monopolistas, e, acima de tudo, pela cegueira dos trabalhadores aos benefícios que a liberdade económica sem limitações comportaria para todos os interesses humanos, entre os quais os deles próprios. (Polanyi, 2012: 315-6)

No fundo, “[u]m progresso intelectual e moral imenso foi assim, segundo os liberais, frustrado pela fraqueza intelectual e moral da massa da população – uma grande realização do espírito das Luzes viu-se destruído pelas forças do egoísmo.” (Polanyi, 2012: 316)

Face a esta alegação, Polanyi defende que o movimento de proteção da sociedade resultou antes da “perceção das fraquezas e dos perigos inerentes a um sistema de mercado autorregulado.” (Polanyi, 2012: 317) E usa, para o efeito, quatro linhas de argumentação:⁶ por um lado, contra a existência de uma conspiração antiliberal, Polanyi recorda a diversidade de campos de ação em que o contramovimento ocorreu: a legislação incidiu sobre a previsão de analistas dos bens alimentares e bebidas, a inspeção das fábricas de gás, penalizações para o emprego de crianças com menos de doze anos e que não frequentassem a escola, a imposição de vacinação, a fixação de tarifa para o aluguer de meios de transporte, etc.; em segundo lugar, Polanyi destaca as decisões tomadas de modo improvisado e pragmático, como a legislação aprovada sobre o Ato de Compensação dos Trabalhadores, que ficou dever-se exclusivamente “à evolução das condições em que o problema passava a pôr-se e uma solução a ser procurada para ele” (Polanyi, 2012: 319); por outro lado, Polanyi assinala que as fases de livre-câmbio e *laissez faire* seguidas de períodos de legislação antiliberal em

⁶ Cf. páginas 317 e ss.

matérias como a saúde pública, a segurança social, as condições de trabalho, etc., ocorreram em vários países europeus o que pressuporia uma conspiração iliberal internacional dificilmente defensável.

Notemos que, apenas dando sentido à ideia de espontaneidade e recusando a versão da conspiração coletivista, pode Polanyi defender o carácter utópico do projeto liberal nos termos que apresentamos na secção anterior: a tese central do nosso autor é a de que o projeto liberal é *impossível nos seus próprios termos*, suportando-a na consideração de que ele próprio gera uma reacção espontânea. Por isso, a última linha do seu argumentário é a mais poderosa:

Não seria possível apresentar prova mais concludente da inevitabilidade do recurso a métodos antiliberais ou “coletivistas” nas condições da moderna sociedade industrial do que o facto de os próprios adeptos do liberalismo económico terem regularmente recorrido a métodos desse tipo, aplicando-os em domínios decisivamente importantes da organização industrial. (Polanyi, 2012: 321-2)

São os problemas gerados pelo projeto liberal a levar os próprios liberais a apresentarem medidas avulsas para a resolução dos problemas que iam surgindo – na sua consideração pragmática de que algo tinha de ser feito e não sendo possível “estabelecer ligações entre os interesses diretos que [as] motivaram ou coerência ideológica entre as suas diversas ações.” (Polanyi, 2012: 323)

Ora, o nosso propósito é o de reapreciar este segundo movimento a partir de uma reflexão democrática, isto é, considerando as medidas tomadas pelos liberais, *contra o ideal liberal*, como sendo condicionadas por uma cultura democrática. Queremos com isto dizer que, para lá dos ideais liberais que dão forma ao posicionamento político dos decisores ou governantes, a percepção de que vivemos num espaço democrático e de que não é possível aplicar um projeto ideológico a qualquer preço condicionou e determinou o recuo ou a adoção, dependendo dos casos, de medidas concretas. Desenvolveremos este aspeto de seguida, a partir de uma consideração das reflexões polanyianas para o contexto presente.

3. Pensar o presente a partir de Polanyi

Como enunciamos a título introdutório, o nosso intuito passa por utilizar estes contributos de Karl Polanyi para refletir sobre a crise desencadeada nos

anos de 2007-8, em especial com a transmutação da crise financeira em crise das dívidas soberanas. Dessa transmutação resultou, em Portugal, um plano de políticas austeritárias justificadas⁷ pelo pedido de empréstimo realizado pelo governo português às três instituições que ficariam conhecidas como *troika*: Comissão Europeia, Banco Central Europeu e Fundo Monetário Internacional. Os quatro anos vividos pelo país, entre o período que antecedeu a chamada da *troika* e os três sujeitos à sua intervenção, parece-nos constituir um interessante laboratório de ideias para as reflexões de Polanyi.

Contudo, não pretendemos afirmar que vivemos hoje uma repetição daquilo que aconteceu no contexto de Polanyi, isto é, no final da década de 20 e anos subsequentes. As circunstâncias históricas e as mudanças sociais que se sucedem, incorporando os acontecimentos passados, excluem a possibilidade de que a história se repita. Mas não excluem a possibilidade de que certos acontecimentos provoquem reminiscências fortes em relação a situações passadas e pensamos que é precisamente isso que acontece com a recente crise financeira e bancária.

Que fatores despertam essas reminiscências? Esses fatores prendem-se essencialmente com o facto de podermos identificar hoje um consenso político que designaremos por *neoliberal*. O uso deste qualificativo é problemático: a designação é geralmente recusada não só por aqueles que não se reveem nos seus ideais, como também por aqueles que, aceitando a qualificação, entendem que a realidade política atual está muito longe daquilo que seria um projeto neoliberal. Queremos fazer, no entanto, uso deste qualificativo e defender que ele constitui, mesmo contra aquelas recusas, o consenso hegemónico atual. Mas, para tal, importa clarificar o seu uso.

Ao usarmos o termo *neoliberalismo* estamos a remeter para a sua origem histórica, recordando que foi cunhado por Alexander Rüstow em 1938, no Colóquio Walter Lippmann, cuja motivação passava por reunir os intelectuais que se opunham à progressiva adoção de políticas coletivistas na Europa e nos Estados Unidos. Nesse colóquio participaram figuras como Raymond Aron, Friedrich Hayek, o próprio Walter Lippmann, Ludwig von Mises e Rüstow, que

⁷ Embora as medidas políticas tenham sido justificadas pela imposição da *troika*, muitas foram aprovadas aproveitando aquela intervenção para escapar ao escrutínio público e democrático. O recente estudo realizado por Catherine Moury e Adam Standing, publicado no *European Journal of Political Research* no início deste ano, segue exatamente neste sentido. (cf. <https://www.publico.pt/2017/06/24/politica/noticia/a-troika-tambem-fomos-nos-1776711> - último acesso: 16/07/2017)

procuraram desenhar um projeto político que recuperasse os ideais do liberalismo clássico. No final da reunião, o termo proposto por Rüstow foi o escolhido para designar o movimento, embora se pudessem identificar, dentro dele, correntes distintas. Após a segunda guerra mundial, o grupo volta a reunir, desta vez em Mont Pèlerin, Suíça, tendo sido criado, por referência local, a Mont Pèlerin Society. Os seus membros comungam dos ideais de mercado livre, máxima liberdade para a iniciativa privada e forte limitação da intervenção estatal. Dois autores assumem destaque: Hayek, com especial influência junto dos países anglo-saxónicos, e Milton Friedman, que será particularmente influente nas políticas adotadas fora do mundo ocidental, sobretudo em virtude da sucessiva formação de jovens estudantes na Universidade de Chicago, que foram cumprindo a missão de levar as suas ideias para os países de origem. E como dá conta David Harvey, em *A Brief History of Neoliberalism*, as crises que marcaram o keynesianismo durante a década de 1970 levaram a que tivesse acontecido uma *neoliberal turn*.⁸

É a estes autores e a este conjunto de ideias que nos referimos quando designamos o atual consenso político como *neoliberal* ou como *Consenso de Washington*. O seu núcleo de ideias gira em torno de um processo de mercantilização crescente de todas as dimensões da vida (em especial daquelas esferas que tinham sido retiradas do mercado durante o consenso keynesiano, como a saúde, a educação, a energia, etc.). E dará forma aos princípios da política de austeridade, que passou a ser defendida na Europa como a reação adequada à grave crise económico-financeira: liberalização dos movimentos de capital, privatização alargada dos setores produtivos estratégicos, intensificação da exploração dos recursos ambientais e ecológicos e enfraquecimento do papel do Estado na economia em diversas áreas, particularmente no domínio laboral. A partir destes princípios do cânone neoliberal, o *austeritarismo* assenta na ideia de que apenas com uma forte contração das despesas do Estado (com congelamento de salários, pensões, diminuição das despesas sociais, etc.) e a diminuição do seu peso na economia, se torna possível resolver o problema da crise das dívidas soberanas.

⁸ Harvey, David (2005). *A Brief History of Neoliberalism*, Oxford, Oxford University Press.

Clarificado este aspeto, como podemos analisar o projeto neoliberal e a política de austeridade à luz de *A Grande Transformação*?

Em primeiro lugar, importa notar que um dos aspetos mais evidentes que o neoliberalismo herdou do liberalismo clássico e, naturalmente, do projeto liberal filosófico, prende-se com a *pretensão de naturalização*. A narrativa do liberalismo económico, como vimos, assenta na ideia de naturalidade dos mercados, *condição natural* que o Estado perturbaria com a sua atividade – e esta é a justificação para a eliminação da interferência estatal e a constituição de mercados autorregulados. O neoliberalismo do século XX recupera precisamente este espírito de naturalização e a tentativa de encontrar as leis rigorosas, e não hipóteses debatíveis, que devem regular a economia. Apresenta-se, para o efeito, como imbuído de imparcialidade científica e orientado para uma abordagem económica à imagem das ciências naturais. Afinal, o sistema económico seria igualmente um sistema natural. Nas palavras de Naomi Klein:

O núcleo de tais sagrados ensinamentos da Escola de Chicago era que as forças económicas da oferta, procura, inflação e desemprego eram como forças da natureza, fixas e imutáveis. (...) Tal como os ecossistemas se autorregulam, mantendo-se em equilíbrio, o mercado, deixado em paz, iria criar o número exato de produtos aos preços precisamente corretos, produzidos por trabalhadores que receberiam os salários à medida para poderem comprar esses produtos – um Éden de emprego abundante, criatividade sem limites e inflação zero. (Klein, 2009: 64-5)

Se o Estado intervém cria uma perturbação e está “a fazer mal indiscriminado ao equilíbrio do mercado e à capacidade dos seus vários sinais comunicarem uns com os outros. A missão da Escola de Chicago era, assim, a de purificar – despir os mercados dessas interrupções para que o mercado livre pudesse singrar.” (Klein, 2009: 67)

Esta estratégia, constituindo uma tentativa de se colocar no *campo neutro da realidade*, reivindica conhecer as regras objetivas que fazem a economia funcionar. É neste contexto que surge a famosa expressão de Margaret Thatcher de acordo com a qual não existe alternativa: não existe alternativa ao capitalismo liberal, não existe alternativa a uma economia baseada em mercados livres – eis como a realidade funciona, por isso não há outro caminho.

Contudo, e como vimos, Polanyi expõe o *mito da naturalidade dos mercados*, revelando, por um lado, como a lógica de mercado teve um domínio de aplicação extremamente reduzido ao longo da história da humanidade e, por

outro, como o sistema de mercado oitocentista foi uma criação do próprio Estado, i.e., resultado de políticas específicas e opcionais: longe de se tratar de um estado natural, o mercado seria antes resultado de uma construção social e política concreta. E este é o primeiro grande contributo de Polanyi, fornecendo-nos a possibilidade de desmontar a reivindicação neoliberal de ter encontrado as leis objetivas e científicas que permitem conhecer o modo de funcionamento da realidade e, por isso, as medidas económicas corretas a adotar. Demonstrando a artificialidade da lógica de mercado, Polanyi faz valer a ideia de que as condições políticas e económicas resultam da vontade de governação política e não de terem sido encontradas as regras objetivas de regulação da realidade e que se impõem inexoravelmente. Este aspeto é particularmente relevante quando avançamos para a justificação imprimida, em termos discursivos, às políticas de austeridade: estas seriam *a única solução possível* para a chamada crise das dívidas soberanas que resultou da crise bancária e financeira. Desconstruindo essa estratégia discursiva, retirando o tapete à pretensão de naturalidade e, por isso, de acesso privilegiado ao modo como as coisas (a economia) realmente funcionam, podemos recuperar então a discussão para o plano político e debate democrático. E nesse sentido, as implicações de Polanyi são profundamente democráticas.

Em segundo lugar, importa considerar o período de austeridade em Portugal, que se iniciou ainda antes da intervenção da *troika*.⁹ É possível encontrar, durante esse período, a dinâmica do duplo movimento descrita por Polanyi: após uma aplicação inicial do pacote austeritário, os momentos seguintes foram marcados por medidas governativas de recuo e correção, por um lado, e por forte contestação social, por outro. Na verdade, as medidas de contração económica geraram desemprego, aumento de pobreza, elevados níveis de emigração e uma forte queda da economia nacional. O flagelo social originou a necessidade de adoção de pacotes sociais de apoio, como forma de correção e proteção da sociedade.¹⁰ Tal como indica Polanyi, o prosseguimento de um plano liberal gera ele mesmo, e de forma quase espontânea, medidas de defesa e

⁹ Notemos que a primeira grande manifestação de rua em Portugal marcada pela contestação às consequências políticas e económicas da crise económico-financeira foi o protesto da Geração à Rasca, no dia 12 de março de 2011. Representou o maior sintoma de um tempo de crise: menos de um mês depois, o governo de José Sócrates cairia com o pedido de intervenção financeira à *troika*. A dinamização que o protesto da Geração à Rasca conseguiu em Portugal galvanizou os protestantes espanhóis já mobilizados: a 15 de maio seguinte, o grande protesto em Madrid dará origem aos *acampados* e *indignados* espanhóis, semente do movimento *Occupy*.

¹⁰ O maior partido político membro da coligação que constituiu o XIX Governo Constitucional, eleito com base num programa político de cariz austeritário, reconhece declaradamente a necessidade de adoção dessas medidas de proteção, conforme se pode encontrar aqui: <http://www.psd.pt/noticia/1743>.

proteção. E aqui reside, como vimos, a sua dimensão *u-tópica*, aspeto que se torna ainda mais evidente com a leitura da carta de demissão do Ministro das Finanças Vítor Gaspar.¹¹ Para além de reconhecer a necessidade de apoio popular para as medidas a adotar, algo que estava posto em causa em resultado das constantes manifestações, Gaspar reconhece *a necessidade de intervenção estatal* em diferentes domínios por forma a resolver os problemas criados pelas suas próprias políticas.

Os grandes custos de ajustamento são, em larga medida, incontornáveis, dada a profundidade e persistência dos desequilíbrios, estruturais e institucionais, que determinaram a crise orçamental e financeira. No entanto, o nível de desemprego e de desemprego jovem são muito graves. Requerem uma resposta efetiva e urgente a nível europeu e nacional. Pela nossa parte exigem a rápida transição para uma nova fase do ajustamento: a fase do investimento!¹²

Mais uma vez: um projeto que procura implementar uma sociedade de mercado não é concretizável, começa a falhar nos seus próprios termos.

Como referimos, era essencial para a posição de Polanyi destacar a espontaneidade da reação de proteção, por forma a afastar a narrativa que apresentava a conspiração coletivista como responsável pela destruição das possibilidades de êxito do projeto liberal.¹³ O nosso propósito passa, no entanto, por oferecer uma leitura do segundo movimento como revelando a cultura democrática da sociedade europeia. De acordo com essa leitura, as reações contrárias aos princípios liberais podem ser perspetivadas como consequências da pressão social e institucional contra as medidas que implicavam custos percecionados como inaceitáveis. Em Portugal, na medida em que as medidas de austeridade tiveram implicações extremamente amplas, a contestação foi sentida um pouco por toda a sociedade, revelando a cultura democrática em que assentam as nossas instituições. Tal aconteceu com as decisões do Tribunal Constitucional a um nível mais institucional como garante do Estado de direito democrático, mas também com um clima de contestação social constante, marcado por grandes manifestações, greves, *grandoladas* e outras estratégias afins.

¹¹ Vítor Gaspar foi Ministro das Finanças do XIX Governo Constitucional até 1 de julho de 2013. Após cessar as suas funções governativas, Atualmente ocupa o cargo de Diretor do Departamento de Assuntos Orçamentais do Fundo Monetário Internacional.

¹² Cf: <https://www.publico.pt/destaque/jornal/carta-de-vitor-gaspar-a-passos-coelho-na-integra-26769219> (último acesso: 16/07/2017).

¹³ O facto de o mesmo tipo de reivindicação se ter feito ouvir um século depois torna o contributo de Polanyi ainda mais precioso.

É deste modo que os contributos de Polanyi permitem reflexões democráticas: permitem-nos falar de uma espécie de *des-cobrimto* da estrutura democrática da sociedade europeia. Interpretado nas suas duas dimensões, enquanto resultado pragmático dos próprios defensores do liberalismo e enquanto resultado das pressões sociais e institucionais, encontramos o reconhecimento de que um projeto político não pode ser feito a qualquer preço, à custa de um sofrimento individual e social inaceitável.¹⁴ Afinal, como chama a atenção Alain de Benoist no seu pequeno texto sobre democracia – *Démocratie: le problème* –, “em contraste com o Oriente o despotismo absoluto foi sempre excessivamente raro na Europa.” (Benoist, 2011: 15) Todas as antigas monarquias eram eletivas e em geral foi apenas a partir do século XII que se tornaram hereditárias; mesmo aí, os reis governaram sempre com a constituição de parlamentos, assembleias ou, como em Portugal, de cortes. Nos momentos em que a legitimidade dos governantes foi posta em causa, quebrando-se o elo de confiança com o povo, a contestação popular conduziu a movimentos revoltosos e à reorganização dessa legitimidade, como aconteceu com a redação da Magna Carta em Inglaterra no século XIII. Esses momentos de contestação recordaram sempre aos governantes que eles respondiam perante as suas populações. O amadurecimento democrático da sociedade europeia e a institucionalização de uma democracia liberal e representativa limitou-se a tornar mais evidente essa estrutura. E de acordo com ela, nem o liberalismo nem o neoliberalismo, com as suas pretensões de naturalidade, escapam ao espírito democrático que podemos reconhecer no princípio do duplo movimento de Polanyi.

4. Pensar o futuro a partir de Polanyi

Como procuramos destacar, as considerações de Karl Polanyi constituem importantes contributos para o domínio democrático. Por um lado, permitem desmontar a estratégia discursiva de pretensão de naturalidade e acesso privilegiado ao modo como a realidade funciona, que justificaria em última instância uma solução inevitável. Alarga-se, desse modo, o espaço de discussão

¹⁴ Outra reminiscência provocada pelos tempos presentes prende-se com o surgimento dos movimentos populistas, sobretudo de cariz nacionalista, que têm pontuado o panorama ocidental. Também aqui, na sua ligação à recente crise neoliberal, os contributos de Polanyi são relevantes.

democrática e repolitiza-se o domínio das decisões coletivas e a própria economia (de acordo com a ideia de incrustação que Polanyi celebrizou). Por outro lado, permitem o aprofundamento da reflexão democrática na medida em que as suas apreciações sobre a dimensão utópica e o princípio do duplo movimento revelam os diferentes modos a partir dos quais podemos avaliar o funcionamento democrático de uma sociedade. Em Portugal, e embora a ocupação da rua como espaço político não tenha conduzido à exigência recorrente de demissão do governo, ela condicionou os decisores políticos, fazendo-os muitas vezes recuar nas suas medidas.¹⁵ Se o período austeritário foi, em certo sentido, um movimento de diminuição do espaço democrático (sobretudo a um nível discursivo, como fizemos notar), o contramovimento agiu em sentido contrário, isto é, abrindo espaço para consequências democráticas. E as reflexões de Polanyi permitem esta complexa consideração dos tipos de discurso que coexistem sempre no espaço público.

Gostaríamos, no entanto, de terminar com os contributos polanyianos para o futuro, destacando esse aspeto a partir de dois elementos: por um lado, a introdução do fator de contingência e, por outro, a reabilitação dos estudos humanísticos. Uma das lições mais ricas que um trabalho como o de Polanyi nos proporciona vai no sentido de deslocar a nossa atenção para a dimensão contingente dos projetos políticos: não há algo como uma condição natural a que possamos ou devemos aspirar ou algo como um projeto político que deva ser imposto pela reivindicação de resultar de um conhecimento especial da Verdade.¹⁶ Esta noção de contingência é introduzida por Polanyi com as observações de como as diferentes sociedades foram encontrando, historicamente, os seus modos particulares de organização, a partir da adaptação ao meio ambiente e da consideração dos seus projetos coletivos. Ora, esta lição é conseguida por uma metodologia que usa trabalhos realizados por diferentes áreas do saber, como a história, a antropologia, o direito ou a etnografia.¹⁷ E é na

¹⁵ O melhor exemplo deste facto é o recuo na medida apresentada pelo governo, em setembro de 2012, em relação à Taxa Social Única, que provocou uma das maiores manifestações desse período, no dia 15 de setembro.

¹⁶ Polanyi é particularmente previdente quando chama a nossa atenção para a ideia de que implementar “em todo o planeta um novo modo de vida que [aspira] à universalidade em termos sem precedentes desde os primeiros tempos de afirmação do cristianismo, sendo a diferença que, desta feita, o movimento se situava num plano puramente material,” gera terríveis consequências na medida em que constitui uma forma de etnocídio: “A desagregação das instituições tem lugar devido ao simples facto da imposição de uma economia de mercado a uma comunidade organizada em termos completamente diferentes.” (Polanyi, 2012: 337)

¹⁷ A análise de Karl Polanyi assume, neste sentido, um cariz distinto de outras obras de autores seus contemporâneos como, por exemplo, a de Karl Mannheim (*Ideologie und Utopie*, de 1929), que se apresenta como uma obra teórica no campo da sociologia do conhecimento.

combinação destes dois aspetos que reside, na nossa perspetiva, o grande contributo de Polanyi para o futuro: não estamos condenados a um projeto político específico, somos sempre livres de re-pensar e re-formular as nossas ideias políticas – e a melhor forma de aumentar a nossa capacidade para imaginar possibilidades alternativas¹⁸ é aproveitar os trabalhos realizados pelas diferentes áreas do saber humanístico.

Bibliografia

Bauman, Zygmunt. (2000). *Liquid Modernity*. Cambridge/Malden, UK: Polity Press.

Beckert, Jens. (2016). *Imagined Futures: fictional expectations and capitalist dynamics*. Harvard, Mass: Harvard University Press.

Benoist, Alain de. (2011). *The problem of democracy*. UK: Arktos Media.

Gaspar, Vítor .(2013). Carta de demissão. *Público*. Retirado de: <https://www.publico.pt/destaque/jornal/carta-de-vitor-gaspar-a-passos-coelho-na-integra-26769219>

Harvey, David. (2005). *A Brief History of Neoliberalism*. Oxford, UK: Oxford University Press.

Klein, Naomi. (2009). *A Doutrina do Choque. A Ascensão do Capitalismo de Desastre*. (Trad. port. Marco de São João. Lisboa). Portugal: Smartbook.

Mannheim, Karl. (1936). *Ideology and Utopia. An introduction to the Sociology of Knowledge*. London/Henley, UK: Routledge & Kegan Paul.

Moury, Catherine e Standring, Adam. (2017). ‘Going beyond the Troika’: Power and discourse in Portuguese austerity politics. *European Journal of Political Research*, 56(3), 660-679.

¹⁸ O papel da imaginação na prática social foi alvo de particular atenção por parte de Paul Ricoeur, especialmente em *L'Idéologie et L'Utopie* (1997), a partir da dinâmica entre ideologia e utopia que nos permite considerar possíveis mundos futuros.

Polanyi, Karl (2012). *A Grande Transformação. As origens políticas e económicas do nosso tempo* (Trad. port. Miguel Serras Pereira). Lisboa, Portugal: Edições 70.

Ricoeur, Paul (1986). *Ideologia e Utopia*. Trad. port. Teresa Louro Perez. Lisboa: Edições 70.

Rothbard, Murray. (1961). *Down With Primitivism: A Thorough Critique of Polanyi*. Retirado de: <https://mises.org/library/down-primitivism-thorough-critique-polanyi>

Soromenho-Marques, Viriato. (2014). *Portugal na queda da Europa*. Lisboa, Portugal: Temas e Debates – Círculo de Leitores, 2014.



POLANYI'S ECONOMIC EMBEDDEDNESS, COUNTERMOVEMENT, AND REPUBLICAN POLITICAL ECONOMY

A INCRUSTAÇÃO ECONÓMICA, O DUPLO MOVIMENTO E A ECONOMIA POLÍTICA REPUBLICANA EM POLANYI

Bru Lain*
bru.lain@ub.edu

Abstract. *The aim of this article is to contend that Karl Polanyi's work bears significant coincidences with the republican tradition of thought. The first of them is one of a methodological or epistemological kind, and it consists of the use of a very similar "social ontology". The second one is of a substantive sort, and it is related to the "material conditions for freedom" which derive from a very similar conception of freedom and property. In the third section, we propose a republican reading of Polanyi's work based on three of the author's primary notions, those of "economic embeddedness", "double movement", and "political economy". The article concludes by arguing that such concepts may prove useful both for understanding the democratic-republican program and for underlining its present validity in contemporary market societies.*

Keywords: *Polanyi, Republicanism, Property, Freedom, Material Independence, Political Economy.*

Sumário. *O objetivo deste artigo é defender que a obra de Karl Polanyi guarda importantes coincidências com o pensamento republicano. A primeira, de tipo metodológico ou epistemológico, é o uso de uma "ontologia social" muito similar. A segunda, de tipo substantivo, gira em torno da ideia das "condições materiais da liberdade" baseada numa concepção muito parecida da liberdade e da propriedade. A terceira seção sugere a leitura republicana da obra de Polanyi baseada em três de suas ideias fundamentais: a "incrustação económica", o "duplo movimento" e a "economia política". O artigo conclui defendendo que tais conceitos podem ser úteis tanto para compreender o programa republicano-democrático como para sublinhar sua vigência nas sociedades de mercado contemporâneas.*

Palavras-chave: *Polanyi, republicanismo, propriedade, liberdade, independência material, economia política.*

* Associated professor of Sociology at the University of Barcelona and of Economy at Elisava University.

o. Introduction: Karl Polanyi's Possible Republican Alignment

Hungarian historian, sociologist, economist and anthropologist Karl Polanyi (1886-1964) has been alternatively described as a “left-wing liberal”, a “democratic socialist”, a “functional socialist” or a “Christian humanist” (Baum, 1996; Block, 2003; Burawoy, 2003; Dale, 2016; Laín, 2014; Mendell, 2003). However, his possible links to the republican tradition of thought have never been duly acknowledged. Because of such lack of consensus among scholars, we feel compelled to defend the hypothesis that his work is significantly connected to the republican political-philosophical tradition and, in particular, to the more democratic version of it.

Since the first third of the twentieth century republicanism “has been the object of renewed interest, first and most decidedly among historians, and then, in a more restricted manner, among law students, philosophers and political scientists” (Goodin, 2003: 55). As it has been pointed out by several authors, “there is no doubt that a re-emergence of republicanism is under way” (Dagger, 2006: 151).¹ Beyond the academic realm, we should also take into account a broad group of classical and modern thinkers, from Aristotle and Cicero, to Machiavelli, Rousseau, Locke, Kant or Smith, among others. Additionally, we should also take heed of the contributions of such activists as Overton and Libourne, as well as Paine, Jefferson and Robespierre, to name only a few. There are significant differences among the authors commonly associated with this tradition in both the academic domain and the political one. However, in its more than 2,500 years of existence, republican thought has always exhibited a certain set of shared methodological and ontological principles that enable us to speak of the existence of a “social ontology” of a republican kind. Our question here is: to what degree does Polanyi’s work share this social ontology?

¹ Some of the contributors to the re-emergence of republicanism is the so-called Cambridge School, headed by J. G. A. Pocock (1975) and Quentin Skinner (1998, 2008), who were the target of devastating criticism by historian Ellen Meiksins Wood (2008). Philip Pettit (1997) was also a major contributor, while authors such as William H. Simon (1991); Stephen Holmes & Cass R. Sunstein (1999); Sunstein (2004), or Bruce Ackerman (1993) flared up the debate in legal theory and constitutionalism. More recently, the interest in republicanism has also reached the areas of public policies, institutional design, theories of justice or productive democracy, e.g., with the works of Stuart White (2011); Richard Dagger (2006), or Alex Gourevitch (2013). In Spanish, the following authors stand out: Antoni Domènech (2000, 2004a, 2005); Jordi Mundó (2005, 2015); María Julia Bertomeu (2004, 2005); Daniel Raventós (2007), and David Casassas (2005, 2011).

1. Methodological Coincidences: A Shared Social Ontology

There are several coincidences between Polanyi's work and the republican tradition, at least concerning their general worldview, i.e., their use of a very similar "social ontology" (Casassas, 2013). Both for Polanyi and for the republican tradition, social analysis starts from a set of epistemological and methodological assumptions that could be summarized as follows:

- a) Polanyi and the republican tradition coincide in their recognition of the plurimotivational quality of human cognition, and in the implications that this has for the political life and the economic one.² Polanyi contributed to the undermining of the "atomist" notion of cognition which marks the theory of rational choice and which sees human cognition as operating on a unimotivational pattern—the maximizing of choice. This critique is particularly present in *Our Obsolete Market Mentality* (1947), where he argues that "economic motives never formed with him the sole incentive to work", although it was "exhorted by economists and utilitarian moralists alike to discount in business all other motives than material ones." Actually, man "was still found to be acting on remarkably mixed motives, not excluding those of duty towards himself and others" (Polanyi, 1957: 115). This plurimotivational structure of human cognition has significant effects, not only on the conceptualization of the relations between individuals and markets, but also on the determination of how subjects relate to the cognitive and anthropological patterns under which social institutions —families, communities, markets, states— operate.³
- b) Both Polanyi and republicanism display a similar interest in the issue of the distribution of social output, understood as the result of more or less conflicting interactions between different social strata and classes. Certainly, the achievement of a fair or equitable distribution of social wealth was never

² To confirm his opinion in this respect, see Polanyi, *The Great Transformation* ([1944] 2001, I, XVIII; *GT* onwards), where he refers to "economic determinism" as a false axiom used to justify the belief in individuals' "motivational monism" and aimed at providing the necessary base for the defense of putatively "self-regulated markets" and the liberal vision of them.

³ In fact, Asad Zaman (2016: 47) states, "An individual's behaviour and motivations cannot be understood outside of his community and his social and historical milieu."

a republican goal. On the contrary, its concerns always revolved around the protection of freedom understood as the absence of relations of domination. For republicanism, a more or less equitable distribution of social output is rather a direct economic consequence of a broader or narrower extension of freedom. In other words, for republican thinkers, the fair distribution of social wealth has a strong instrumental component; it does not constitute an end in itself and, simultaneously, it depends on the ideological option which is adopted with respect to the degree of extension that is should to the political domain of freedom. As we shall see, this is the reason why republicanism can be historically divided into a more aristocratic trend and a more democratic one.

- c) Both Polanyi's work and the republican line of thought also share a common perspective in their understanding the problems of output distribution in societies, which they tackle from the standpoint of the role that historically contingent institutional structures and, particularly, class relations have played in them. As Polanyi pointed out, from an economic point of view, while "free competition leads to monopoly, freedom in property rights leads to the exploitation of the masses" (Polanyi, 1931: 1). So that, according to him and to republican thought, the distributive pattern of a particular society depends to large extent on its social class structure, on these classes' particular and conflictive interests, and on the different equilibrium of their bargain power.⁴
- d) Closely related to the previous point, they also coincide in the fact that they always bear in mind that the social life is actually traversed, and constituted, by multiple asymmetrical power relations, basically between the poor and the wealthy. For this reason, Polanyi argues that we must always start from the realization that "a human society from which power and coercion are absent

⁴ It is true, as Fred Block suggests, that "in the writing of the GT, there are few explicit references to this [class] perspective. Terms such as 'productive forces' and 'ruling classes' are completely absent" (Block, 2003: 280). However, "adopting a position vis-a-vis the middle class analogous to that of Gramsci's toward the working class", Gareth Dale suggests, "Polanyi made a case for its hegemonic role in a broad democratic bloc, to ensure leadership of the intellectual forces on a democratic foundation" (Dale, 2016: 51-2). In doing so, he reframed the classical republican goal, that is: the dissolution of all social classes –and their divergent interests– in a community of free and equal citizens, since according to Aristotle, citizens "in most cases of republican government (...) tend to be on an equal level in their nature and to have no difference at all" (*Pol. I*, 1259b).

is not possible” (Polanyi, 1947: 116). Thus, both the bulk of the republican thought and Polanyi’s own ideas are marked by a constant concern for the development and the adaptation of institutional arrangements whose design is properly grounded in such reality.⁵

These four analytical-methodological guidelines, which are characteristic of republicanism, are also very present in the writings of the Hungarian thinker.⁶ However, we shall not linger on them here, since our purpose was nothing more than briefly pointing out their similar methodologies and the use of a very similar “social ontology”. Now, our real interest in this writing is in the coincidences of a more substantive kind between Polanyi’s work and the republican tradition.

2. Substantive Coincidences: The Material Conditions for Freedom

The first substantive coincidence that can be noted is the use of a similar definition of freedom as a starting point. According to the classical republican notion, freedom emerges when no arbitrary interferences, that is, when no domination relationships, can be observed between private individuals, or between them and public institutions.⁷ Eric MacGilvray (2011: 28) describes it as follows:

One is free if one is able to act under one’s own initiative instead of merely reacting to the deeds of others, benevolent though they might be (...) To depend on another person, whether tyrant or patron, master or benefactor, is to be unfree (...) the free man must be economically independent, so that he does not rely on others for his (...) the secure ownership of property (...) is essential to the enjoyment of freedom.

Hence, freedom will be present as long as there are no relationships of dependence or domination between individuals, so that these are able to “act under one’s own initiative.” For this to happen, an individual must be materially independent—he must own property—, as absence of such property will make him inevitably dependent on the arbitrary decisions of others, “benevolent

⁵ As he defended, this kind of conflict “seemed inherent in economy, whether as competition of individuals or as struggle of classes (...) in present, or perhaps future, society” (GT: 89).

⁶ For a more detailed exposition of these four points, see Domènech & Bertomeu (2005).

⁷ As Philip Pettit (1997: 52) points out, “one agent dominates another if and only if they have a certain power over that other, in particular a power of interference on an arbitrary basis.”

though they might be.”⁸ We should also take notice that, so that we can speak of freedom in society, two conditions must be present: “no citizen should be rich enough to be able to buy another”, and no other citizen should be “so poor that he has to sell himself” (Rousseau, 1999, II, XI). This is why, “if a republican state is committed to advancing the cause of freedom as non-domination among its citizens, then it must embrace a policy of promoting socioeconomic independence” (Pettit, 1997: 158-9).

To fulfil both conditions, a Republic has to acquire an institutional order, which allows for the weaving of a social fabric free of domination relationships. Thus, it is a normative requirement of freedom the establishment of a political-economic order designed to function as a safeguard that mitigates the sort of power asymmetries leading to both dependency and relationships of domination. To do so, property is a necessary condition; those who own some — and enjoy independence — will not be forced to subject themselves to the will of third parties. This is why Arthur Lee held, in 1775, that property is “the guardian of every other right,⁹” and Carol Rose (1996) asks herself whether property is “the keystone right.” Several writers refer to this thesis — the idea of property as freedom — as the “material conditions for freedom” (Raventós, 2007; Casassas & Raventós, 2007, Bertomeu, 2005), a notion deeply rooted in the classical republican legal order. The Roman legal category that identifies such condition of the material independent citizen is that of *sui iuris* (or material, and thus legally, autonomous person endowed with his own rights), in contrast with the *alieni iuris* (or person lacking all material, and hence, legal autonomy). It is not surprising then that, for Aristotle, and also for Marx, a *misthotés* or wage

⁸ Property should be understood here “as durable control over a set of material resources or assets that leverages one’s freedom in economic exchanges by effectively securing a form of bargaining power vis-à-vis other agents. In other words, property allows one the freedom to choose who to contract with and under what conditions. In extreme cases, it offers an exit option from those exchanges that are deemed alienating, exploitative or otherwise detrimental to one’s freedom, ideally by opening the door to more congenial ones” (Casassas & De Wispelaere, 2016: 287). In sum, “property” is used here as synonym of “material independence”, since this property gives individuals “a significant degree of material independence” (White, 2011: 567). Nevertheless, there might be plenty of modes of securing such material independence, even within the republican framework. One of them is by promoting the so-called Universal Basic Income (Casassas, 2007; Raventós, 2007), others might be the schemes of Property-Owning Democracy famously stated by James Meade in 1964 (Meade, 1964; Rawls 2001; O’Neill & Williamson, 2012).

⁹ Arthur Lee, quoted in Ely (1992: 26).

labourer, being materially dependent on another person—his employer, who is the owner of the means of production — is little more than a “part-time slave.”¹⁰

We have just pointed out Polanyi’s agreement with the thesis of the existence of “material conditions for freedom”. However, in order to demonstrate such coincidence, now we must move on to explore in more depth the Polanyian links to republican idea of freedom and its relation to property.

2.1 Freedom as Non-Domination, and Institutional Designs

It is commonly understood that “the freedom of one person stops where that of another person begins.” The logical, formal consequence of this would be that “the more limited the freedom of others is, the broader mine is”. The result — one definitely lacking any substantive meaning — would be that “one would be the only absolutely free person if all the rest lost such privilege completely”. This article takes a significantly opposed stance on the issue, and adopts the republican notion of freedom, where this is not understood as absence of interferences, but as absence of arbitrary interferences, i.e., the notion of freedom as non-domination.

Republican freedom tends to be analysed by contrasting it with liberal one.¹¹ This is often done by distinguishing between a negative, putatively republican freedom (the right not to be interfered with) and a positive, supposedly liberal one (the right *of* or *to*), developed by Isaiah Berlin (1958) on reinterpreting the binary distinction between the “liberty of the ancients” and that of the “moderns”, brandished by Benjamin Constant in 1819. Nevertheless, such categorization

¹⁰ Aristotle defended the existence of this connection between material dispossession and servitude, according to which the wage labour that was needed due to lack of—property—alternatives was nothing but a “sort of limited slavery” (*Pol.*, 1260a). In 1656, almost two thousand years later, in his famous *A Commonwealth of Oceana*, James Harrington repeats that: “The man that cannot live upon his own must be a servant; but that can live upon his own may be a freeman” (Harrington, 1992: 269). Two hundred years on, in 1875, a very Aristotelian Karl Marx will insist on this same topic in his *Critique of the Gotha Program*: “The man who possesses no other property than his labor power must, in all conditions of society and culture, be the slave of other men who have made themselves the owners of the material conditions of labor. *He can only work with their permission, hence live only with their permission*” (Marx, 1994: 316, emphasis added).

¹¹ In contrast with the liberal idea of a purely negative freedom, republican freedom is a “dispositional notion” (Domènech, 2000: 30).

ended up sparking off an academic debate that was utterly disconnected both from the reality of the ancients' and from that of the moderns'.¹² More recently, the republican type of freedom has been interpreted as a third kind of freedom or a somehow more refined kind of negative liberty: "Being unfree does not consist in being restrained; on the contrary," says Pettit (1997: 5)

the restraint of a fair system of law—a non-arbitrary regime—does not make you unfree. Being unfree consists rather in being subject to arbitrary sway: being subject to the potentially capricious will or the potentially idiosyncratic judgment of another. Freedom involves emancipation from any such subordination, liberation from any such dependency. It requires the capacity to stand eye to eye with your fellow citizens, in a shared awareness that none of you has a power of arbitrary interference over another.

Pettit chooses a definition which is halfway between purely negative freedom and purely positive one in the vein of Berlin. Thus, republican freedom would be a person's (A's) capacity not to be arbitrarily interfered with by one or several other persons.¹³ Consequently, interference of a non-arbitrary kind in A's set of opportunities would be admissible. Antoni Domènech (2000: 30) insisted on this when distinguishing between, on the one hand, domination relationships or arbitrary interferences and, on the other, non-arbitrary ones:

I am free when I am under no one's thumb or authority, when no one—whether he/she actually does it or not—could interfere with my life plans at his or her discretion (...) my liberty is promoted when other free people can interfere with my life in a non-arbitrary way.

Both definitions may prove prolific when it comes to re-thinking the political-normative implications which derive from them. The key point here, though, is to establish which of A's spheres is (or can be) interfered with. In the republican tradition, such sphere is always institutionally marked—notice the methodological assumptions *b* and *c* in the republican "social ontology" at the beginning of this writing—and is inextricably linked to the material conditions of both A and his potential dominators (not only B, but also C, D...). Let us see it this way: an interference of B in the set of A's opportunities which did not affect the conditions for A's autonomous existence (e.g., a white lie) might be ethically or

¹² A critique of this binary conception of freedom can be found in Domènech & Bertomeu (2005), and Holmes & Sunstein (1999).

¹³ Pettit defines the republican freedom with his metaphor of the "eyeball test", which describes that particular social position from which one is able "to look one another in the eyes without reason for fear or deference" (Pettit, 2014: 82).

morally reprehensible but, for republicanism, it would still constitute a political irrelevant (though arbitrary) interference. It would not be the same if B (or C, or D...) is able to make use of A at his own will (as a wage labourer, for instance), as A finds himself institutionally forced to accept this dependent relationship because he lacks the resources or the material conditions that are needed to subsist in an autonomous and independent manner. What is rather politically relevant for republicanism is the fact that which sphere A (and B, and C...) can be interfered with (his array of available opportunities) is something which is always politically and economically regulated.

This is why, once we acknowledge that property (or material independence) is the basis for republican freedom, Berlin's distinction between negative liberty (the freedom *from*) and positive liberty (the freedom *to*) loses much of its meaning. Domènech & Bertomeu (2005: 70) put it this way:

On one hand, it is their freedom *to* ("positive") govern themselves by administering the material bases for his autonomous existence what trains people in virtue, what ultimately enables them to be free citizens. On the other hand, the state is compelled to intervene "positively" (and, sometimes, at great expense) in the set of opportunities of the myriad of individuals who might try to destroy X's freedom *from* being interfered with ("negative") in the self-governance of his property ("positive"), as well as it is to "assist" ("positively") X in his freedom *to* ("positive") resist the assault in a licit manner.

We shall look into this further below. What we should like to stress now is that (a) Polanyi was always deeply interested in the status of freedom in the industrial society, and (b) his definition of freedom is quite similar to the one we have exposed.¹⁴ However, reconstructing his notion of freedom may prove a rather intricate task, since it involves the tracking of countless fragments, chapters, letters, and passages from his journalist articles, opinion notes and books.¹⁵ However, what is not as difficult to come across is the antagonist of this kind of freedom. For Polanyi, in line with the republican tradition, the opposite of liberty is slavery, i.e., the condition of "being forced or compelled by someone

¹⁴ Although Polanyi's theory might be linked with Socialism tradition in the light of several of his assumptions, according to Michael Burawoy (2003: 229), one important divergence with this theory might be that he emphasized "a tendency inherent in industrial civilization rather than in capitalism" in itself in talking about social class domination and, thus, the status of freedom in society.

¹⁵This intricate reconstruction of Polanyi's freedom can be found in: Polanyi (1931; GT: 389-408; 1947; 2012: 317-326; and 1959).

else's will.¹⁶ In this sense, Polanyi was a harsh critic of the liberal conception of freedom, since "with the liberal the idea of freedom thus degenerates into a mere advocacy of free enterprise" (GT: 265).¹⁷

Far from the logic of nineteenth century liberalism, for Polanyi and for republicanism, governments can (and must) interfere in a non-arbitrary manner in the set of opportunities of individuals in order to ensure the material conditions in which freedom can be thought of.¹⁸ For this, it is essential for the institutional order (basically, the political arrangement of markets and the economy) to recover its "substantive" character and be designed as a means to sustain a certain distribution of wealth, and therefore, of freedom in itself. Consequently, large private powers (brought about by large accumulations of wealth) that are capable of imposing domination relations should not be allowed to exist. So that, "if the property system or distribution has the contingent effect of allowing domination, then that makes a case for institutional adjustment" (Pettit, 2006: 139). Thus, it is possible to understand how Polanyi's concerns always revolved around the material "conditions for freedom." What is really at stake, he declares, is that "the very possibility of freedom is in doubt. It appears that the means of maintaining freedom are themselves adulterating and destroying it;" hence, "the key to the problem of freedom in our age must be sought on this latter plane" (GT: 262).

¹⁶ "Jean Jacques Rousseau or ¿it is possible a free society?", in: Polanyi (2012: 321). Polanyi's domination relationships (or arbitrary interferences, as republicanism would term them) can be summarized in the (false) "liberty to exploit one's equals, the liberty to obtain exorbitant earnings without providing the community with comparable services, the liberty to prevent technological innovation from being used for public good, or the liberty of benefiting from public calamities by secretly manipulating them for private profit. If such liberty ever disappears, it will be for the good of all" (Polanyi, 1947: 116).

¹⁷ Similarly, Erik MacGilvray observes that "From a republican standpoint, the freedom to sell one's labor to the highest bidder is not properly speaking freedom at all, as the various critics of wage slavery never tired of pointing out. From the standpoint of market freedom, by contrast, the effort to preserve or enhance the security of the working class by blocking or regulating voluntary transactions between consenting adults is a manifest and even paradigmatic encroachment on individual liberty" (MacGilvray, 2011: 165).

¹⁸ There is a large amount of work arguing about the more or the less radical-interventionist character of Polanyi's framework. Someone interested in should takes into account: Block & Somers (2014), Dale (2010, 2016), De Castro & Pereño (2012), Fraser (2012), Halperin (2004), Hodgson (2016), and Mendell (2003).

Emphasizing the need to establish a democratic institutional order that fosters such material conditions, Polanyi (GT: 236) adds that personal freedoms will exist as long as they

deliberately create new safeguards for its maintenance and, indeed, extension. In an established society the right to nonconformity must be institutionally protected. The individual must be free to follow his conscience without fear of the powers.

Freedom, therefore, must be discretionarily protected. Moreover, its protection must achieve its universalization. In this way, the achievement of freedom becomes inextricably linked to the regulation of freedom by producing “spheres of arbitrary freedom protected by unbreakable rules” (GT: 254). Such regulation, nonetheless, must not be understood as some sort of arbitrariness or despotism, neither a loss of freedom, on the part of the political body; it is rather the only way to ensure the material conditions which are needed to universalize freedom.¹⁹ This is the reason why Polanyi (GT: 262-3) insists that, in the institutional level,

regulation both extends and restricts freedom (...) The comfortable classes enjoy the freedom provided by leisure in security (...) They talk of slavery, while in effect only an extension to the others of the vested freedom they themselves enjoy is intended.

The republican-Polanyian type of freedom demands a “substantive” concept of freedom, or freedom understood as a *de iure* status which is simultaneously sustained by the material conditions that make it ultimately possible as a *de facto* reality. Thus, Polanyi’s vision is not based on a merely isonomic conception of freedom (or freedom as formal equality in law). Since “no mere declaration of rights can suffice,” “institutions are (also) required to make the rights effective.” Actually, his contention is that (GT: 264-5):

regulation and control can achieve freedom not only for the few, but for all. Freedom not as an appurtenance of privilege, tainted at the source, but as a prescriptive right extending far beyond the narrow confines of the political sphere into the intimate organization of society itself.

¹⁹ In order to develop this point—something which is beyond the scope of this article—, we should take into account the following explanation: “Thus, implementing the republican ideals does not only demand coping with the *imperium*—those relations of domination which are derived from the state—but also, and very particularly, with the *dominium*, i.e., those relations of domination created by dependency bonds rooted in the core of the so-called *civil* world and which largely give rise to the *imperium* as well” (Casassas, 2005: 238).

In summary, the normative requirement of the republican-Polanyian kind of freedom demands certain political and legal instruments capable of sustaining it (the required legal-institutional regulation of freedom) but, above all and most importantly, it demands the kind of economic order and material basis that can make it possible. “Hence,” Polanyi goes on, the “transcending importance of the institutional aspect of the economy” (Polanyi, 1957: 249). It is not striking that he identified the way in which the market is instituted in society as the prevailing factor in the shaping and ensuring of freedom, since “liberty, in a complex society, demands an inviolable passport. It is necessary to protect the individual against undue pressure, be it by a person or by a company” (Polanyi, 2012: 341). This is how he identifies the sudden emergence of economic liberalism and market society in the nineteenth century as the main historical process by which freedom came to be restricted to the enjoyment of the wealthy classes.

One of Polanyi’s major contributions to the study of this period’s economy was, no doubt, his description of the processes of communal dispossession or “enclosures of open fields and conversions of arable land to pasture during the earlier Tudor period in England, when fields and commons were hedged by the lords, and whole counties were threatened by depopulation” (GT: 36). His analysis of such processes (termed “primitive accumulation” in the Marx’s *Capital*,²⁰ and “accumulation by dispossession” by David Harvey²¹) gained him an insight into the true character of the new mode of production: the fact that the expansion of capitalist-type markets involved an accelerated and virulent destruction of classical freedom based on material independence grounded on the communal and small private property. Such expansion implied a previously unseen increase in the forms of social domination, which were no longer mediated by the feudal bond between a liege lord and a small free tenant, but rather by the formally free contractual relationship between employers and the new industrial proletariat.

²⁰ In the chapter *Historical Tendency of Capitalist Accumulation*, Marx himself estimates that around 3,511,770 acres of land were taken away or expropriated by large landowners in England between 1810 and 1831. This is what he called the “dissolution of private property based on the labour of its owner” (Marx, 1906, VIII, XXXII).

²¹ Harvey (2003: 158).

With this process of dispossession, Polanyi says, the old basic pillars of the whole economy turn to commodities: nature becomes land, fiduciary money becomes an object for speculation, and human beings become labour power. In other words, “the extension of the market mechanism to the elements of industry—labour, land, and money—was the inevitable consequence of the introduction of the factory system in a commercial society.” The emergence of such “fictitious commodities²²” will cause the collapse of all communities, whose essential institutions will become subsumed under the new logic of commoditization. The most dramatic thing about this is that “the fiction of their being so produced became the organizing principle of society” (GT: 78; 79).

All in all, the process of commoditization of the old forms of popular economy based on common resources represented an unprecedented and traumatic transformation based on the gradual abandonment of all economy in its substantive sense.²³ When we take this into account, we discover that Polanyi’s analysis of such process constitutes one of his major points of connection with the republican tradition, for which freedom is established by and through a basic mainstay: the fact that people cannot be free if their material existence is not politically ensured. And particularly, what links Polanyi with the republican-democratic tradition is that this must be made extensive to the population as a whole, and not limited to a few. This is the bond between freedom and property in the republican-Polanyian line of thought, which we shall explore in more depth in the following section.

²² In *The Economy As Instituted Process* (Polanyi, Arensberg, & Pearson, 1957: 255), Polanyi himself states that “the rise of the market to a ruling force in the economy can be traced by noting the extent to which land and food were mobilized through exchange, and labor was turned into a commodity free to be purchased in the market.” Despite the fact that, from a Polanyian perspective, labour is considered as a fictitious commodity, it must be noticed that “it is not even a commodity at all: the relationship established by the owner of labor power when he enters a labor relationship is not a commercial relationship (he is not really selling a specified amount of labor) but a social relationship (he is alienating his labor capacity for a specified time)” (López Calle, 2012: 82).

²³ In this transformation, Polanyi holds, “human society had become an accessory of the economic system. We recall our parallel between the ravages of the enclosures in English history and the social catastrophe that followed the Industrial Revolution. Improvements (...) bought at the price of social dislocation. If the rate of dislocation is too great, the community must succumb in the process” (GT: 79).

2.2 The Material Conditions for Freedom: Property and Material Independence

According to Marx, the use of the common resources and small private property in land allowed the bulk of the agrarian population to live “without having to ask for permission.” What the process of enclosure meant was the “dissolution of small private property based on personal labour,” a condition which had ensured small peasants a certain amount of material independence that, up to a certain extent, made them free from having to sell their labour power in order to subsist.²⁴ Clearly, Polanyi subscribed this Marxian logic: the Poor Law Reform of 1834 had led the working poor to beggary. Indigence, massive unemployment, land privatization, and the extinction of the poor laws, mean that the “right to live was abolished” (GT: 86).

As we have already insisted, a republican civil society can only be made up of free individuals, those who are materially autonomous and are not compelled to ask for favours or sell themselves to others. As Aristotle contends (*Pol.*, I, 1317b), the main “factor of this freedom is to govern and be governed in turn.²⁵” Property as material independence appears once again as the necessary institution by which the popular classes can become part of the civil body as full-right, or *sui iuris*, citizens (hence the Marxian term “alienation” as an expression of *alieni iuris*). David Casassas (2005: 239) puts it this way:

For the bulk of the republican tradition, whether aristocratic or democratic, we can only speak of civil society when there exists an association of individuals who are free and equal—*equal* meaning *equally free*—within a community where each and every one of its full-right members enjoys material independence, that is, a guaranteed sphere of autonomous social existence.

The constitution of the civil body, the establishment of republican citizenship, is then founded upon the relationship between freedom and property,

²⁴Marx might add that “a worker’s private property in his own means of production is the foundation of petty industry (this) is an essential condition for the development of social production and of the *free individuality of the labourer himself*. (...) it flourishes (...) it attains its adequate classical form, only where the labourer is the private owner of his own means of labour set in action by himself: the peasant of the land which he cultivates, the artisan of the tool he handles as a virtuous” (Marx, 1906, VIII, XXXII; emphasis added).

²⁵ In fact, it was not only Aristotle who contended this, since “almost all the great theorists of citizenship (...) have believed that, in order to be a citizen of a *polis*, in order to be able to participate fully in public life, one needed to be in a certain socio-economic position (...) People could not act as citizens at all, or could not be expected to act well in the political sphere and to make adequate decisions, unless some attention was paid to matters of their wealth, their well-being and their social and economic status” (King & Waldron, 1988: 425).

where political-institutional arrangements must always foster and ensure the “autonomous social existence” — or material independence — of the whole group of citizens. According to Birnbaum & Casassas (2008: 77), “material independence”:

constitutes a necessary condition for republican freedom as non-domination, this being understood as the freedom a person enjoys when capable of making choices in all domains of life with the security that nobody will have the possibility to arbitrarily interfere in the decisions she might make with regard to her own life plans.

This is why the republican tradition advocates the creation of an institutional order which would be capable of fulfilling such requirements. According to this, for Polanyi's view the key point here would be not only that self-regulating markets — and, therefore, absolute freedom over one's property — produce desolation in the form of poverty or generalized hunger, but essentially that the final consequence of economic disembeddedness is the inability of the dispossessed to govern themselves and others, i.e., the negation of democracy itself. For this reason, Polanyi advocated the promulgation of those “public interferences” (democratically enacted laws) needed to sustain and extend freedom, since “he who obeys the law he himself sanctioned is free” (Polanyi, 2012: 321).

Surely, speaking of property in the classical world amounts to speaking mostly of (individual or collective) land ownership. However, the modern version of republican material independence is quite different.²⁶ However, “the basic idea

²⁶ There are deep differences between classical and modern property. While “the purpose of feudal property seems to have been essentially to organize a territory politically and economically,” [in the] “modern property” (...) “the accent has been placed exclusively on the economic organization of production and the facilitation of a system of private exchanges” (Trazegnies, 1978: 78). The pre-modern logical unity of property has been broken in contemporary legislations, through its dissociation into a bundle of rights (use, disposition, organization, withdrawal, alienation, etc.) that all together constitute the essence of the classical idea of property. Present property also refers to the right to dispose freely of one's own (free) labour. Polanyi, like modern republican thought, was aware that the Jeffersonian free-holding citizenship ideal was no longer feasible. In short, the industrial developments “placed the republican association of freedom with economic independence [property] under enormous pressure, and the later 19th century saw a series of efforts to reconcile the ideal of independent proprietorship with the demands of industrial production. The most notable of these was the effort to promote workers' cooperatives as an alternative to wage labour, and thus to replace the hierarchical and authoritarian capitalist factory with a workplace based on principles of individual autonomy and collective self-rule” (MacGilvray, 2011: 164-5). It is by replacing the classical individual ownership of land by a collective and cooperative ownership of the means of production that Polanyi forged his ideal of industrial democracy and cooperative production, which in turn, should be understood as heirs of the republican tradition. In order to follow the historical transition of the republican material independence, see: Domènech (2004a) and Gourevitch (2013)

is still there” suggests Carole Rose (1996: 347), “that property nurtures the independence necessary for political participation,”

but in its modern permutation, this idea becomes a platform for distributive rights. The modern form of the independence argument is that all people should have a voice in the political order, but to acquire that voice they need a secure baseline of property –and if necessary, this baseline must be secured by redistribution. (...) baseline entitlements are just as apt to be described as the right to employment, health care, shelter, or the right to such human capital as education.

As Rose points out, to ensure a certain amount of protection vis-à-vis arbitrary interference in the market society is also part of the duty of the so-called welfare state regimes. In effect, the labour legislations of a protective and universalizing kind implemented and constitutionalized from 1945 on (including the rights to strike, to organize trade unions, to collective bargaining, to unemployment and redundancy compensations, to retirement pensions, etc.) may well be understood as a further exponent of the legal codification of the principles of republicanism, principles which were in turn inherited by the mid twentieth century socialist and trade union movements.²⁷ The right to “free labour”, William H. Simon contends, is a “set of basic and uncontroversial rights associated with (republican) citizenship (...) These rights are reflected in the thirteenth amendment prohibition of ‘involuntary servitude,’ the statutory criminalization of ‘peonage,’ and the common law’s refusal to specifically enforce ‘personal service’ contracts” (Simon, 1991: 1335).

In short, wage labour contracts are a legal — and above all political — institution by which workers are partially protected from possible despotic behaviours by their employers, i.e., workers are collectively provided with a set of political, labour and economic rights which prevent them from being arbitrarily interfered with by their employers. In its modern version, “freedom as non-domination” has been updated in the form of politically and institutionally protected fundamental rights such as the right to free labour subject to political

²⁷ Antoni Domènech expresses this same idea — that of the “political bridling” generated by the political pressures that socialism of republican-democratic tradition exercised upon the despotic relations characterizing the economic-labour sphere — in a more suitable manner when he argues that “the core contribution of the socialist labour movement to contemporary republican democracy (has been) the bottom-up control — however imperfect it be — of the potentially despotic power of large private corporative empires” (Domènech, 2004b: 7). Certainly, Polanyi was pretty aware that “the replacement of private property with socialist relations of production” is a process that “requires first the destruction of the capitalist state and then the creation of a new form of state” (...) “then the passage to socialism cannot be automatic. It can only be the result of a deliberate, collective effort” (Buroway, 2003: 2010).

regulation.²⁸ Polanyi acknowledges the particularity of post-war Europe when he adapts the normative codification of classic republicanism to the new context and highlights the importance of labour regulations. Thus, he argues that

His indefeasible rights must be enforceable under the law even against the supreme powers, whether they be personal or anonymous. (...) No mere declaration of rights can suffice: *institutions are required to make the rights effective*. (...) They must be made to prevail against all authorities (...) The list should be headed by *the right of the individual to a job under approved conditions* (...) This implies *guarantees against victimization however subtle it be*.²⁹

The idea of the need for “institutional protection” of fundamental rights (above all, the right to non-domination) is a concern already expressed in classic republicanism. In the ancient world, the basic issue revolved around the ways of gaining access to property and thus obtaining the condition of citizenship. In market society, though, “citizenship rights are not determined by one’s socioeconomic position — and, in this sense, capitalism can coexist with formal democracy —,” while “civil equality does not directly affect class inequality, and formal democracy leaves class exploitation essentially intact” (Wood, 1995: 201). For this reason, it is in the economic sphere where institutions, laws and political safeguards have to be erected in order to enhance the material independence of citizens and, therefore, minimize the domination relationships which are inherent to market society.

For Polanyi, the ultimate point about the “industrial society” is not the abolition of markets, but their regulation and adaptation to the real needs of the populations. What he proposes is to conduct and to embed economic activity and institutions in order to articulate a “substantive economy” (Polanyi, 2009: 75-97) which enhances the material existence of individuals and of society as a whole. And it is precisely this notion of “substantive economy” that most strongly connects his work to republicanism, since, on the contrary, a “formal” conception of the economy “suffocates all possibility of thinking politically about the relationship between the economic and the social spheres,” when, actually, “the

²⁸ For an examination of the right to “free labor” and its constitutionalization as a legacy of the republican legal code, as well as of the broader legacy of republican property rights, see Rose (2003, 1996); Sunstein (1993); Michelman (1987); Simon (1991), and Gourevitch (2013).

²⁹ Polanyi (GT: 264; emphasis added).

market economy (...) is always a political-economic system (...) a way of organizing the production of goods which demands a social and political organization to suit its structure and its operation” (Prieto, 1996: 23-24). Both for Polanyi and for republicanism, unveiling the political dimension of the economy becomes central in order to construct and give shape to their political-normative project.

3. A Republican Reading of Polanyi’s Work: Economic Disembeddedness, Countermovement and Political Economy

In the previous sections, we have limited ourselves to pointing out how Polanyi’s work is similar to much of the republican line of thought. Such similarity can be inferred from an analysis of, on the one hand, their use of a very similar social ontology and the adoption of an analogous methodological stance, and on the other hand, their equally similar definitions of the concepts of freedom (as non-domination) and property (as a necessary condition for the institution of freedom, i.e., for the prevention of domination relationships). Next, we shall provide a reading in republican-democratic terms of his two greatest contributions: the concept of economic embeddedness and that of countermovement.

One of Polanyi’s major contributions to political thought is his notion of “economic embeddedness” (Granovetter, 1985).³⁰ As we have seen, in archaic societies, land, labour power and money, taken as constitutive elements of economic systems, were subordinated to the logic and operation of other social institutions such as the family, the community or the political institutions (Polanyi, 1957). “Economic disembeddedness,” on the contrary, is the phenomenon by which economic activity became dissociated — disembedded — from these same institutions. The emergence of self-regulated markets required

³⁰ Authors like Geoffrey Hodgson, have pointed out that the Polanyian conception of embeddedness might be contradictory. “Even sympathizers have noted some of the conceptual and analytical problems in Polanyi’s work. Among others, Gareth Dale (2010: 246) wrote that his argument concerning ‘how economies are embedded in societies (...) lacks precision’. Fred Block and Margaret Somers (2014: 91, 94) noted the ‘inconsistencies’ and the ‘ambiguities in his discussion of embeddedness’ in *The Great Transformation*” (Hodgson, 2016: 14).

nothing less than “the division of society into an economic and a political sphere” and, in effect, “rejected political action on that account.” The consequence of it all implied “the recognition of the principle of gain and profit as the organizing force in society” (GT: 178) and, with this, the increasing commoditization of both productive and reproductive social activity. In fact, this self-regulated or pure free market is a utopia, since “free-markets could never have come into being merely by allowing things to take their course” (GT: 145).³¹

According to Carlos de Castro & Andrés Pedreño (2012: 11), the process of “disembedding” responds to two closely related factors: on the one hand, a process of “institutional separation” and, on the other, one of “subordination” of one institution to the other. While economic disembeddedness involved the adoption of differential operational and regulative logics by, on one side, the political institution — the state — and, on the other, the economic institution — the markets —, this same separation led to the subsumption of the former under the mercantile logic of the latter. In other words, “such an institutional pattern [self-regulated market] could not have functioned unless society was somehow subordinated to its requirements” as, he added, “a market economy can exist only in a market society” (GT: 74). The commoditization processes fostered by this disembedding are also de-democratization processes by which the political sphere becomes disembedded from and subordinated to the economic activity. That is why commoditization presupposes the de-democratization, not only of the economic life, but more especially of the political-civil life.

Simultaneously, and as a reaction to this commoditization process, heterogeneous and contradictory movements operating with an opposite logic emerged. This is the famous Polanyian “counter” or “double movement.” As he put it (GT: 136),

the dynamics of modern society has been governed by a double movement: the market expanded continuously but this movement was met by a countermovement checking the expansion in definite directions. Vital though such a countermovement

³¹ In fact, “no serious defender of market mechanisms (...) would defend [absolute] laissez-faire markets, however, but would most likely endorse market mechanisms as complemented and assisted by whatever political institutions are necessary to ensure that they work well.” (Furendal, 2017: 7). It was such a double discourse that Polanyi precisely criticized: “to the politician and administrator laissez-faire was simply a principle of the ensurance of law and order, at minimum cost. Let the market be given charge of the poor, and things will look after themselves” (GT: 135)

was for the protection of society, in the last analysis it was incompatible with the self-regulation of the market, and thus with the market system itself.

It is important to add a few considerations to this. The movement of social defence or “countermovement” constituted an effort to re-embed the economic sphere into the social domain. It was not only an attempt to re-unite both spheres, but to subordinate the economic realm to the political one. Historically, however, it would be a mistake to think that all this countermovement was aimed at re-democratizing the economic and the political lives.³²

And again, it is important to notice the similarity between Polanyi’s understanding and that of democratic republicanism in this point. The process of disembedding involved the dissolution of small private property based on personal labour due to the enclosure and expropriation of natural resources used for common exploitation. This was followed by the extinction of the ways of life, the sociability patterns, the bonds and the communitarian forms of organization that sustained the so-called “moral economy of the English crowd” (Thompson, 1971, 1991) and the French “popular political economy” (Gauthier, 1992, 2015).³³

³² It is rather frequent to mistake this double movement in Polanyi’s theory for progressive movements of emancipation, or even for class struggle in Marxian terms. A good illustration of this is Nancy Fraser’s (2012) critique of the “double movement” for including practices of exclusion or subjugation of category groups such as women. Historian Sandra Halperin also criticizes the idea of a double movement as an essentially cultural phenomenon, which would not recognize “the existence of exploiters and producers and their differential capacities, limitations, and potentialities, (Polanyi’s double movement) is conceived and elaborated largely without references to specific social relations or interest” (Halperin, 2004: 13). If Polanyi’s theory of double movement is useful, it is because it contributes to our understanding of the essentially contradictory dynamics arising around the middle of the nineteenth century and triggered by the colonization of the social mechanisms by the practices of the market system. Class struggle, then, cannot be likened to double movement; rather, it should be understood as an integral part—together with other sociohistorical dynamics and phenomena— of the logic of double movement.

³³ The political economy of the crowd, as well as the popular political economy, were “grounded upon a consistent traditional view of social norms and obligations, of the proper economic functions of several parties within the community, which, taken together, can be said to constitute the moral economy of the poor” (Thompson, 1991: 188-9). The expression “popular political economy” and its opposite, the “tyrannical political economy”, were both originally coined by J. J. Rousseau in his article *Political Economy*, in volume V of Diderot’s and d’Alambert’s *Encyclopédie* of 1755, where he stated: “It would therefore be appropriate to add a further distinction, between the popular and the tyrannical forms of public economy. The first kind is that of every state in which a unity of will and interest reigns between the people and its chiefs; and the second kind necessarily exists everywhere where the government and the people have different interests, and contrary wills as a result” (Rousseau, 1999: 9). Later on, this dichotomistic opposition would be employed by Maximilien Robespierre in his *Discours Sur la Constitution* May 10, 1793, in Robespierre (1910-67, IX: 508).

Indeed, “the embeddedness position is associated with (...) the idea of ‘moral economy’ in history and political science” (Granovetter, 1985: 482).³⁴

Both expressions of this political economy, in turn, exhibited a strong philosophical substratum coming from the tradition of natural law,³⁵ including writers such as Bartolomé de las Casas and Francisco de Vitoria, Locke, Paine, Kant, Robespierre or Jefferson, among others. What we should like to highlight here is that the central element around which this natural law tradition revolves is the defence of the right to existence as the guiding principle for both the civil and the economic life. In short, the *natural* right to existence is given precedence over the *civil* right to property, which can only be justified as a means to guarantee the former. Two of the main exponents of this natural law trend within modern republicanism were Maximilien Robespierre and Thomas Jefferson. In 1792, the French explicitly incorporated natural law to his political program to defend that the most important of all imprescriptible rights is the right to existence:

The first social law is therefore the one that guarantees to all the members of the society the means to exist. All others are subordinate to this. The property has not been instituted nor guaranteed for anything else than to cement it. They have properties, first, to live. It is not true that property can ever oppose to the subsistence of men. The foods necessary for man are as sacred as life itself.³⁶

Three years later, Jefferson expressed the same natural law base, although he adopted a certainly different ideological stance:

Whenever there is in any country, uncultivated lands and unemployed poor, it is clear that the laws of property have been so far extended as to violate natural right. The earth is given as a common stock for man to labour & live on.³⁷

As we can see, the main concern of both authors is the violation of natural law (first and foremost, of the right to existence). For the Jacobin, the unlimited freedom of commerce and the idea of property as a “sacred right” that were consecrated in the protoliberal Constitution of 1791 constituted an affront to

³⁴ Other authors, like Jeremy Adelman, also assign to the Hungarian a framework considered as “moral economy”. Nevertheless, Adelman considers that his “moral economics did as much to obscure the nature of global interdependence as it did to reveal the perils of leaving the invisible hand to its own devices” (Adelman, 2017).

³⁵ About this natural law substratum, see Bosc (2009, 2011); Gauthier (1992, 2015); and Gauthier & Ikni (1988). For a sound historical reconstruction of the role of natural law in political thought, see Tierney (1982).

³⁶ *Discours Sur les Subsistances*, December 2, 1792, in Robespierre (1910-67, IX: 112-3).

³⁷ Letter to Reverend James Madison, October 28, 1795, in Jefferson (1904-5, VIII: 196).

natural law (Bosc, 2013: 92-114). For the Virginian, who brandished the flag of the yeomanry, with their freedom and their self-sufficiency based on land ownership, the growing industry, the “moneyed interest” and the corruption of the “pseudo-aristoi” from the North perverted both the virtue and the self-government of the republican citizenry, undermining their material base of existence in favour of an increasingly plutocratic and oligarchic society, as it had happened in senatorial Rome (Domènech, 2004a: 59-72). This is why both revolutionaries can be ascribed to the same tradition of republican freedom (Richard, 1995: 29; Rosenberg, 1966: 17-26).

Polanyi never referred to the notion of natural law.³⁸ However, this did not prevent him from agreeing with both republican thinkers in his analysis: the enclosures, the gradual commoditization of the economy and of the basic resources (among these, the “fictitious commodities”), and the consequent “disembedding” of economy are irremediably at odds with the very possibility of existence of a democratic system and, ultimately, they threaten to put an end to society itself. Once again, Polanyi does not employ the language of natural law. His point of departure is different here from that of natural law republicans. Still, he gets to the same point of arrival: the new enclosures (Boyle, 2003), the mercantilization of both natural resources (Federici, 2004; Rai & Eisenberg, 2003) and artificial ones (Epstein, 2001; Orsi, 2002), the liberalization of property rights (Singer, 1996; Simon, 1991), and the consequent economic uncoupling and self-regulated markets, do not only constitute a threat to the “natural right” to existence, but also to the very existence of society. He uses a different language indeed, but with the same meaning.

Freedom was always the normative goal of republicanism — be it of an aristocratic or a democratic kind. And, in order to achieve it, a necessary condition was the establishment of suitable, politically designed markets. This is the basic outline of the so-called republican political economy, or the “political economy of democracy” (Casassas & De Wispelaere, 2016). As some scholars has

³⁸ Among other reasons, because the reactionary-Thermidorian program leveraged by Benthamian utilitarianism had already erased all traces of it, as, for Bentham (1931: I, 7–9), rights are “nonsense” and human rights are “nonsense upon stilts.”

also noted, when understood as an “instituted processes of economic democratization” (Mendell, 2003: 7)³⁹, Polanyi’s “countermovement” bears a close resemblance to this republican political economy because of the fact that he proscribes the erosion that self-regulated markets inflict on society and on collective and individual self-government. In particular, the most democratic wing of this “countermovement” as this republican political economy alike, seem to be normatively motivated by the necessity to politically bridle the economic activity and institutions in order to offer a more democratic response to the productive and reproductive needs of contemporary societies. Nevertheless, in general, the republican stance can be seen as an option that falls within the “countermovement” logic, which is to say that the republican political economy can be interpreted as an effort to re-embed the economic activity and institutions (especially the institution of property) in its “social ontology” and, consequently, as an effort to subject these to its own political-philosophical principles, most notably, the principle of safeguarding the freedom of citizens, whether the many or the few.

If we have suggested that Polanyi bears a special connection to the more democratic wing of republicanism, this is not because he believes that the aim of the movement for social self-protection should be limited to that of (re-)integrating the economic sphere into society, subordinating it and politicizing it. It is because such re-politicization was aimed at re- democratizing the operation and the logic of markets themselves. As a consequence, Polanyi’s yearning for “democratic socialism,⁴⁰” when understood as a possible reformulation of the democratic republican political economy, could be interpreted as a kind of

³⁹ Marguerite Mendell defends that “Polanyi’s writings on economic democracy, his proposal for a functional democracy (functional socialism), influenced by the guild socialism of G. D. H. Cole, the writings of Robert Owen, and especially those of Otto Bauer and the experience of ‘Red Vienna’ (1917-34), and his writings on education, contribute towards a conceptualization of contemporary processes of institutionalization, in particular, to what I have called instituted processes of economic democratization”.

⁴⁰ For a more in-depth account of Polanyi’s “democratic socialism”, see Polanyi (1960); and Dale (2016: 218-87). As Marguerite Mendell states, “Karl Polanyi’s writings on economic democracy, his proposal for a functional democracy (functional socialism), influenced by the guild socialism of G. D. H. Cole, the writings of Robert Owen, and especially those of Otto Bauer and the experience of ‘Red Vienna’ (1917-34), and his writings on education, contribute towards a conceptualization of contemporary processes of institutionalization, in particular, to what I have called instituted processes of economic democratization” (Mendell, 2003 :7).

countermovement, not only to re-embed the economy and subordinate it to the social institutions, but also to re-democratize the working of markets and, by extension, the present market society as a whole.

4. Concluding Remarks: Freedom and Self-Regulated Markets

In the same manner as the republican tradition was always concerned about the material conditions of existence in order to implement its ideal of freedom, Polanyi always focused on trying to understand the way that the economy was organized within society, or “its place in society” (Polanyi, Arensberg, & Pearson, 1957: 239-241; Polanyi-Levitt & Mendell, 1987), in order to include that same requirement. The key element in Polanyi’s vision is his determination of how the economic processes become institutionalized throughout history and how they constrain or promote (the conditions for) freedom. In this respect, he always contended that the institutionalization of the economic processes should be carried out by universalizing (the conditions for) freedom. This should be done by regulating the markets and politically interfering in them in order to avoid the possibility that certain actors (using the power conferred to them by their economic wealth) impose their own criteria upon all economic activity and, ultimately, upon the political activity as well.⁴¹ While the immediate consequence of such lack of regulation would be a disproportionate accumulation of wealth and the subsequent imposition of arbitrary powers, the most devastating result would be the tragedy of fascism and, with it, the extinction of any kind of freedom:

Nowhere did the liberals in fact succeed in reestablishing free enterprise, which was doomed to fail for intrinsic reasons. (...) Planning, regulation, and control, which they wanted to see banned as dangers to freedom, were then employed by the confessed enemies of freedom to abolish it altogether. Yet the victory of fascism was made practically unavoidable by the liberals’ obstruction of any reform involving planning, regulation, or control (GT: 265).

As this forceful passage reveals, Polanyian freedom is not an abstract, epistemic notion. It has an intrinsically institutional character contingent on

⁴¹ The main interventions or regulations that Polanyi supported are: laws on industrial activity, laws concerning the organization of trade unions, safety regulations for workers, compensations for accidents, housing regulations, the limitation of market prices, laws concerning the exploitation of intercontinental resources, exchange rates, etc. For a more detailed exposition of his proposals, see Polanyi (1936).

society's socioeconomic development and the evolution of its social forces and conflicts. What both Polanyi and the republican tradition are mainly interested in is "the material conditions for freedom." This is what gives so much relevance to the status of economy in society. And it is only after recognizing such a fact, that Polanyi is ready to point out that the dynamics which characterize the "self-regulated" market system constitute a threat to these same "material conditions for freedom". This is also the reason why he — and republicanism — always supported political-economic schemes aimed at re-embedding the economic sphere in the political sphere and subordinating the former to the latter.

As republicanism always believed, and as Polanyi impassively contended, "self-regulated market" and "freedom" are mutually exclusive terms. The republican, and Polanyian, freedom requires the establishment of non-arbitrary mechanisms of intervention in the set of opportunities of the different actors and social groups. In particular, Polanyi and republicanism alike will contend that freedom requires the existence of democratically sanctioned institutional safeguards to restrict the set of opportunities of individuals if such set of opportunities gives them the power to impose his particular will on society as a whole. Because, indeed, freedom "will only exist to the degree in which we will deliberately create new safeguards for its maintenance and, indeed, extension" (GT: 263).

References

Ackerman, B. (1993). *We, the People, Vol. I, Foundations*. Cambridge, Mass: Harvard University Press.

Adelman, J. (2017). Polanyi, the Failed Prophet of Moral Economics. *Boston Review*, May 30 2017. [Online available: <http://bostonreview.net/class-inequality/jeremy-adelman-polanyi-failed-prophet-moral-economics>, access retrieved: September 12, 2017].

Aristotle (1997). *Politics* (trans. A. Rackham). London: William Heinemann, Ltd.

Baum, G. (1996). *Karl Polanyi on Ethics and Economics*. Montreal: McGill-Queen's University Press.

Bentham, J. (1931). *The Theory of Legislation: Civil Code*. London: Ogden.

Berlin, I. (2009 [1958]). *Sobre la libertad* (trans. Á. Rivero). Madrid: Alianza Editorial.

Bertomeu, M. J. (2005). Republicanismo y propiedad. *El Viejo Topo*, 205-206, 84–89.

Bertomeu, M. J. (2004). De la apropiación privada a la adquisición común originaria del suelo. Un cambio metodológico 'menor' con consecuencias políticas revolucionarias. *Isegoría*, 30, 141–147.

Birnbaum, S. & Casassas, D. (2008). Social Republicanism and Basic Income, in: D. LEIGHTON & S. WHITE, (eds.). *Building a Citizen Society: The Emerging Politics of Republican Democracy*. London: Lawrence & Wishart, 75–82.

Block, F. (2003). Karl Polanyi and the Writing of 'The Great Transformation'. *Theory & Society*, 32 (3), 275–306.

Block, F. & Somers, M. (2014). *The Power of Market Fundamentalism: Karl Polanyi's Critique*. Cambridge, MA: Harvard University Press.

Bosc, Y. (2013). Robespierre Libéral. *Annales historiques de la Révolution française*, 1, 95–114.

Bosc, Y. (2011). Liberté et propriété. Sur l'économie politique et le républicanisme de Condorcet. *Annales historiques de la Révolution française*, 4, 53–82.

Bosc, Y. (2009). Le conflit des conceptions de la république et de la liberté : Thomas Paine contre Boissy d'Anglas, in: M. Belissa Y. Bosc, & F. Gauthier (eds.). *Républicanismes et droits naturels à l'époque moderne. Des humanistes aux révolutions des droits de l'homme et du citoyen*. Paris: Kimé, 101–115.

Boyle, J. (2003). The Second Enclosure Movement and the Construction of the Public Domain. *Law and Contemporary Problems*, 66, 33–74.

Burawoy, M. (2003). For a Sociological Marxism: The Complementary Convergence of Antonio Gramsci and Karl Polanyi. *Politics & Society*, 31, 193–261.

Casassas, D. (2013). Adam Smith's Republican Moment: Lessons for Today's Emancipatory Thought. *Economic Thought*, 2 (2), 1–19.

Casassas, D. (2011). *La ciudad en llamas. La vigencia del republicanismo comercial de Adam Smith*. Barcelona: Montesinos.

Casassas, D. (2007). Basic Income and the Republican Ideal: Rethinking Material Independence in Contemporary Societies. *Basic Income Studies*, 2 (2), 1–7.

Casassas, D. (2005). Sociologías de la elección y nociones de libertad: la Renta Básica como proyecto republicano para sociedades de Mercado. *Isegoría*, 33, 235–248.

Casassas, D. & De Wispelaere, J. (2016). Republicanism and the political economy of democracy. *European Journal of Social Theory*, 19 (2), 283–300.

Casassas, D. & Raventós, D. (2007). Propiedad y libertad republicana: La Renta Básica como derecho de existencia para el mundo contemporáneo, *Sin Permiso*, 2, 35–69.

Dale, G. (2016). *Karl Polanyi. A Life on the Left*. New York: Columbia University Press.

Dale, G. (2010). *Karl Polanyi: The Limits to the Market*. Cambridge: Polity Press.

Dagger, R. (2006). Neo-republicanism and the civic economy. *Politics, Philosophy & Economics*, 5 (2), 151–173.

De Castro, C. & Pereño, A. (2012). El péndulo de Polanyi: de la desdemocratización a la resistencia social. *AREAS. Revista Internacional de Ciencias Sociales*, 31, 8–24.

Domènech, A. (2005). Entrevista político-filosófica, in: M. J. Bertomeu, A. Domènech, & A. De Francisco (eds.). *Republicanism y democracia*. Buenos Aires: Miño y Dávila, 281–314.

Domènech, A. (2004a). *El eclipse de la fraternidad: una revisión republicana de la tradición socialista*. Barcelona: Crítica.

Domènech, A. (2004b). Basic Income and the Present Threats to Democracy, Workshop procedure n. 10, Basic Income European Network (September 19-20, 2004), Barcelona. (Available online: www.basicincome.org/bien/pdf/2004Domenech.pdf).

Domènech, A. (2000). Individuo, comunidad y ciudadanía, *Contrastes. Revista Interdisciplinar de Filosofía*, 5, 27–42.

Domènech, A. & Bertomeu, M. J. (2005). Algunas observaciones sobre el método y la substancia normativa en el debate republican, in: M. J. Bertomeu, A. Domènech, & A. De Francisco (eds.). *Republicanism y Democracia*. Buenos Aires: Miño y Dávila, 21–40.

Ely, J. W. Jr. (1992 [1938]). *The Guardian of Every Other Right: A Constitutional History of Property Rights*. Oxford: Oxford University Press.

Epstein, R. A. (2001). Intellectual Property: Old Boundaries and New Frontiers, *Indiana Law Journal*, 76 (4), 803–827.

Federici, S. (2004). *Caliban and the Witch. Women. The Body and Primitive Accumulation*. New York: Autonomedia.

Fraser, N. (2012). Can society be commodities all the way down? Polanyian reflections on capitalist crisis, Working Paper n.18, Fondation Maison des Sciences de l'Homme, Paris.

Furendal, M. (2017). Defining the duty to contribute: Against the market solution. *European Journal of Political Theory*, February 20, 2017 [Online available: <https://doi.org/10.1177/1474885117693401>, access retrieved: September 12, 2017].

Gauthier, F. (2015). Political Economy in the Eighteenth Century: Popular or Despotic? The Physiocrats Against the Right to Existence. *Economic Thought*, 4, 47–66.

Gauthier, F. (1992). *Triomphe et mort du droit naturel en Révolution, 1789-1795-1802*. Paris: Press Universitaires de France.

Gauthier, F. & Ikni, G. R. (comps.). (1988). *La Guerre du blé au XVIIIe siècle. La critique populaire contre le libéralisme économique au XVIIIe siècle*. Paris: La Passion.

Goodin, R. E. (2003). Folie Républicaine. *Annual Review of Political Science*, 6, 55–76.

Gourevitch, A. (2013). Labor Republicanism and the Transformation of Work. *Political Theory*, 41 (4), 591–617.

Granovetter, M. (1985). Economic Action and Social Structure: The Problem of Embeddedness. *The American Journal of Sociology*, 91 (3), 481–510.

Halperin, S. (2004). *War and Social Change in Modern Europe. The Great transformation Revisited*. Cambridge: Cambridge University Press.

Harrington, J. (1992 [1656]). *The Commonwealth of Oceana and A System of Politics*, J. G. A. Pocock (ed.). New York: Cambridge University Press.

Harvey, D. (2003). *The New Imperialism*. Oxford: Oxford University Press.

Hodgson, G. (2016). Karl Polanyi on economy and society: a critical analysis of core concepts. *Review of Social Economy*, 75, (1), 12–40.

Holmes, S. & Sunstein, C. (1999). *The Cost of Rights. Why Liberty depends on taxes*. New York & London: W.W. Norton & Company.

Jefferson, T. (1904-5). *The Works of Thomas Jefferson*. XII Vols., P. L. Ford (ed.). New York & London: The Knickerbocker Press.

King, D. & Waldron, J. (1988). Citizenship, Social Citizenship and the Defence of Welfare Provision. *British Journal of Political Science*, 18, 415–443.

Laín, B. (2014). Karl Polanyi, republicanismo democrático y los fundamentos materiales de la libertad. *Encrucijadas*, 7, 112–133.

López Calle, P. (2012). La constitución del trabajo como mercancía: el caso español. *AREAS. Revista Internacional de Ciencias Sociales*, 31, 81–94.

MacGilvray, E. (2011). *The Invention of Market Freedom*. Cambridge: Cambridge University Press.

Marx, K. (1994 [1875]). Critique of the Gotha Programme. L. H. Simon (ed.). *Karl Marx. Selected Writings*. Indianapolis & Cambridge: Hackett Publishing Company, Inc.

Marx, K. (1906). *Capital. A Critique of Political Economy*. (trans. S. Moore & E. Aveling). New York: The Modern Library, Random House, Inc.

Meade, J. (1964). *Efficiency, Equality, and the Ownership of Property*. London: George Allen & Unwin.

Mendell, M. (2003). Karl Polanyi and Instituted Process of Economic Democratization, Polanyian Perspectives on Instituted Economic Processes, Development and Transformation. Workshop proceeding, Centre for Research on Innovation and Competition, University of Manchester October 23-25, 2003.

Michelman, F. I. (1987). Possession vs. Distribution in the Constitutional Idea of Property. *Iowa Law Review*, 72, 1319–1350.

Mundó, J. (2015). Political Freedom in Locke's Republicanism. in: Y. Bosc, R. Dalisson; J. Y. Frégné, C. Hamel y C. Lounissi (dirs.). *Cultures des républicanismes*. París: Éditions Kimé, 103–116.

Mundó, J. (2005). Autopropiedad, derechos y libertad (¿debería estar permitido que uno pudiera tratarse a sí mismo como a un esclavo?). in: M. J. Bertomeu, A. Domènech, & A. de Francisco (eds.). *Republicanism y democracia*. Buenos Aires: Miño y Dávila Editores, 187–208.

O'Neill, M. & Williamson, T. (2012). *Property-Owning Democracy: Rawls and Beyond*. Boston: Wiley-Blackwell.

Orsi, F. (2002). La constitution d'un droit de propriété intellectuelle sur le vivant aux Etats-Unis: origine et ignification économiques d'un dépassement de frontière. *Revue d'Economie Industrielle*, 99, 65–86.

Pettit, P. (2014). *Just Freedom: A Moral Compass for a Complex World*. New York: W. W. Norton & Company, Inc.

Pettit, P. (2006). Freedom in the Market. *Politics, Philosophy & Economics*, 5 (2), 131–149.

Pettit, P. (1997). *Republicanism. A theory of Freedom and Government*. Oxford: Clarendon Press.

Pocock, J. G. A. (1975). *The Machiavellian Moment. Florentine Political Thought and the Atlantic Republican Tradition*. Princeton: Princeton University Press.

Polanyi, K. (2012). *Karl Polanyi. Textos Escogidos*. J. L. Laville, M. Mendell, K. Polanyi Levitt, & J. L. Coraggio (eds.). Buenos Aires: Clacso.

Polanyi, K. (2009 [1977]). *El sustento del hombre*. (trans. E. G. Parro). Madrid: Capitán Swing.

Polanyi, K. (1960). Personal Statements, Karl Polanyi Archive, Container 30, Folder 2, Karl Polanyi Institute of Political Economy, Concordia University, Montreal, Canada.

Polanyi, K. (1959). Galbraith's Farewell to Poverty. Karl Polanyi Archive, Container 37, Folder 11, Karl Polanyi Institute of Political Economy, Concordia University, Montreal, Canada.

Polanyi, K. (1957). Aristotle Discovers The Economy, in: K. Polanyi, C. M. Arensberg, & H. W. Pearson (eds.). *Trade and Market in the Early Empires, Economies in History and Theory*. Illinois: The Free Press & The Falcon's Wing Press, 64–94

Polanyi, K. (1947). Our Obsolete Market Mentality. *Commentary*, 3, 109–117.

Polanyi, K. (2001 [1944]). *The Great Transformation. The Political and Economic of Our Time*. Boston, Mass: Beacon Press.

Polanyi, K. (1936). Economic Liberalism vs Interventionism. Karl Polanyi Archive, Container 15, Folder 4, Karl Polanyi Institute of Political Economy, Concordia University, Montreal, Canada.

Polanyi, K. (1934). Marxism Re-Stated. *New Britain*, July 4, 1934.

Polanyi, K. (1931). The Paradox of Freedom. Conference at Birkbeck College, Container 21, Folder 1, Karl Polanyi Institute of Political Economy, Concordia University, Montreal, Canada.

Polanyi, K.; Arensberg, C. M. & Pearson, H. W. (eds.). (1957). *Trade and Markets in the Early Empires*. Illinois: The Free Press & Falcon's Wing Press.

Polanyi-Levitt, K. & Mendell, M. (1987). Karl Polanyi: His Life and Times. *Studies in Political Economy*, 22, 7–39.

Prieto, C. (1996). Karl Polanyi: crítica del mercado, crítica de la economía. *Política y Sociedad*, 21, 23–34.

Rai, A. K. & Eisenberg, R. S. (2003). Bayh-Dole Reform and the Progress of Biomedicine. *Law and Contemporary Problems*, 66 (288), 289–314.

Raventós, D. (2007). *Basic Income: The Material Conditions of Freedom*. London: Pluto Press.

Rawls, J. (2001). *Justice as Fairness : A Restatement*. Cambridge, MA: Harvard University Press.

Richard, C. J. (1995). *The Founders and the Classics. Greece, Rome, and the American Enlightenment*. Cambridge, Mass: Harvard University Press.

Robespierre, M. (1910-67). *Oeuvres complètes de Maximilien Robespierre*. X Vols., E. Leroux, G. Laurent, M. Bouloiseau, G. Lefebvre, A. Soboul, & J. Dautry (eds.). Paris: Société des Études Robespierristes, Libraire Félix Alcan, & Press Universitaires de France.

Rose, C. M. (2003). Romans, Roads, and Romantic Creators: Traditions of Public Property in the Information Age. *Law and Contemporary Problems*, 66, 89–110.

Rose, C. M. (1996). Property as the Keystone Right? *Notre Dame Law Review*, 71 (3), 329–366.

Rosenberg, A. (1966 [1937]). *Democracia y socialismo. Aporte a la historia política de los últimos 150 años*. Buenos Aires: Editorial Claridad.

Rousseau, J. J. (1999). *The Social Contract*. (trans. by Christopher Betts). Oxford: Oxford University Press.

Simon, W. H. (1991). Social-Republican Property. *University of California Law Review*, 38, 1335–1415.

Singer, J. W. (1996). No Right to Exclude: Public Accommodations and Private Property. *Northwestern University Law Review*, 90 (4), 1283–1499.

Skinner, Q. (2008). *Hobbes and Republican Liberty*. Cambridge: Cambridge University Press.

Skinner, Q. (1998). *Liberty before Liberalism*. Cambridge: Cambridge University Press.

Sunstein, C. R. (2004). *The Second Bill of Rights: FDR's Unfinished Revolution, and Why We Need It More than Ever*. New York: Basic Books.

Sunstein, C. R. (1993). On Property and Constitutionalism. *Cardozo Law Review*, 14, 907–35.

Trazegnies, G, F. (1978): La transformación del derecho de propiedad. *Derecho PUCP*, 33, 75–104.

Thompson, E. P. (1991). *Customs in Common*. London: Penguin Books.

Thompson, E. P. (1971). The Moral Economy of the English Crowd in the Eighteenth Century. *Past & Present*, 50, 76–136.

Tierney, B. (1982). *Religion, Law, and the Growth of Constitutional Thought. 1150-1650*. Cambridge: Cambridge University Press.

White, S. (2011). The Republican Critic of Liberalism. *Critical Review of International Social and Political Philosophy*, 14 (5), 561–579.

Wood, E. M. (2008). Why it Matters. *London Review of Books*, 30 (18), 3–6.

Wood, E. M. (1995). *Democracy Against Capitalism*. Cambridge, UK: Cambridge University Press.

Zaman, A. (2016). The Methodology of Polanyi's Great Transformation. *Economic Thought*, 5 (1), 44–63.



O DEBATE NA GRANDE TRANSFORMAÇÃO: LUDWIG VON MISES, FRIEDRICH HAYEK E KARL POLANYI SOBRE CAPITALISMO, SOCIALISMO E MERCADOS

THE DEBATE IN THE GREAT TRANSFORMATION: LUDWIG VON MISES, FRIEDRICH HAYEK E KARL POLANYI ON CAPITALISM, SOCIALISM AND MARKETS

João Rodrigues*
joaorodrigues@ces.uc.pt

Sumário. *O debate relativo ao cálculo económico em socialismo, iniciado por Ludwig von Mises, contou com a participação de Karl Polanyi e de Friedrich Hayek, moldando de diferentes formas o seu pensamento. Este artigo sintetiza alguns dos contributos para este debate sobre o estatuto e o lugar dos mercados, indicando como contribuiu para consolidar uma abordagem neoliberal à economia política institucionalista, mas sobretudo como moldou a original economia política institucionalista crítica de Karl Polanyi. Esta última foi construída numa discussão acesa com o liberalismo económico mais intransigente, a qual continua a enquadrar muitas questões actualmente relevantes.*

Palavras-chaves: *Capitalismo, Socialismo, Mercado, Neoliberalismo.*

Abstract. *The socialist calculation debate, initiated by Ludwig von Mises, counted with Karl Polanyi and Friedrich Hayek among its participants, molding their intellectual endeavors in multiple ways. This article reviews some of their contributions for this debate about the nature and place of markets, indicating how it contributed to consolidate a neoliberal approach to institutionalist political economy, but also Karl Polanyi's original critical institutionalist political economy. The latter was built through an intense debate with intransigent economic liberalism. This discussion still frames numerous important issues.*

Keywords: *Capitalism, Socialism, Market, Neoliberalism.*

* Professor Auxiliar na Faculdade de Economia da Universidade de Coimbra e Investigador no Centro de Estudos Sociais.

o. Introdução¹

Em 1920, Ludwig von Mises inicia um dos debates mais relevantes da história da economia política do chamado breve século XX, o debate relativo ao cálculo económico em socialismo, ao responder, em artigo, aos socialistas que defendiam uma economia sem mercados, totalmente guiada pela deliberação política explícita. Karl Polanyi seria um dos participantes socialistas da primeira fase, em língua alemã, do debate, antes da sua mais conhecida fase em língua inglesa, a partir dos anos trinta, com Friedrich Hayek, discípulo de Mises, no Reino Unido a partir de 1931, ou Oskar Lange, economista neoclássico e socialista de origem polaca, a ocuparem aí posições antagónicas.

Este debate foi importante para a transformação do liberalismo clássico em neoliberalismo no período entre as guerras. O neoliberalismo foi em parte definido pela nova retórica, também função da natureza dos oponentes socialistas, a favor de uma ordem capitalista de mercado renovada. Estava centrado numa visão essencialista sobre os poderes desta instituição de coordenação na geração de acção económica racional, orientada pelo cálculo egoísta, por um lado, e na difusão de informação e aproveitamento do conhecimento disperso, por outro. A contribuição da Escola Austríaca da economia política seria central para este projecto. O debate inicial com Ludwig von Mises e o engajamento com os argumentos de Friedrich Hayek foi também importante para moldar o pensamento socialista de Polanyi e os seus contributos para uma economia política institucionalista crítica das apologias do capitalismo utópico. Este artigo contribui para aprofundar o conhecimento nesta dupla dimensão: por um lado, sintetiza o contributo da Escola Austríaca para a emergência do neoliberalismo, um conjunto de ideias económicas em movimento político, também em oposição ao socialismo, muito antes do seu triunfo no final dos anos setenta; por outro lado, e mais importante na economia do texto, tenta compreender como é que o pensamento de Karl Polanyi foi moldado pelo debate com a tradição do liberalismo económico, em processo de reinvenção, sem deixar

¹ Agradeço os comentários e sugestões de Ana Cordeiro Santos, de Nuno Teles, de Pedro Teixeira e de um avaliador anónimo. Obviamente, assumo sozinho todos os erros e omissões que possam constar deste artigo.

de procurar responder a algumas das dicotomias armadilhadas que esta entretanto procuraria criar.

O artigo está dividido em três secções. Na primeira, farei uma breve síntese, em parte comparativa, dos contributos de Ludwig von Mises e de Friedrich Hayek no quadro do debate do cálculo económico socialista e para lá dele. Na segunda, procederei a uma descrição de como o pensamento de Karl Polanyi foi explicitamente moldado pelo debate com Mises. Finalmente, na terceira parte, farei uma breve apreciação da forma como o quadro conceptual de Karl Polanyi também resulta de um debate implícito com Hayek, ajudando a explicar a sua inadvertida capacidade de nos legar um aparato conceptual capaz de pensar criticamente o neoliberalismo.

1. O socialismo não saberia calcular

É comum afirmar-se que o debate do cálculo económico em socialismo, que decorreu entre o início da década de vinte e o fim da de quarenta do século passado, contribuiu para a consolidação da Escola Austríaca de economia política, entendida agora como abordagem intransigentemente oposta a todas e quaisquer formas de socialismo, mas também à economia neoclássica e à sua fixação com o equilíbrio, geral ou parcial (Hodgson, 1999). Aliás, a economia convencional oferecia, precisamente, uma justificação teórica para a economia socialista, dotada de mecanismos que mimetizariam os mercados, graças a contributos nos anos trinta, em resposta a Ludwig von Mises, como os de Oskar Lange. Entretanto, os veredictos sobre este debate têm mudado: a ideia de que os socialistas neoclássicos, como Lange (1936), tinham conseguido responder com sucesso aos argumentos teóricos acerca da inexequibilidade de uma economia socialista racional, e de que os de Hayek tinham considerações de índole mais prática, deu lugar a uma opinião hegemónica que considera que a Escola Austríaca conseguiu defender de forma eficaz a inferioridade económica da planificação socialista, quando comparada com as instituições económicas concretas associadas ao capitalismo (Hodgson, 1999). De facto, em 1942, o economista e historiador do pensamento económico liberal Joseph Schumpeter alinhava com a ideia da plausibilidade do argumento socialista contra a crítica à

planificação, enquanto em 1989, por contraste, o economista e historiador do pensamento económico social-democrata Robert Heilbroner afirmava que Mises e Hayek estavam correctos na sua crítica (Schumpeter, 1942/1972; Heilbroner, 1989).

O contributo da Escola Austríaca pode agora ser entendido como representando a ala neoliberal da tradição institucionalista na economia, e a trajetória socioeconómica das experiências socialistas “realmente existentes”, assim como a viragem neoliberal nos países ocidentais, tanto ao nível intelectual como político, também ajudaram a consolidar esta perspectiva (Rodrigues, 2013). É por isso importante tentar sistematizar os contributos de Mises e de Hayek contra o socialismo, mas também, ainda que muito mais brevemente, as suas visões sobre as funções do Estado no capitalismo.

Publicado em 1920, em alemão, o artigo de Ludwig von Mises, “cálculo económico numa comunidade socialista”, é justamente apresentado como o artigo fundador da controvérsia sobre a viabilidade do cálculo económico em socialismo. A este artigo seguiu-se a publicação, em 1922, também em alemão, de uma das suas principais obras: *Socialismo* (Mises, 1981). Mises (1920/1935) tinha um duplo objectivo. Em primeiro lugar, confrontar o movimento socialista e a tradição marxista em ascensão com a sua insuficiente reflexão sobre as consequências económicas que poderiam advir da socialização integral dos meios de produção e da superação das transacções mercantis, ou seja, do que para ele seria a implementação de uma economia socialista. Em segundo lugar, demonstrar que a propriedade privada e o mercado – as instituições centrais do capitalismo – são absolutamente indispensáveis e insubstituíveis para a existência do cálculo económico racional, base do progresso civilizacional: “o socialismo é a abolição da economia racional” (Mises, 1920/1935, p. 110).

O argumento de Mises estrutura-se em torno de duas ideias. Em primeiro lugar, o cálculo económico racional, que pode e deve guiar a acção humana em economia, pressupõe a existência de preços monetários para os bens de consumo, para o trabalho e para os bens de capital, formados em mercados reais, ou seja, segundo um ideal de mercadorização quase universal. Os mercados, em especial para os bens de capital, são inseparáveis da propriedade privada dos meios de produção. Assim, o cálculo económico racional envolve comparação, o que

pressupõe uma unidade comum – valor monetário dos custos potenciais das diferentes alternativas – ou seja, a comensurabilidade. A comensurabilidade, por sua vez, depende da existência da propriedade privada dos meios de produção e dos mercados para gerarem preços monetários – o nexos comensurabilidade-mercadorização –, central em todo o pensamento de Mises (Rodrigues, 2013). Em segundo lugar, o cálculo económico aparece estreitamente associado ao “problema da responsabilidade e da iniciativa empresariais em socialismo (...) à exclusão da livre iniciativa e responsabilidade individual de que depende o sucesso da empresa privada” (Mises, 1920/1935, p. 116). Estamos aqui perante a relação entre os arranjos institucionais de uma economia, os incentivos gerados e as motivações para a acção económica racional de indivíduos, vistos como psicologicamente egoístas: a esfera dos mercados favoreceria um comportamento empreendedor motivado “só e apenas pelo interesse egoísta em obter lucros e adquirir riqueza” (Mises, 1949, p. 288). O capitalismo, em versão pura, de rivalidade generalizada, é assim a condição para o esforço, compatível com a hipótese egoísta, para avaliar os custos e benefícios pecuniários das diferentes alternativas e fazer escolhas económicas entre as várias afectações possíveis dos recursos, com os resultados a posteriori, expressos em termos de ganhos e perdas, a ditar o acerto dessas escolhas. Isto acontece num cenário de incerteza genuína e numa economia em perpétua mudança, duas características incontornáveis. O socialismo eliminaria as condições institucionais para resolver o problema económico, pelo que seria uma utopia, que produziria efeitos distópicos, socioeconómica e politicamente altamente regressivos, na sua tentativa de institucionalização. Uma impossibilidade a prazo, em suma.

Entretanto, note-se também que, ao contrário da maioria dos seus discípulos no Instituto que leva hoje o seu nome na Universidade de Auburn, no Alabama, Mises não é um libertário de direita, defensor da abolição do Estado, já que a existência de uma verdadeira esfera dos mercados, que coincide exatamente com a esfera onde os indivíduos podem fazer cálculos monetários de custo-benefício relevantes, tem necessariamente de depender da existência de uma esfera burocrática, estatal, regida por considerações de outra índole: “existem áreas de actividade humana em que não pode haver questões de gestão de lucro e onde tem de prevalecer a gestão burocrática” (Mises, 1944, p. v). Para Mises, a esfera do Estado capitalista, que equivaleria por definição à esfera da burocracia,

deve ser reduzida à tarefa de garantir a integridade dos mecanismos de mercado, tarefa essa que considera bastante simples, até dada a ausência de falhas de mercado, se os princípios racionalistas liberais de base utilitarista, por si expostos, não forem dissipados em compromissos com toda a espécie de ideologias “intervencionistas”. Note-se, entretanto, que “medidas que são tomadas com o objectivo de preservar a ordem da propriedade privada não são intervenções” (Mises, 1927, p. 17). Esta divisão de esferas confere intransigência e rigor aparentes aos esforços de Mises, independentemente da sua implausibilidade política perante os dados conhecidos relativos ao dinamismo da chamada economia mista em parte do século XX – como reconhece até quem trabalha no âmbito da Escola Austríaca (Ikeda, 1997). Seja como for, o ideal de uma “sociedade de mercado pura” (Mises, 1949, p. 239) deve estar na base de um esforço deliberado de acção e de persuasão intelectuais, que não dispensa a força e logo o apoio de circunstância a forças iliberais, que salvem o princípio da propriedade privada contra ameaças socialistas, como foi o caso do fascismo italiano (Mises, 1927). Apesar disso, e repetidamente, Mises apresentava nos debates os assuntos em discussão em termos de escolhas e previsões absolutas: “O Homem tem de escolher entre uma economia de mercado ou o socialismo. Não pode escapar a optar entre estas alternativas, adoptando uma posição ‘a meio caminho’, qualquer que seja o nome que lhe dêem” (Mises, 1949, p. 861).

A natureza institucionalista deste argumento, que depende de um sistema socioeconómico historicamente específico – o capitalismo –, ainda que abstractamente definido por Mises, não passou despercebida aos socialistas neoclássicos (Lange, 1936, p. 62). Lange (1936), partindo da ideia da universalidade do problema económico, e apoiando-se na teoria neoclássica dos preços, defendeu que o socialismo poderia criar um procedimento de tentativa e erro, que mimetizaria os mercados, em modo de leiloeiro Walrasiano, no que diz respeito aos bens de capital, e que se combinaria com os mercados de trabalho e de bens de consumo. Deste modo, assegurar-se-ia a coexistência do planeamento, racional e igualitário, com a comensurabilidade. A “concepção algorítmica da racionalidade”, que partilhava com Mises, pelo menos no que diz respeito à vida económica, não seria apenas salvaguardada, mas ainda elevada a um patamar superior em termos de eficiência e de equidade (O’Neill, 1998, p. 115).

Friedrich Hayek, discípulo de Mises, cujo livro *Socialismo* terá sido responsável pela sua conversão “liberal”, foi responsável por trazer contributos como os do seu mestre para um debate que, a partir da década de trinta, se trava fundamentalmente em língua inglesa, beneficiando da sua posição como Professor na *London School of Economics*, a partir de 1931 (Hayek, 1935). Hayek desenvolveu um argumento epistemológico contra a planificação socialista na sua crítica do modelo neoclássico. Este modelo esteve na base do contra-ataque socialista de Lange e de outros. Segundo ele, o modelo neoclássico, com a sua obsessão com a tendência para o equilíbrio final, assume irrealisticamente, desde a génese, a possibilidade da existência de um conhecimento perfeito (Hayek, 1937/1948a). Esta premissa oculta a função real, e, na perspectiva de Hayek (1946/1948d), insubstituível, dos processos de mercado e da concorrência possível a eles associada: gerar preços em permanente mudança, que funcionam como mecanismos de informação para os empreendedores descobrirem as melhores maneiras de ir ao encontro das preferências dos consumidores, com o mínimo de custos e o máximo de benefícios possíveis. Hayek (1945/1948c) assinalou a natureza dispersa e tácita de algumas formas de conhecimento economicamente relevante, bem como a indispensabilidade da descentralização associada aos mercados para mobilizar esse conhecimento, compatibilizando assim os planos de consumidores e produtores. Era esta a base do seu argumento de que os responsáveis pelos planos centrais não poderiam ter qualquer expectativa, mesmo através dos processos de tentativa e erro delineados por Lange, de aceder ou ser capazes de agregar o conhecimento económico contextual e subjetivo de quem participa nos mercados. Além disso, a informação objectiva contida nos preços de mercado (que seriam, simultaneamente, a consequência das escolhas e poder dos agentes económicos e parte da informação que teriam de ter em conta ao tomar as suas decisões) também não poderia existir sem a configuração institucional dos verdadeiros mercados.

Ao encetar esta agenda de investigação contra o socialismo, Hayek tinha bem presente que o trabalho de Mises era “o ponto de partida de todas as discussões acerca dos problemas económicos do socialismo, tanto construtivas como críticas, que quisessem ser levadas a sério” (Hayek, 1935: 33). Porém, dois passos acabariam por afastar Hayek de Mises, logo no debate do cálculo socialista. O primeiro foi metodológico: Hayek (1937/1948a) rejeitava a lógica da

escolha pura, a base da forma de raciocínio apriorística de Mises sobre questões económicas, e defendia que perceber como é que os agentes económicos aprendem implicava dar conta da divisão de facto do trabalho e do conhecimento na sociedade, assim como do processo com que os agentes se deparariam para tornar os seus planos compatíveis; ou seja, implicava reconhecer a dimensão empírica da economia (Caldwell, 2004). O segundo passo foi tanto epistemológico como institucional: Hayek defendeu, por um lado, que a concepção de racionalidade na economia era demasiado exigente; e, por outro, que a dependência da configuração institucional não fora suficientemente escrutinada. Assim sendo, a sua crítica ao socialismo tornou-se simultaneamente uma crítica ao “racionalismo” utilitarista, ou seja, uma crítica à ideia de que, nos mercados, os indivíduos poderiam, através de um quadro de meios e de fins pré-definido, avaliar com toda a clareza as alternativas económicas e os efeitos das instituições particulares, e meticulosamente desenhadas, em que poderiam exercê-las. Acresce que a natureza limitada, tácita e incomensurável de grande parte do conhecimento na posse dos participantes nos mercados significa que a racionalidade não é apenas mais frágil, mas também fundamentalmente dependente das regras sociais e políticas em vigor, que teriam de estar em evolução progressiva, configurando uma ordem dita espontânea e que só um certo tipo de capitalismo poderia garantir, bem como a indispensável liberdade individual para descobrir, ou seja para empreender, que lhe estaria unicamente associada (Hayek, 1944/2009).

Dito isto, é preciso acrescentar que a convergência substancial, tanto intelectual como política, entre Mises e Hayek, no que confere ao tópico do socialismo, é avassaladora (Boettke, 2002). Tal como a insistência de ambos em que cada tentativa de combinação de elementos do capitalismo e do socialismo, numa espécie de solução “a meio caminho”, não teria pernas para andar, degenerando na planificação autoritária e ineficiente ou sendo salva por uma restauração do capitalismo, com todas as instituições associadas (Hayek, 1944/2009; Mises, 1949). Onde Hayek se separa de Mises de forma mais saliente, mas menos escrutinada, é na questão para si central e que lhe tinha sido colocada por Keynes (1944) na sua resposta ao *Caminho para a Servidão*: onde traçar a linha entre os mercados e o Estado – por outras palavras, a estrutura concreta das sociedades de mercado. Como Keynes (1944) assinalou, Hayek (1944/2009)

defendeu que a intervenção estatal na economia corria o risco de levar, em jeito de plano inclinado, ao autoritarismo político, para este indissociável do socialismo, ao mesmo tempo que advogou que o liberalismo não se podia confundir com o *laissez-faire* e que a agenda do Estado liberal era legitimamente ambiciosa, dado que os mercados dependiam de uma esfera não-mercantil relevante.²

Embora Hayek refira várias vezes, e de forma aprovadora, ao longo da sua obra a fórmula de Mises sobre as acções do Estado que seriam necessárias, porque conformes à institucionalização de uma ordem capitalista de mercado, e que por isso se poderiam distinguir do “intervencionismo” anti-mercantil de tipo socialista, a verdade é que Hayek foi muito para lá da ideia de Mises, dita liberal clássica, de um Estado guarda-nocturno. Isto foi assim, devido a um movimento, de novo duplo, de separação em relação a Mises. Em primeiro lugar, Hayek (1944/2009, 1960/2006, 1982) acabou por reconhecer que as instituições centrais do capitalismo eram bem mais variadas e plásticas e que a sua criação e evolução era o resultado de um trabalho político mais exigente, apesar da sua retórica sobre a ordem espontânea. Em segundo lugar, Hayek, ao contrário de Mises, aceitou abrir “a caixa de Pandora das falhas de mercado”, na apta fórmula de um crítico libertário (Block, 1996, p. 347), que o Estado poderia eventualmente corrigir. Na sua síntese sobre a arquitectura institucional da sociedade de mercado, Hayek (1960/2006: 194) defendeu que “é o carácter e não o volume da actividade estatal que é importante”, aproximando-se implicitamente do célebre diagnóstico de Polanyi (1944/2012), segundo o qual a expansão dos mercados não se faz sem a expansão dum certo tipo de intervenção estatal. Antes de escrutinar esta questão, à guisa de conclusão, vale a pena detalhar a forma como pensamento de Karl Polanyi foi moldado por um envolvimento cuidado e continuado com os argumentos da Escola Austríaca.

² Para uma apreciação das implicações da observação de Keynes sobre a evolução do pensamento de Hayek relativamente às contestadas fronteiras entre Estado e mercados, ver Rodrigues (2012).

2. O socialismo saberia não só calcular, mas também deliberar

Uma leitura dos contributos de Karl Polanyi na fase austríaca do debate do cálculo socialista, nos anos vinte, torna clara a importância de Ludwig von Mises neste e noutros debates ao longo do seu percurso intelectual, a maneira como a sua obra foi, de um modo mais geral, visível no seu único livro publicado em vida, *A Grande Transformação*, marcada pelo confronto permanente com o liberalismo económico mais intransigente. Polanyi sempre recusou colocar o debate nos temas propostos por Mises (1920/1935): ou o capitalismo *laissez-faire* baseado na propriedade privada, concebida como o controlo absoluto, assente numa economia de base racional onde o cálculo económico é possível, dada a existência de preços de mercado, ou uma economia centralizada, onde não há guias para a distribuição racional dos meios de produção, porque não existem os alicerces institucionais necessários. Polanyi também não aceitou o conceito central da “economia utópica” de Otto Neurath, ou seja, a ideia de uma economia socializada sem mercados, onde o critério monetário seria substituído por deliberações levadas a cabo por uma agência central – um dos principais alvos da crítica inicial de Mises ao socialismo (Uebel, 2008). Esta dupla rejeição foi formulada em termos de uma crítica dirigida aos “partidários dogmáticos”, tanto liberais como socialistas, que aderiam à tese de uma economia pura, fosse de mercado ou sem mercados (Polanyi, 1922/2008a).

Todavia, Polanyi absorveu ideias de ambas as posições. Na verdade, convergiu com Mises: como este último afirmou em artigo de 1923, republicado como apêndice numa reedição do seu livro *Socialismo*, Polanyi tinha compreendido “que o cálculo económico se relevaria impossível numa economia com uma administração centralizada e sem mercados” (Mises, 1981, p. 474).³ Polanyi (1922/2008a) sustenta que o problema central do capitalismo – o de não dar conta de todos os custos relevantes do processo de provisão – só poderia ser resolvido através de um sistema democrático e associativo fortemente inspirado na proposta do socialismo de guildas de G. D. H. Cole (1920), procurando alargar

³ Esta convergência não legitima uma leitura da crítica de Mises ao socialismo que a reduza à condição de “ácida no seu melhor, histeria ideológica formulada em jargão científico no pior” (Mendell, 1990, p. 67). Polanyi sempre levou Mises mais a sério do que alguns dos seus intérpretes e apoiantes mais influentes.

a voz democrática da esfera política a todas as esferas relevantes para a vida socioeconómica (Mendell, 1990; Maucourant, 2005; Dale, 2010).

Mais ainda, concorda com Mises quando este reconhece que a tradição anti-capitalista, a partir de Marx, não desenvolveu uma “teoria económica positiva” para o socialismo, capaz de fornecer uma orientação teórica clara para o exercício de reconfiguração institucional da economia. Significa isto que a única teoria económica formal disponível era o marginalismo. Apesar disso, a asserção de Rosner (1990, p. 57) de que Polanyi acreditava numa economia planeada baseada em “conceitos de utilidade marginal” é equivocada. Polanyi não antecipou a estratégia de Lange (1936) de recorrer à teoria económica neoclássica para propor um modelo de planificação central com imitações dos mercados ideais da teoria. Ao contrário de Lange, Polanyi (1922/2008a) considerou que a teoria económica não poderia servir de guia, neste estágio de desenvolvimento, para a discussão sobre as instituições económicas socialistas. A solução de Polanyi para este impasse foi defender a possibilidade de uma separação “formal” da contabilidade e da teoria económica. A contabilidade – a “visão numérica da economia” – seria pensada de acordo com as questões fundamentais levantadas pelas prioridades que revelam a natureza institucional dos dois sistemas socioeconómicos exequíveis disponíveis: lucros para o capitalismo; produtividade máxima e produção com objectivos sociais, garantida pelas regras sociais necessárias para promover a justiça social na distribuição, para o socialismo (Polanyi, 1922/2008a).

Estes conceitos ajudam a perceber os contornos institucionais do modelo de socialismo associativo proposto por Polanyi. Antes de nos dedicarmos a eles, vale a pena assinalar que também partilhava com Mises, pelo menos nesta fase inicial, a ideia de que um cálculo como deve ser de todos os custos do processo de provisão, tanto de natureza “técnica” como “social”, requereria a existência de um único critério pecuniário. Só assim se garantiria a sua comensurabilidade e se permitiria uma apreciação racional dos custos “dos ideais da humanidade”, mesmo se os benefícios desses ideais não possam ser traduzidos por esse denominador comum: “um sistema contabilístico que não permita a redução de todas as somas à mesma unidade de cálculo [monetário] não faria muito sentido” (Polanyi, 1922/2008a, p. 299). A existência de preços, sejam negociados entre

associações de produtores e consumidores ou determinados por uma autoridade política, é assim um elemento crucial no sistema de cálculo socialista de Polanyi.

É na crítica ao capitalismo que Polanyi defende uma economia plenamente socializada, pela eliminação da propriedade privada dos meios de produção na base do capitalismo, que considerava estar na origem da maioria dos problemas políticos e socioeconómicos. Não obstante, isto não significa que no socialismo as unidades económicas tivessem de ser geridas completamente de acordo com um plano definido centralmente: o facto de uma autoridade política (a “Comuna”, na terminologia de Polanyi) ser formalmente detentora dos meios de produção não tem de impedir a descentralização de certos direitos, tipicamente associados à propriedade, bem como das respectivas decisões económicas (Polanyi, 1922/2008a).

Rosner (1990) refere que Polanyi criou dois conceitos de produtividade – técnica e social – e usou-os para criticar as instituições do capitalismo e delinear os contornos institucionais do socialismo. A produtividade técnica é convencionalmente definida como a minimização dos factores de produção – incluindo trabalho e recursos naturais – para produzir a quantidade máxima possível de resultados, e o capitalismo é criticado, como costumava acontecer nos círculos socialistas, tanto por promover formas de competição esbanjadoras e propícias a crises entre empresas demasiado pequenas e míopes, como por criar monopólios que contraem deliberadamente a produção. Enquanto a produtividade técnica diz respeito ao campo da produção, a produtividade social pertenceria ao campo do consumo, sendo a questão a de saber se as necessidades do indivíduo, “enquanto ser social consciente”, podem influenciar o processo de provisão. No capitalismo não existe “um órgão de percepção” capaz de identificar os “objectivos sociais, intelectuais e morais da comunidade política na medida em que dependem de meios materiais” (Polanyi, 1922/2008a, p. 293-294). Segundo Polanyi (1922/2008a), apenas as preferências do indivíduo isolado, se apoiadas por dinheiro, são reconhecidas neste sistema, independentemente da perda de “utilidade social” originada pelos custos sociais, geralmente não reconhecidos, do capitalismo. Assim sendo, o desafio perante os socialistas consistiria em criar vários “órgãos de percepção”, capazes de inculcar nos produtores a consciência das prioridades sociais das pessoas, não tanto enquanto consumidores isolados, mas

como consumidores organizados; cidadãos a quem o “direito a viver”, numa perspectiva alargada, é garantido, e que devem ter a capacidade de deliberar acerca das regras sociais que influenciarão os bens que serão fornecidos, como serão fornecidos e com base em que critérios de acesso, para além do modo como os custos e benefícios do trabalho podem ser repartidos de forma justa (Polanyi, 1922/2008a). Polanyi acaba concluindo que o conceito de produtividade só será operacional dentro da esfera “técnica” da produção, e que o conceito de produtividade social terá, na prática, de ser substituído pelo que designa de “legislação social”, incorporando os ideais de justiça e utilidade social duma comunidade socialista dotada de “órgãos de percepção” capazes de deliberar acerca destes fins. No entanto, estes ideais geram custos económicos específicos, e Polanyi está confiante em que alguns deles possam ser distinguidos dos (e depois acrescentados aos) custos “técnicos”, “naturais”, da produção.

A identificação fundamental de todos os custos significativos gerados pelos aspectos técnicos da produção e pela legislação social depende duma configuração institucional específica e original, que Polanyi (1922/2008a) apenas descreve em traços largos. Unidades autogeridas de forma democrática, formando associações locais e regionais através de ramos da indústria representados num congresso nacional de associações de produção de todo o Estado, seriam um dos pilares. Em paralelo com a organização democrática da produção, com os seus diferentes níveis, existiriam associações semelhantes de consumidores, incluindo cooperativas, cujos interesses seriam representados, à escala nacional, pela principal organização política responsável pelos interesses gerais de toda a comunidade política, a “Comuna”. Seria esta a definir, em articulação com o congresso dos produtores, todos os aspectos relativos à legislação social, incluindo as prioridades sociais que regulam a produção e o investimento; e também a definição dos salários, que seriam decididos de acordo com as concepções dominantes de salário justo, assim como a definição de alguns preços, tais como o das matérias-primas e outros recursos naturais.

Alguns dos custos sociais resultantes destas negociações políticas (salários e o preço das matérias-primas, por exemplo) fariam então parte dos parâmetros “técnicos” que as associações de produtores teriam de ter em consideração aquando das decisões de produção, assinalando-as como “naturais” ou “técnicas”

nos seus balanços. As restantes decisões políticas, que constituem o perímetro social que pressiona directamente os meios e os fins técnicos da produção, gerando, por conseguinte, custos adicionais, deveriam, por uma questão de transparência e prestação de contas, ser considerados como tal nos balanços das unidades de produção afetadas. Isto significa que a tentativa de Maucourant (2005) de clarificar o modelo de Polanyi, segundo a qual os processos “naturais” são processos económicos que emergem espontaneamente no seio do quadro institucional criado pela comunidade política, só está correcta se o conceito de quadro institucional for alargado para incluir salários e muitos outros preços fixados, que aparecem como custos naturais ou técnicos, para recorrer à taxonomia deveras intrincada de Polanyi. No interior deste quadro institucional, os preços dos bens seriam o resultado das negociações entre as associações de produtores e as associações de consumidores.

Polanyi não estabelece uma articulação clara dos dispositivos institucionais através dos quais este processo de negociação aconteceria. Maucourant (2005) considera que o modelo de Polanyi, enquanto recusa da ideia de um grande mercado auto-regulado, deixa imenso espaço para verdadeiros ajustamentos de mercado, uma vez que as unidades económicas autogeridas descentralizadas, representadas pelas suas associações, têm de vender os bens que produzem aos consumidores ou às outras unidades económicas, através das associações que as representam. As associações têm de gerar excedentes acima dos custos técnicos de produção para poderem sustentar os custos sociais impostos pela comunidade e financiar novos investimentos. A reconstrução paciente, feita por Maucourant, do modelo algo vago de Polanyi sugere que existe um esforço para combinar o planeamento democrático com mercados reais. Os parâmetros, definidos pelos representantes dos consumidores e cidadãos na “Comuna” e pelo congresso que representa as associações de produtores, delimitariam o espaço onde as decisões de produção e as trocas de mercado poderiam ocorrer.

Na sua resposta a Polanyi, de 1923, Mises considera que o modelo de socialismo proposto, deixando um espaço ambíguo para as trocas de mercado, não poderá jamais ser estável. A principal razão indicada está em conformidade com o seu argumento global acerca da instabilidade política e institucional de qualquer economia que vá para lá das linhas de demarcação por si traçadas entre

o Estado e os mercados: a incoerência gerada pela existência de instituições económicas baseadas em princípios diferentes cria perturbações económicas, que podem levar tanto ao restabelecimento de uma economia baseada na propriedade privada como ao aumento do intervencionismo. Em última instância, a superação da propriedade privada inviabiliza qualquer tipo de transacção de mercado.

Todavia, o argumento de Mises contra Polanyi tem uma dimensão mais interessante: a acusação de que o sistema de Polanyi leva ao “sindicalismo”, se o congresso dos produtores se sobrepuser, na luta pelo poder, à “Comuna” (Mises, 1981). Para Mises, o sindicalismo é uma forma de capitalismo de mercado menos produtiva, onde são os próprios trabalhadores quem detém os meios de produção. Aqui, o argumento não se baseia tanto na impossibilidade de cálculo económico, mas sobretudo na ideia de que este sistema deixa de poder ser considerado socialista: se não existe nenhuma forma de planificação central, as desigualdades económicas fundamentais entre as diferentes associações em competição nos mercados acabariam, inevitavelmente, por surgir. E só poderiam ser precariamente contrariadas através de redistribuições periódicas de riqueza, que teriam efeitos negativos nos incentivos de mercado. Mais ainda, o rumo colectivo de cada unidade económica seria menos eficiente e eficaz do que um comando único por parte do empresário, o capitalista ou o senhorio, por causa da disseminação dos incentivos pecuniários.

Ou seja, a posição de Mises confirma a importância dada, as mais das vezes implicitamente, a asserções acerca das motivações humanas dominantes – o egoísmo – e a necessidade associada à existência de incentivos pecuniários selectivos, gerados pelos mercados, que são a força motriz para uma elite de empresários, os quais monopolizam a iniciativa, sendo os trabalhadores remetidos para funções passivas (Rodrigues, 2013). Esta questão das motivações é contraditória com a posição expressa, em 1922, de que o argumento contra o socialismo e a favor do capitalismo é neutro nesta matéria. O argumento usado contra o socialismo de Polanyi torna ainda mais claros três dos pilares do pensamento de Mises que foram integrados de alguma forma na economia política neoliberal posterior: a ideia de que a política se reduz a relações de poder, de força, sem espaço para deliberações ou compromissos estáveis entre interesses e valores; a ideia de que os mercados, baseados nos direitos de propriedade

capitalistas aos quais estão indissociavelmente ligados, são o melhor, na realidade o único, mecanismo para resolver apoliticamente os conflitos de interesses; e a subordinação do Estado, enquanto poder coercivo unitário, à missão de expandir e produzir os alicerces institucionais dos mercados para, potencialmente, a generalidade dos bens e serviços. Visando confrontar estes argumentos, Polanyi (1924/2008b) reforça a sua tese inicial – de que era mesmo necessária uma teoria positiva do socialismo, que podia ser assegurada pela sua tradição associativa, de forma a ultrapassar “as disputas escolásticas entre os marxistas ortodoxos e os seus inimigos burgueses” (Polanyi, 1924/2008b, p. 318) – e acrescenta-lhe mais duas linhas de argumentação, desenvolvidas também alhures (Polanyi 1925/2008c).

Em primeiro lugar, clarifica o paralelo entre a coexistência, em cada pessoa, como ser em várias relações sociais, de diferentes motivações – enquanto trabalhador, consumidor ou cidadão –, o que não impediria a consubstanciação “de um processo vital único, ou seja, a actividade económica individual”, e a representação institucional desta diversidade motivacional num modelo funcionalista de socialismo. Este processo seria constituído de forma tal que “o conflito entre os interesses idênticos de diferentes grupos de homens, como no caso de uma sociedade de classes sob a égide do capitalismo, é substituído pelo conflito entre interesses diferenciados de um e mesmo grupo de homens – conflito esse que se tornaria o princípio fundamental de acção na sociedade, e, por conseguinte, na economia” (Polanyi, 1924/2008b, p. 322). Polanyi (1924/2008b) pensa que o espaço político poderia ser estruturado de forma a substituir as relações de poder de dominação permanente – de “supremacia” – em que uma hierarquia acaba por levar a melhor, tal como Mises prevê, seja a Comuna ou o congresso dos produtores, por “relações de reconhecimento” legalmente respaldadas. Estas relações expressam a ideia de que é necessário, tanto ao nível individual como social, acomodar motivações e interesses divergentes: “cada indivíduo deve ser confrontado consigo mesmo, através das suas diferentes representações funcionais” (Polanyi, 1924/2008b, p. 323). O processo de decisão ao nível político, embora reconheça de forma transparente os interesses antagónicos, não impede a existência de equilíbrios consensuais, expressos na coexistência de diferentes disposições institucionais no seio da

economia, impulsionando uma ética individual de responsabilidade pelos resultados colectivos (Dale, 2010).

Em segundo lugar, Polanyi (1925/2008c), que, por esta altura, estava a tentar dissuadir os sociais-democratas austríacos de aderir a uma economia puramente administrada, com base no planeamento central, pensava que a descentralização necessária exigiria uma sociedade civil socialista forte e autónoma em relação ao Estado, assente em sindicatos independentes e numa série de outras associações que reflectissem as capacidades organizativas da classe trabalhadora – cooperativas, associações profissionais, municípios democráticos ou partidos operários (Chaloupek, 1990). Estes seriam parte de todo um esforço educacional, consciente, misturando teoria e prática, capaz de promover simultaneamente a autonomia individual e o sentido de simpatia e responsabilização ético-política pelo que acontece aos outros, dada a natureza das interdependências em sociedade. Tanto o planeamento central como as transacções de mercado eram consideradas, por si sós, incapazes de chegar a uma visão de conjunto da economia “a partir do seu interior”, sem estarem incrustadas numa rede de movimentos sociais organizados e articulados entre si (Polanyi, 1925/2008c; Mendell, 1990). Esta rede, valorizada por Mendell (1990), ou não fosse no período mais recente uma teorizadora da chamada economia social, era nos anos vinte vista por Polanyi como um elemento crucial da economia socialista, entendida como o resultado do processo de alastramento e aprofundamento da democracia. Polanyi (1925/2008c) reconhecia, mais uma vez, que não existia uma teoria económica própria para esta economia; e acrescentava até que talvez nunca viesse a haver: teorias parciais das instituições, que fornecessem instrumentos utilizáveis “na acção do movimento operário”, eram o melhor que os socialistas podiam obter.

Polanyi (1925/2008c) aprofundou a posição, relativamente moderada para a altura, que não se inscrevia na abordagem então dominante do socialismo (Chaloupek, 1990). Os argumentos que apresenta contêm o embrião da potencial resposta socialista ao futuro argumento de Hayek a favor dos mercados e em oposição à planificação central, como explanado na sequela anglo-saxónica do debate do cálculo socialista nos anos trinta e quarenta. Esta resposta aceita, em parte, os termos que enquadrarão o argumento epistémico de Hayek a favor do

mercado: os responsáveis pelo plano central podem apenas adquirir, recorrendo a sistemas de controlo traduzíveis em estatísticas, uma útil, mas estreita e limitada, “visão da economia a partir do exterior”, sobretudo acerca dos meios de produção disponíveis e das suas possibilidades técnicas passadas (Polanyi, 1925/2008c, pp. 330-331). Polanyi defendeu que, para lá dos meios de produção, a economia engloba também elementos qualitativos, que se esquivam a qualquer tentativa de quantificação ou medida, seja através dos mercados ou da planificação central (Congdon, 1990). Polanyi (1925/2008c) estava sobretudo preocupado com a identificação de mecanismos institucionais que permitissem revelar as necessidades humanas mais prementes e os custos sociais, de natureza necessariamente qualitativa, dos processos de produção, por exemplo em termos de condições de trabalho indignas. Se o “rumo ao socialismo é uma questão de organização”, então a dimensão moral da economia socialista, expressa na capacidade para chegar a uma “visão interna global da economia” exigia formas apropriadas de organização (Polanyi, 1925/2008c, p. 335). E as mais apropriadas de todas seriam democráticas, isto é, associações igualitárias, no espírito do socialismo de guildas. Estas proporcionariam um local de destaque para “ter voz” e educação, de forma a promover os sentimentos morais adequados entre os próprios trabalhadores e a criar mecanismos não-mercantis, capazes de produzir e espalhar fluxos de informação e conhecimento, para gerar as respostas políticas apropriadas por parte das lideranças políticas, até dadas as possibilidades que seriam abertas para se exercer pressão de baixo para cima.

Polanyi presumia que a evolução do princípio cooperativo de organização dos trabalhadores promovesse as motivações adequadas, para lá do egoísmo, entre os que geriam os bens de produção em nome da comunidade que os detinha formalmente. A existência de um conjunto de associações de consumidores e produtores, organizadas em federações nacionais, acabaria por, na expectativa de Polanyi, replicar a um nível mais vasto a atmosfera de fraternidade e interesse mútuo entre os trabalhadores que ele associava às experiências socialistas de sucesso como a “Viena Vermelha”.⁴

⁴ Polanyi (1944/2012) refere a experiência da capital austríaca no pós-guerra, gerida pelos sociais-democratas entre 1918 e 1934, como um enorme sucesso social e ético-político, provisionando bens essenciais, como a habitação, e assegurando a participação e dignificação das classes trabalhadoras. Apesar disso, ou talvez por causa disso, esta experiência foi “violentamente atacada pelos adeptos do liberalismo económico” (Polanyi, 1944/2012, p. 525), como Mises, mas só sucumbiu “ante o ataque de forças políticas poderosamente sustentadas por argumentos puramente económicos”. Esta formulação é exemplo de uma intuição significativa que

A existência de uma esfera de transacções de mercado, enquadrada por regras adequadas e por alguns preços, incluindo salários, definidos politicamente entre associações, tornaria o cálculo económico possível e indicaria os custos “naturais” e “sociais” da economia socialista, facilitando também a deliberação política na economia. Esta é, assim, uma súpula do projecto de Polanyi (1922/2008a, 1924/2008b, 1925/2008c) para uma “sociedade socialista organizada em termos funcionais”, um meio-termo entre a economia associativa sem mercados de Otto Neurath, assente na planificação integral da economia, e os mercados sem restrições de Mises, onde as relações de mercado são promovidas ao máximo e as únicas “associações” valorizadas são empresas capitalistas controladas pelos capitalistas.

É possível, entretanto, afirmar que o artigo de Polanyi de 1925 já contém o embrião do conceito de mercadorias fictícias, tão importante no argumento de *A Grande Transformação*, em 1944, quando parece abandonar a distinção entre custos sociais e naturais, tal como fora previamente definida, ainda que mantendo a intuição de que teriam de ser criados “órgãos de percepção” para garantir que os custos sobre terceiros gerados pelas actividades económicas seriam reconhecidos como tal e as instituições fundamentais do capitalismo reformadas – e, eventualmente, substituídas. Para fundamentar esta afirmação, é importante começar por assinalar que existe um importante elemento de continuidade no pensamento de Polanyi, desde a participação no debate do cálculo socialista até ao seu livro *A Grande Transformação*: a defesa de que os preços mais importantes – salários, taxas de juro, matérias-primas – têm sempre de ser definidos, directa ou indirectamente, fora dos mercados e através de um processo político. A diferença reside no facto de, nos primeiros anos da década de vinte, Polanyi estar convicto de que todos os custos relevantes do socialismo, em oposição ao capitalismo, podiam ser contabilizados através de um único critério – a hipótese da comensurabilidade, partilhada também por Mises, ainda que para justificar a racionalidade de um sistema socioeconómico radicalmente distinto. A partir de meados da década de vinte, Polanyi começou a dar nota do facto de que existem elementos que não podem ser reduzidos a um critério quantitativo

Polanyi desenvolveu em relação ao liberalismo económico (que considerava, erradamente, como tendo passado à história), aplicável à sua reconstrução como neoliberalismo: a combinação do poder de Estado e de teoria económica, com o objectivo de destruir as instituições não-mercantis igualitárias e os valores que as sustentam.

comum, pelo menos não sem perda de informação e conhecimentos fundamentais e vitais em termos de políticas públicas (Polanyi, 1925/2008c). Em *A Grande Transformação*, estes temas são tratados através da exploração do conceito de mercadorias fictícias, que pode ser interpretado como a principal resposta de Polanyi ao liberalismo económico mais intransigente, ou não fosse Mises, de novo, um dos seus alvos neste contexto.

Contra a ideia de que os mercados ditos livres tendem a ser a única maneira de induzir nas interações humanas comportamentos cooperativos eficientes, Polanyi denuncia a natureza utópica da engenharia política que procura transformar em mercadorias, submetidas a um sistema de mercado auto-regulado, com preços supostamente determinados pelas forças da oferta e da procura, três elementos que, na sua essência, não podem ser considerados mercadorias, já que não existem para ser comprados e vendidos: a moeda é um produto combinado da acção coerciva de um poder público e da confiança social; a terra é parte integrante da natureza; e o trabalho é “somente outro nome da actividade humana que acompanha a própria vida” (Polanyi, 1944/2012: 215). A tentativa de mercadorização integral destes elementos é tão artificial, tão contrária à “substância humana e natural” das sociedades, que não pode deixar de introduzir rupturas históricas profundas com outros sistemas de provisão, de exigir transformações institucionais de largo alcance e que passaram pela demolição sistemática de múltiplas instituições de protecção: o princípio prudente da “habitação”, saudavelmente desconfiado em relação ao alastramento dos mercados, foi substituído pela crença no progresso, na “beneficiação”, mesmo que este implicasse uma deterioração multidimensional das condições de vida, em particular das comunidades alvo de destruição, em última instância cultural (Polanyi, 1944/2012). Polanyi indica alguns dos mecanismos responsáveis por essa destruição, ao apontar de novo a inexistência de um órgão adequado no capitalismo liberal para dar conta dos custos sociais, invisíveis e ubíquos, gerados pelo esforço de criar uma sociedade de mercado: “as classes ligadas ao comércio não estavam dotadas de um órgão que as fizesse aperceber-se dos perigos acarretados pela exploração física dos trabalhadores, a destruição da vida familiar, a devastação dos meios circundantes, a ruína das florestas a poluição dos rios (...) bem como as inúmeras formas de vida privadas e públicas que não

tinham a ver com os lucros” (Polanyi, 1944/2012, p. 301).⁵ Estes custos irão progressivamente ser considerados incomensuráveis e, portanto, não abarcáveis pelos mecanismos de preços. A insustentabilidade das “ficções grosseiras”, associada ao imaginário mercantil de certas esferas da vida, é de tal ordem que o seu esforço de institucionalização tende a gerar um “contramovimento de protecção social”, plural do ponto de vista político e ideológico, muitas vezes conduzido por razões puramente pragmáticas, e plural do ponto de vista social, dada a multiplicidade de classes, suas fracções e alianças. O socialismo inscrever-se-ia na ala mais radical e consequentemente democrática deste contramovimento (Polanyi, 1944/2012).

3. O que é utópico e o que é realista na economia política?

Em contraste com as relações entre Mises e Polanyi, nunca houve qualquer debate intelectual explícito entre Polanyi e Hayek. Na obra de ambos existem apenas breves referências ao outro. Não obstante, podemos olhar para elas como sinais de uma sobreposição mais extensa de interesses e até de posições. Assim, Hayek menciona Polanyi somente duas vezes, ao longo da sua volumosa obra. A primeira referência é feita na edição inglesa de uma colecção de ensaios de críticos do socialismo em língua alemã (Hayek, 1935). Embora aponte os artigos de Polanyi nos anos vinte entre as mais importantes contribuições socialistas, e considere as experiências factuais e os contributos teóricos produzidos na Áustria dotados “de uma considerável força no âmbito da história intelectual do nosso tempo” (Hayek, 1935, p. 30), não há qualquer sinal, ao contrário do que sugere Congdon (1990), de que se tenha debatido a sério com os argumentos de Polanyi. Aliás, ele considerava que “os primeiros sistemas de um socialismo mais descentralizado, como o socialismo corporativo, ou sindicalismo, não devem fazer-nos perder tempo, uma vez que é agora bastante claro que não apresentam mecanismo algum para a direcção racional da actividade económica” (Hayek, 1935, p. 19). Mais de cinquenta anos depois, Hayek (1988) voltaria a mencionar Polanyi, apenas para criticar a sua visão de que a prosperidade de Atenas na

⁵ K. William Kapp, o economista institucionalista que sistematizou a noção de custo social, enquanto alternativa ao conceito mais circunscrito e convencional de externalidade negativa, considerou que *A Grande Transformação* indica “o muito que pode ser alcançado quando a história do capitalismo é reescrita tendo em vista tornar visível o fenómeno dos custos sociais” (Kapp, 1950, p. 45).

Antiguidade pudesse ter alguma coisa a ver com o modo particular como as autoridades públicas enquadraram e controlaram as actividades de trocas de mercado. Apesar da distância temporal considerável e dos diferentes contextos da discussão, existem elementos de continuidade, que dizem muito acerca da principal preocupação de ambos: a relação entre autoridade política e mercados, bem como o papel e a essência dos preços.

Por sua vez, Polanyi nomeia Hayek também duas vezes: numa crítica ao que ele via como o “preconceito económico” do argumento sobre a perversidade política do “intervencionismo” económico na obra *O Caminho para a Servidão* (Polanyi, 1947), e na crítica à edição em inglês, da responsabilidade de Hayek, da primeira edição dos *Princípios de Economia Política* de Carl Menger, em vez da amplamente revista segunda edição (Polanyi, 1977). Segundo Polanyi, esta última continha a distinção entre os significados formal e substantivo do termo económico, que ele desenvolveu, de acordo com a tradição aristotélica (Polanyi, 1977). Estas críticas merecem ser analisadas com mais atenção, pois contêm elementos que podem ilustrar a análise comparativa.

De facto, Polanyi (1947) posiciona os argumentos de Hayek (1944/2009) numa tradição liberal anacrónica, atolada no determinismo económico, que ele considerava ser o cunho intelectual distintivo da sociedade do século XIX, dominada por um sistema disfuncional de mercados. Esta sociedade estava em processo de transformação, por via do reconhecimento político e moral da natureza ficcional das mercadorias mais cruciais. O determinismo económico liberal manifestava-se num cepticismo profundo em relação ao futuro da liberdade nesta conjuntura histórica: “há quem sustente, como Hayek, que, uma vez que as instituições livres foram um produto da economia de mercado, devem dar lugar à escravidão quando essa economia desaparecer (...) Não é lá muito lógico atribuir os efeitos de uma necessidade económica à força de um motivo económico que depende da sua existência” (Polanyi, 1947, p. 117). No entanto, Polanyi reconhecia um aspecto moral importante na economia política de Hayek: que as denominadas “liberdades prezadas” – liberdade de consciência, liberdade de reunião, liberdade de associação, liberdade de escolher o emprego – eram um subproduto não intencional do domínio do sistema de mercados e da hegemonia da prática liberal de pensar as esferas política e económica como estando

separadas, por forma a subordinar a primeira às exigências da segunda (Polanyi, 1947, p. 117). A diferença, para Polanyi, é que estas “liberdades prezadas” estavam acompanhadas de “liberdades perversas”, por exemplo as ligadas ao poder dos empresários para transferir os custos sociais para os trabalhadores e para o conjunto da comunidade. Baseando-se na experiência da economia de guerra britânica, onde planeamento, liberdades civis e uma diminuição da desigualdade económica estiveram em sintonia, Polanyi declarou que uma reconfiguração institucional da economia talvez fosse capaz de criar uma separação entre as liberdades que os cidadãos têm razões para valorizar e as liberdades que têm razões para detestar.⁶ Uma vez que, como defendeu extensivamente em *A Grande Transformação*, os moldes das sociedades humanas não poderão jamais ser determinados apenas pelo que acontece na esfera económica, Polanyi concluiu que “as garantias institucionais de liberdade pessoal são compatíveis com qualquer sistema económico” e que as “liberdades perversas” podiam ser distinguidas das liberdades prezadas, e progressivamente eliminadas pela extensão destas ao reino da “indústria” (Polanyi, 1947, p. 117). É a pertinência desta distinção e o interesse desta expansão que, entre outras coisas, Hayek (1944/2009) nega.⁷ No entanto, Hayek (1944/2009) reconheceu que toda a economia é estruturada por regras, e nesse sentido regulada, procurando distanciar-se do *laissez-faire*. E como defendeu Polanyi (1944/2012, p. 466), toda a regulação “expande e ao mesmo tempo restringe a liberdade”, sendo que o que importa, em termos de economia política, “é a comparação entre as liberdades perdidas e as conquistadas”.

A outra referência de Polanyi a Hayek acontece no contexto da tentativa de resgatar do esquecimento intelectual a segunda edição, póstuma, dos *Princípios* de Menger, pela qual Hayek seria parcialmente responsável. O livro – *The Livelihood of Man* – onde surgem os comentários críticos de Polanyi foi publicado depois da sua morte, dando mostras do seu interesse constante pelos

⁶ Não deixa de ser interessante assinalar que Polanyi (1947) refira a necessidade de uma intervenção planeada e democrática por parte de consumidores e produtores para resolver o “problema da indústria” e para transcender “a mentalidade obsoleta de mercado”, fazendo deste modo eco, no contexto do pós-guerra e com menos pormenores institucionais, da posição original aquando do debate sobre o cálculo em socialismo.

⁷ Fá-lo através de um argumento que rejeita, abertamente, qualquer modelo de determinismo histórico, em particular a tese de que os desenvolvimentos tecnológicos estavam a levar os sistemas socioeconómicos para direções específicas, nomeadamente na direção do planeamento central. Ao mesmo tempo, afirma a primazia das ideias (Hayek, 1944/2009). Porém, Polanyi, que até podia estar de acordo com a posição de Hayek, tinha em mente uma forma mais específica de determinismo: a associação necessária, estabelecida por Hayek, entre a sociedade de mercado, por um lado, e a liberdade, por outro. Para Polanyi, esta relação era muito mais contingente e problemática.

contributos dos economistas na tradição da Escola Austríaca. Polanyi (1977) afiança que Menger formulou uma distinção entre dois entendimentos do económico e tentou, a partir dos anos quarenta, desenvolver a opinião do autor, a cujas implicações, tanto teóricas como políticas, atribui absoluta importância. O significado formal do económico sublinhava a importância do chamado comportamento económico racional, das escolhas racionais entre meios escassos para satisfazer preferências não-controladas, dando assim origem a uma visão a-histórica do indivíduo isolado, submerso num mercado altamente idealizado. Este significado cativou a imaginação social e foi posteriormente aplicado para avaliar toda a espécie de interações sociais, independentemente do tempo ou do espaço. A abordagem substantiva, por outro lado, centra-se no lado material da economia, nos “processos institucionalizados através dos quais se garante o sustento”, isto é, na combinação específica, em termos históricos e espaciais, das formas de integração – redistribuição, reciprocidade, economia doméstica e trocas de mercado – que a estruturam (Polanyi, 1977, p.20). Dedicar-se também à variedade de motivações humanas presentes no âmago das expressões institucionais concretas destas formas e ao tipo de interações entre as pessoas, e entre pessoas e a natureza, que são promovidos.

A dicotomia de Polanyi pode ser comparada com a (e em parte contraposta à) distinção que Hayek e de certo modo Mises fazem entre a *cataláxia*, enquanto ordem de mercado aberta e espontânea, e a economia, enquanto organização deliberada e fechada. Hayek (1982) considera que o termo “economia” é erradamente usado para descrever as sociedades de mercado, porque essas sociedades são, ao invés, compostas por miríades de economias, seja em modelo familiar, empresarial ou governamental. Nas empresas predominam a cooperação e o comando explícitos, baseados em fins específicos, assumidos e mais ou menos partilhados. O mesmo se passa nessa “peça da máquina utilitária” que é o Estado (Hayek, 1944/2009, p. 80). Os fins variam imenso ao nível social e só podem ser coordenados e tornados de algum modo compatíveis entre si pelo nexos monetário de mercados enquadrados por regras apropriadas (Hayek, 1982). Estamos perante a ordem económica aberta que Hayek contrapõe à ordem económica fechada, concebida à imagem da “concepção aristotélica da economia como economia doméstica”, ou seja, uma ordem geral onde as instituições económicas são pensadas e defendidas de acordo com os fins éticos que

promovem e as necessidades e capacidades humanas que favorecem (O'Neill, 1995, p. 426). O socialismo, a expansão do Estado social, ou formas de planeamento que interfiram com a formação de preços no âmbito de países capitalistas seriam, segundo Hayek, as expressões institucionais contemporâneas da hegemonia da ideia de que uma ordem complexa e em evolução podia ser pensada como uma economia.

Uma vertente importante e bastante estudada do pensamento de Hayek é, precisamente, marcada pela tentativa de construir uma narrativa histórica, em que a oposição entre a *cataláxia* e a economia seja determinante e que acompanhe a oposição, na evolução e selecção de regras, entre ordens espontâneas e construídas. A esta dicotomia institucional é necessário acrescentar uma outra, que ocorre ao nível fundamental das ideias: a oposição entre o construtivismo e aquilo a que Hayek chama várias coisas, desde abordagem evolucionista, racionalismo crítico, ou simplesmente tradição liberal clássica (Hayek, 1960/2006). Hayek defende em várias obras que o construtivismo enfatiza a capacidade da razão humana para configurar e reconfigurar as instituições da economia e controlar a sua evolução, apresentando assim uma visão a-histórica da razão humana. Esta abordagem ignoraria que, com todas as suas limitações, a razão é o produto de um longo processo de evolução cultural, em que os grupos mais bem-sucedidos são os que “tropeçaram” em determinadas configurações institucionais, que conduziram à ordem de mercado, e conseguiram preservá-las e melhorá-las (Hayek, 1944/2009; 1960/2006). A tradição evolucionista que Hayek invoca e reinventa retira várias implicações epistemológicas e ideológicas das “ideias gémeas” de ordem espontânea e de evolução cultural: os racionalistas construtivistas, sobretudo os de persuasão socialista, são incapazes de perceber a natureza da ordem que almejam destruir ou o papel que as regras, consoante vão evoluindo, desempenham em moldar e auxiliar os indivíduos, que detêm um conhecimento irremediavelmente limitado de como se devem comportar o mais racionalmente possível. Hayek, numa tentativa interessante para inverter os termos da discussão, considera o socialismo como sendo, na essência, e apesar da retórica racionalista dos seus proponentes, nada mais do que um anseio nostálgico por um estado de coisas primitivo, em que predominam as economias, em vez das mais avançadas e progressistas *cataláxias* (Hayek, 1982).

Existem duas questões que, apesar de óbvias, têm de ser levantadas neste contexto: em primeiro lugar, poderá Polanyi ser considerado um racionalista construtivista, de persuasão socialista, se aceitarmos os termos da taxonomia de Hayek? Em segundo lugar, será que o pensamento neoliberal, em geral, e em particular o de Hayek, não são vítimas da dicotomia estrita que ele elabora, na medida em que há, como assinalaremos com a ajuda de Polanyi, uma clara dimensão construtivista, assumidamente utópica, no seu seio?

À primeira vista, Polanyi parece incorporar todas as armadilhas do construtivismo racionalista e socialista que Hayek criticou ao longo da sua vida. Aliás, segundo os parâmetros de Hayek, podemos afirmar que Polanyi oscilou entre a tentativa de construir modelos para a sociedade socialista do futuro, nos anos vinte, e uma ânsia idealista e “atávica” por um passado pré-capitalista, dos anos quarenta em diante, em particular nos seus estudos de antropologia histórica.⁸ Poderíamos afirmar que o primeiro Polanyi expressa a “arrogância fatal” intrínseca ao socialismo, ou seja, a pretensão, por parte dos intelectuais, de que possuem o conhecimento suficiente para remodelar o complexo enquadramento institucional da sociedade, de forma a deixá-la de acordo com os seus ideais preconcebidos (Hayek, 1988); e é outrossim possível afirmar que o último expressa a relutância em aceitar a ideia de que a expansão das forças de mercado é uma força motriz na base do crescimento da civilização. Aquilo que Polanyi interpretou como o emergir de uma economia capitalista, minando os laços comunitários de confiança e reciprocidade, e substituindo-os por uma cultura míope e autodestrutiva de ganância, era, segundo Hayek (1982), nada mais do que a difícil e fundamental repressão dos instintos associados às sociedades primitivas de pequena escala, onde os seres humanos começaram a interagir uns com os outros. Esta repressão devia-se sobretudo à expansão do nexos monetário, ligada ao crescimento de uma cultura individualista, e com a transformação da essência das relações sociais. Mais ainda, a análise de Polanyi (1944/2012) acerca da necessária diversidade institucional que florescia no seio do capitalismo e o seu prognóstico de um modelo exequível de socialismo resumem, dirão alguns, a inconsistência intelectual do “pântano do meio” que Hayek (1944/2009) criticou de forma tão incisiva.

⁸ Para estes estudos, ver, por exemplo, Polanyi (1957).

Na realidade, o esboço de modelo para uma economia socialista de Polanyi surgiu em resposta ao desafio de Mises; e nunca houve a intenção de que fosse mais do que um rascunho para justificar e imaginar os padrões gerais de uma transição no contexto duma crise sistémica do capitalismo. Não era muito diferente, em termos da atitude intelectual perante as questões sociais e do conhecimento necessário para intervir nelas, dos momentos de engenharia política deliberada de Hayek, como por exemplo quando tentou “construir um equipamento intelectual de emergência”, capaz de redesenhar as instituições políticas dos países capitalistas, de forma a enfrentar o que considerava serem os problemas da “democracia ilimitada” e a atrofia dos mercados associada (Hayek, 1982, p. 152). Aliás, já nos anos trinta, Hayek (1939/1948b) tinha apresentado um detalhado modelo de organização federal que gerasse enviesamentos institucionais favoráveis à expansão dos mercados.

Pode dizer-se que Polanyi convergiu ainda mais com Hayek, embora chegando a conclusões morais e políticas opostas, em três áreas onde o conhecimento humano e respectivas limitações se manifestam com uma acuidade particular: a identificação de mecanismos e padrões que dêem conta do impacto das instituições nas motivações e carácter humanos; o papel de instituições específicas na identificação e mobilização de certas formas de conhecimento acerca das necessidades humanas; e a importância de avanços espontâneos, imprevisíveis e involuntários nas questões sociais.

Polanyi considerava, tal como Hayek, que era possível identificar os padrões globais e os mecanismos que os explicam, através dos quais instituições e ideias específicas podem influir nas motivações e interesses dos indivíduos. E, na verdade, a sua abordagem, embora de tipo socialista, converge com a do neoliberal Hayek, no sentido em que há em ambos um esforço para ter como ponto de partida o postulado realista de “homens cuja essência e carácter são determinados pela existência em sociedade” (Hayek, 1946/1948e: 6). Para Polanyi (1944/2012), isto implica igualmente levar até às últimas consequências a “realidade da sociedade”, com os seus mecanismos, tanto privados como públicos, de poder, coerção e persuasão, denunciando a divisão liberal entre a esfera da política e da ideologia e a esfera da economia. Esta separação servia apenas para aprofundar a hegemonia de uma visão altamente problemática das

motivações humanas, que as reduz ao egoísmo na esfera económica, incentivado por mercados em constante expansão – a acumulação de bens, sobretudo dinheiro, torna-se perversamente um fim em si mesmo para os indivíduos. Tanto ao nível do discurso, como ao da promoção política de um grau mais elevado de uniformidade institucional na economia, o liberalismo económico é acusado de ter um impacto corrosivo nas bases morais partilhadas em que qualquer sistema socioeconómico decente tem, necessariamente, de assentar (Polanyi, 1944/2012; Rodrigues, 2004). Uma expressão deste carácter corrosivo é a incapacidade dos indivíduos, em particular os que controlam as empresas capitalistas, de reconhecerem a responsabilidade moral e política pelos custos sociais que as suas actividades geram. Na ausência de regulação adequada e de freios e contrapesos sociais organizados, estes custos serão transferidos para os trabalhadores sob a forma de sofrimento humano, tanto no desemprego como no trabalho; e para toda a comunidade, através da criação de obstáculos à possibilidade de assegurar aquilo que Polanyi designou como o princípio da “habitação”, por causa da destruição do ambiente e da degradação dos espaços públicos na base da existência das comunidades.

Existem elementos suficientes para concluir que Polanyi retém uma importante conclusão do postulado da dependência institucional da sociabilidade humana, nomeadamente que não existe posição, instrumento ou forma de integração privilegiada, através da qual a totalidade da realidade social possa ser apreendida por um indivíduo ou grupo de indivíduos. Neste ponto, Polanyi não está muito distante de Hayek. Aliás, Polanyi (1925/2008c) recusa, muito claramente, o projecto de uma economia organizada exclusivamente em torno da planificação central e da sua dependência da recolha e agregação de informação quantitativa sobre a economia, em particular com recurso à estatística – era suposto o plano ser capaz de fornecer, quando muito, uma visão quantitativa a partir do exterior do processo de provisão, mas não conseguiria facultar uma visão qualitativa do interior. Isto, sublinhemo-lo agora, antes de Hayek ter criticado a pretensão de conhecimento por parte dos socialistas. Infelizmente, Hayek ignorou, de forma deliberada, os contributos dos socialistas, como Polanyi (1944/2012), que acabaram por aceitar a presença de mercados e o papel dos preços como instrumentos de informação para determinar os custos de bens e serviços numa sociedade complexa, mas que consideravam que esta instituição,

sobretudo na configuração idealizada pelos liberais, de um mecanismo espontâneo de procura e oferta, só era sustentável ou adequada para algumas áreas da vida social.⁹

O conceito de mercadoria fictícia desenvolvido por Polanyi (1944/2012) tem como objectivo, entre outras coisas, tornar clara a necessidade de instituições de controlo externas aos mercados, no Estado e na sociedade civil – como sindicatos, órgãos de regulação e planeamento e mecanismos de participação e deliberação, todos presentes no processo de provisão. Estes podiam ser veículos para a expressão de importantes preocupações de cariz qualitativo e para a reconfiguração política das instituições de mercado e sua substituição localizada por outros padrões de integração – como a domesticidade, a reciprocidade ou a redistribuição na provisão de bens. Polanyi acreditava, ao contrário de Hayek, que os mercados só seriam viáveis numa economia genuinamente mista, que considerava muito mais sustentável do que qualquer apologia utópica de uma *cataláxia* inescrutável.

Na verdade, Polanyi (1944/2012) inverteu a interpretação global da história de Hayek. Para simplificar, de algum modo, podemos afirmar que, para Hayek, os alegados mercados livres são sobretudo desenvolvimentos espontâneos, com que a humanidade se deparou, e que teria de preservar, aprimorar e expandir. A economia mista, por outro lado, é uma “mixórdia” insustentável, produto da mentalidade construtivista empenhada em reconfigurar a ordem social de forma a aproximá-la do projecto socialista pré-definido. Para Polanyi, é precisamente o contrário: o mercado livre é uma construção ideológica, utópica, que oculta o facto de os verdadeiros mercados serem artefactos maleáveis de poder estatal mais ou menos centralizado, que se expandiram numa determinada direcção no século XIX pela mão da ideologia intransigente do liberalismo económico. O contramovimento de protecção social, que deu a origem a uma série de economias mistas, que poderiam, eventualmente, ser a base de futuras sociedades socialistas, surgiu enquanto resposta política realista, espontânea e não planeada, à devastação causada pelas crises recorrentes das sociedades de mercado.

⁹ De resto, não foi só Hayek a ignorar a presença, ainda que subordinada dos mercados, no pensamento de Polanyi. Mesmo o seu biógrafo afirma que Polanyi defendeu uma “utopia anti-mercantil” (Dale, 2016: 326). Esta fórmula é duplamente infeliz: o termo utopia é sempre usado com uma conotação negativa em Polanyi, sinónimo de distopia liberal, por um lado, e os mercados são retidos nos exercícios prospectivos de Polanyi (1944/2012), por outro lado.

Embora esta resposta tenha sido inicialmente guiada por preocupações pragmáticas, acabou por tornar-se num movimento orientado por um conjunto diverso de ideologias.

Aquilo que Hayek via como “o caminho para a servidão” foi interpretado por Polanyi como a descoberta de modos de incrustar democraticamente a economia. Tal só poderia ser conseguido se houvesse o reconhecimento intelectual de que a “economia humana” não tem outro remédio senão ser “incrustada e emaranhada em instituições, tanto económicas como não económicas” (Polanyi, 1957, p. 250). A partir daqui, Polanyi conclui que o projecto liberal de reduzir a política e a moral a apêndices instrumentais do sistema de mercado, bem patente na ideia de que “em vez de existir uma economia incrustada nas relações sociais, são as relações sociais que estão incrustadas no sistema económico”, é um empreendimento utópico que corrói os requisitos socioeconómicos, políticos e morais de uma sociedade viável (Polanyi, 1944/2012, p. 194). Os objectivos do liberalismo económico podem ser inatingíveis, em última instância, mas as consequências de tentar expandir os mercados a um número cada vez maior de esferas da vida social seriam tremendas e negativas: uma utopia com consequências distópicas, em suma.

Tanto Polanyi como Hayek tinham consciência da dimensão das implicações políticas de escolher entre duas interpretações opostas da evolução histórica das economias políticas. Existe uma ligação entre a narrativa liberal da simplicidade e espontaneidade dos mercados livres e a ideia de que o socialismo não passa de engenharia política radical, um simples produto de evoluções ao nível ideológico. Polanyi estava apostado em destruir esta narrativa. É de notar que Polanyi (1944/2012), neste contexto, não avança com nenhum modelo detalhado, que explicita o modo como o liberalismo económico seria substituído pelo socialismo. Em parte, tal deveu-se ao facto de ele considerar na década de quarenta que o socialismo seria uma tendência favorecida pelo funcionamento não planeado das instituições democráticas – “o socialismo é, essencialmente, a tendência imanente a uma civilização industrial para transcender o mercado autorregulado, subordinando-o conscientemente a uma sociedade democrática” (Polanyi, 1944/2012, p. 440). Os mercados seriam, neste contexto, uma instituição economicamente útil, mas politicamente subordinada. Tendo em

mente a trajectória da sua investigação a partir de *A Grande Transformação*, Polanyi também se convenceu de que os modelos detalhados eram inúteis, perante a complexidade, mas também a diversidade, que as transformações institucionais poderiam adquirir, sobretudo a partir do momento em que os constrangimentos de um sistema-mundo capitalista integrado se estavam a atenuar por via de processos mais ou menos espontâneos de desglobalização. Em vez de dar detalhes institucionais, optou por olhar para as áreas onde a desmercadorização, indissociável da democratização com escala sobretudo nacional, teria de acontecer – trabalho, natureza e dinheiro – e para a defesa genérica de uma visão realista e unificada das motivações complexas dos indivíduos, indo bem para lá do egoísmo, que mimetizaria a nova articulação entre política e economia. Nestas, haveria sempre espaço, ainda que mais reduzido, para os incentivos pecuniários gerados nos mercados, e para promover a imaginação política que está por trás do surgimento de espaços cooperativos. Deste modo, fortalecer-se-ia a ideia de que a sociedade pode ser “um conjunto de relações propriamente humanas entre pessoas” (Polanyi, 1944/2012, p. 440).¹⁰

Ademais, Polanyi desejava igualmente conceber uma plataforma conceptual que pudesse orientar a busca pelas formas globais de integração identificáveis nas “economias empíricas” do passado, presente e futuro (Polanyi, 1957). Contrariamente às premissas liberais acerca da simplicidade das sociedades de pequena escala do passado pré-capitalista, Polanyi tentou demonstrar, na sua pesquisa histórica e antropológica, que todas as economias humanas tendem a ser caracterizadas por padrões de interacção social complexos. O processo de provisão estava, assim, muito longe da simplicidade das premissas associadas à ideia de uma *cataláxia*, que tendiam a olhar para o passado em termos de evolução conjunta, mais ou menos progressiva, de dinheiro, mercados e trocas, e para o presente como nada mais do que uma tentativa equivocada de destruição desta evolução (Polanyi, 1977). Embora não tenham sido concebidas directamente para responder às críticas de Hayek, a análise e categorias de Polanyi podem ser assim facilmente reconstruídas com esse objectivo em mente.

¹⁰ Esta ideia reflecte o horizonte cristão que influenciava Polanyi. Nos anos trinta, defendeu que cristãos e socialistas partilhavam uma mesma preocupação, ambos em conflito com o fascismo e o liberalismo de Mises (este acusado de tolerar o primeiro), nomeadamente quando insistiam no seguinte: “A descoberta do indivíduo é a descoberta da humanidade. A descoberta da humanidade é a descoberta da comunidade. A descoberta da igualdade é a descoberta da sociedade. Cada um está implicado no outro. A descoberta da pessoa é a descoberta de que a sociedade é uma relação de pessoas” (Polanyi, 1935, p. 370).

4. Conclusão

É interessante notar, em jeito de conclusão, que uma das ideias mais fecundas de Polanyi (1944/2012), usada hoje em dia para analisar criticamente a economia política do neoliberalismo, revelou ser a de que o próprio liberalismo económico requereu sempre um activismo político estatal, nacional e imperial, sem precedentes para que a sua precária institucionalização fosse possível. Haveria um desfasamento entre a teoria e a prática políticas.

O neoliberalismo, enquanto reinvenção do liberalismo económico depois da sua queda no período entre guerras, cujo poder e alcance Polanyi claramente subestimou, seria também, na prática, um construtivismo político disfarçado por uma teoria mais ou menos naturalista sobre o mercado. Apesar dessa subestimação, Polanyi legou-nos um arsenal teórico que nos permite compreender as razões para esse construtivismo. Assim sendo, a retórica sobre as ordens espontâneas, analisada num quadro Polanyiano, faria parte de uma “dupla verdade” com intuítos de obnubilação ideológica e propagandística da parte de quem, como Hayek, sempre valorizou o papel das ideias e instituições políticas, e do seu controlo, na da evolução institucional da economia, valorização aliás partilhada com Mises (Mirowski, 2009).

Creio, no entanto, que uma leitura da tradição neoliberal revela como o desfasamento enunciado entre teoria e prática é bem menor do que se julga, o que não quer dizer que não existam aí tensões teóricas relevantes (Rodrigues, 2018). Sendo a questão do Estado e do seu controlo elitista decisiva, sobretudo tendo em conta a resiliência das tendências ditas colectivistas no século XX, a criação de estruturas económicas e políticas de constrangimento, que limitem a democracia, torna-se crucial, dado o diagnóstico: “duvido que um mercado funcional tenha alguma vez surgido no quadro de uma democracia ilimitada e parece provável que uma tal democracia o destruiria se por acaso aí tivesse surgido” (Hayek, 1982, p. 77). Dado que Polanyi nos indica como o “*laissez-faire* fora planeado” ou como “a via do mercado livre foi aberta e manteve-se aberta graças ao incremento do intervencionismo organizado e controlado centralmente” (Polanyi, 1944/2012,

pp. 310-311), tal desconfiança em relação à democracia de alta intensidade não se deve a uma desconfiança em relação ao Estado, bem pelo contrário.

Para os neoliberais é claro que a expansão nacional e interligação global dos mercados exigiram sempre uma expansão da acção multi-escalar dos poderes públicos – sobretudo dos países do centro –, o que seria aliás facilitado por uma condição política: inexistência ou forte limitação da democracia, até porque a condução política é confiada a instituições tanto quanto possível blindadas face às pressões sociais das maiorias nos vários espaços da sua operação, idealmente de natureza supranacional.

Ao definir o socialismo como processo de democratização da economia, as hipóteses de Polanyi contrastam com as que são propagadas pelos neoliberais. Ao mesmo tempo, as suas propostas socialistas foram moldadas e temperadas pelo envolvimento com os poderosos argumentos na origem da renovação do liberalismo económico, culminando na ideia da economia mista com ampla variação institucional, permitida pela redução das pressões globalizadoras. É caso para dizer que o debate continua na actualidade; e que alguns dos seus termos continuam a ser bem enquadrados pelos autores aqui apresentados.

Bibliografia

Block, W. (1996). Hayek's Road to Serfdom. *Journal of Libertarian Studies*, 12: 2, 339-36.

Boetke, P. (2002). *Calculation and Coordination – Essays on Socialism and Transitional Political Economy*. Londres: Routledge.

Caldwell, B. (2004). *Hayek's Challenge – An Intellectual Biography of F. A. Hayek*. Chicago: Chicago University Press.

Chaloupek, G. (1990). The Austrian debate on economic calculation in a socialist economy. *History of Political Economy*, 22: 4, 659-675.

Cole, G. D. H. (1920). *Guild Socialism Restated*. Londres: Leonard Parsons.

Congdon, L. (1990), "The Sovereignty of Society", in K. Polanyi-Levitt (ed.), *The Life and Work of Karl Polanyi*, Montreal: Black Roses Books.

Dale, G. (2010). *Karl Polanyi – The Limits of the Market*. Cambridge: Polity Press.

Dale, G. (2016). *Karl Polanyi – A Life on the Left*. Nova Iorque: Columbia University Press.

Hayek, F. (1935). The Nature and History of the Problem. In F. Hayek (Ed.), *Collectivist Economic Planning* (pp. 1-40). Londres: Routledge and Kegan Paul.

Hayek, F. (1948a). Economics and Knowledge (pp. 33-56). *Individualism and Economic Order*. Chicago: University of Chicago Press (publicado originalmente em 1937).

Hayek, F. (1948b). The Economic Conditions of Interstate Federalism. In *Individualism and Economic Order* (pp. 255-272). Chicago: University of Chicago Press, 1939.

Hayek, F. (1948c). The Use of Knowledge in Society. In *Individualism and Economic Order* (pp. 77-93. Chicago: University of Chicago Press. (Publicado originalmente em 1945).

Hayek, F. (1948d). The Meaning of Competition. *Individualism and Economic Order* (pp. 92-106). Chicago: University of Chicago Press. (Publicado originalmente em 1946).

Hayek, F. (1948e). Individualism: True or False. In *Individualism and Economic Order* (pp. 1-32). Chicago: University of Chicago Press. (Publicado originalmente em 1946).

Hayek, F. (1982). *Law, Legislation and Liberty*. Londres, UK: Routledge.

Hayek, F. (1988). *The Fatal Conceit: The Errors of Socialism*. Chicago, IL: University of Chicago Press.

Hayek, F. (2006). *The Constitution of Liberty*. Londres: Routledge. (Publicado originalmente em 1960).

Hayek, F. (2009). *O Caminho para a Servidão*. Lisboa: Edições 70. (Publicado originalmente em 1944)

Heilbroner, R. (1989). The Triumph of Capitalism. *The New Yorker*, 23 de Janeiro.

Hodgson, G. (1999). *Economics and Utopia*. Londres, UK: Routledge.

Ikedá, S. (1997). *Dynamics of the Mixed Economy – Toward a theory of interventionism*. Londres: Routledge.

Kapp, K. W. (1950). *The Social Cost of Private Enterprise*. Nova Iorque: The Schocken Books, 1971.

Keynes, J. M. (1944). Letter to Hayek. In D. Moggridge (Ed.), *The Collected Writings of John Maynard Keynes*, Volume 27 (pp. 385-388). Londres: MacMillan.

Lange, O. (1936). On the Economic Theory of Socialism. In B. Lippincott (Ed.), *On the Economic Theory of Socialism*. Minnesota: University of Minnesota Press.

Maucourant, J. (2005). *Avez-vous lu Polanyi?*. Paris : La Dispute.

Mendell, M. (1990). Karl Polanyi and Feasible Socialism. In K. Polanyi-Levitt (Ed.), *The Life and Work of Karl Polanyi* (pp. 66-77). Montreal: Black Roses Books.

Mirowski, P. (2009). Postface: Defining Neoliberalism. In P. Mirowski & D. Plehwe (Eds.), *The Road from Mont Pelerin – The Making of the Neoliberal Thought Collective* (pp. 417-455). Cambridge, MA: Harvard University Press.

Mises, L. von (1920). Economic Calculation in the Socialist Commonwealth. In F. Hayek (Ed.), *Collectivist Economic Planning* (pp. 87-130). New York: Augustus M. Kelley, 1935.

Mises, L. von (1981). *Socialism*. Indianapolis: Liberty Fund.

Mises, L. von (1927). *Liberalism*. San Francisco, CA: Cobden Press. (Publicado originalmente em 1985).

Mises, L. von (1944). *Bureaucracy*. Cedar Falls: Center for Futures Education. (Publicado originalmente em 1983).

Mises, L. von (1949). *Human Action*. London: William Hodge.

O'Neill, J. (1995). Polity, Economy, Neutrality. *Political Studies*, 33, 414-431.

O'Neill, J. (1998). *The Market – Ethics, Knowledge and Politics*. Londres: Routledge.

Polanyi, K. (1935). The Essence of Fascism. In J. Lewis, K. Polanyi e D. Kitchin (Eds.), *Christianity and Social Revolution*. Londres : Victor Gollancz Ltd.

Polanyi, K. (1947). Our Obsolete Market Mentality – Civilization Must Find a New Thought Pattern. *Commentary*, 3, 109-117.

Polanyi, K. (1957). The Economy as an Instituted Process. In K. Polanyi, C. M. Arensberg and H. W. Pearson (Eds.), *Trade and Market in the Early Empires—Economies in History and Theory*. Glencoe, Illinois: The Free Press.

Polanyi, K. (1977). *The livelihood of man*. Nova Iorque: Academic Press.

Polanyi, K. (2008a). La comptabilité socialiste. In *Essais de Karl Polanyi* (pp. 283-316). Paris: Seuil (publicado originalmente em 1922).

Polanyi, K. (2008b). La théorie fonctionnelle de la société et le problème de la comptabilité socialiste. In *Essais de Karl Polanyi* (pp. 317-326). Paris: Seuil (publicado originalmente em 1924).

Polanyi, K. (2008c). Nouvelles considérations sur notre théorie et notre pratique. In *Essais de Karl Polanyi* (pp. 327-336), Paris: Seuil (publicado originalmente em 1925).

Polanyi, K. (2012). *A Grande Transformação*. Lisboa: Edições 70. (Publicado originalmente em 1944).

Rodrigues, J. (2004). Endogenous Preferences and Embeddedness. *Journal of Economic Issues*, 30 : 1, 189-200.

Rodrigues, J. (2012). Where to Draw the Line between the State and Markets? Institutional Elements in Hayek's Neoliberal Political Economy. *Journal of Economic Issues*, 66: 4, 1007-1033.

Rodrigues, J. (2013). The Political and Moral Economies of Neoliberalism: Mises and Hayek. *Cambridge Journal of Economics*, 37: 1001-1017.

Rodrigues, J. (2018). Embedding neoliberalism: The theoretical practices of Hayek and Friedman. In D. Cahill, M. Konings, M. Cooper and D. Primrose (Eds.), *SAGE Handbook of Neoliberalism*. Londres: Sage, no prelo.

Rosner, P. (1990). Karl Polanyi on Socialist Accounting. In K. Polanyi-Levitt (Ed.), *The Life and Work of Karl Polanyi* (pp. 55-65). Montreal: Black Roses Books.

Schumpeter; J. (1976). *Capitalism, Socialism and Democracy*. Londres, UK: Routledge. (publicado originalmente em 1942).

Uebel, T. (2008). Calculation in kind and marketless socialism: On Otto Neurath's utopian economics. *European Journal of the History of Economic Thought*, 15: 3, 475-501.



CEPS FCT
Centro de Ética, Política e Sociedade
Centre for Ethics, Politics and Society
Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÃO

COMPETE
PROGRAMA OPERACIONAL FACTORES DE COMPETITIVIDADE

QREN
QUADRO DE REFERÊNCIA
ESTRATÉGICO
NACIONAL
PORTUGAL 2007-2013

 **UNIÃO EUROPEIA**
Fundo Europeu de
Desenvolvimento Regional